

ALMA MATER STUDIORUM · UNIVERSITÀ DI
BOLOGNA

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

QUANTIZZAZIONI
INEQUIVALENTI:
gli effetti della topologia in meccanica
quantistica

Relatore:
Prof.ssa Elisa Ercolessi

Presentata da:
Luca Lumia

Anno Accademico 2017/2018

Sommario

Si discutono i problemi legati alla quantizzazione di un sistema il cui spazio delle configurazioni non è semplicemente connesso. Dopo aver introdotto alcuni concetti preliminari di topologia algebrica, viene descritta la procedura per portare a termine tale quantizzazione, mostrando che, in generale, si ottengono funzioni d'onda polidrome e che allo stesso sistema classico possono corrispondere più quantizzazioni inequivalenti. Infine si mostrano alcune applicazioni, fra cui il caso importante dei sistemi di particelle identiche.

Indice

1	Elementi di topologia algebrica	6
1.1	Cammini e omotopia	6
1.2	Il gruppo fondamentale	10
1.3	I rivestimenti di uno spazio topologico	18
1.4	Fibrati e rivestimento universale	25
2	Spazi multiplamente connessi in meccanica quantistica	35
2.1	Stati quantistici e fasi	36
2.2	L'effetto Aharonov-Bohm	38
2.3	Quantizzazione su spazi non semplicemente connessi	43
2.4	La particella singola sulla circonferenza	51
3	Particelle identiche	57
3.1	Lo spazio delle configurazioni per particelle indistinguibili . . .	57
3.2	Le statistiche su \mathbb{R}^d : bosoni, fermioni e anioni	60
3.3	Le particelle identiche sulla circonferenza	67
	Bibliografia	79

Introduzione

In fisica classica, lo stato dinamico di una particella in un certo istante t è identificato dalla coppia $(\vec{q}(t), \dot{\vec{q}}(t))$ della sua posizione e della sua velocità in quell'istante. Questo stato classico ha una chiara interpretazione fisica e l'oggetto matematico che usiamo per rappresentarlo è in corrispondenza biunivoca con esso: ad una coppia $(\vec{q}(t), \dot{\vec{q}}(t))$ corrisponde un unico stato di un sistema classico e viceversa. Nella meccanica quantistica questa corrispondenza non è così banale. Lo stato è rappresentato da una funzione d'onda ψ , generalmente definita sullo spazio delle configurazioni del sistema e a valori complessi, ma la funzione d'onda non ha un significato fisico immediato come $(\vec{q}(t), \dot{\vec{q}}(t))$. Tipicamente si dà significato solo al suo modulo quadro, dicendo che $|\psi(\vec{q})|^2$ sia la densità di probabilità con cui si distribuisce la probabilità di localizzare la particella intorno a \vec{q} eseguendo una misura di posizione. Questo spiega la condizione di normalizzazione che viene solitamente richiesta: le funzioni d'onda che descrivono stati misurabili dovranno avere integrale del loro modulo quadrato pari ad uno. Nella meccanica quantistica, le osservabili vengono descritte da operatori autoaggiunti che agiscono sulle funzioni d'onda, i cui autovalori corrispondono ai possibili risultati della misura e anche i valori attesi delle misure in un certo stato ψ dipendono solo da $|\psi(\vec{q})|^2$. Per questo, di solito, si dice che i risultati di una misura di un'osservabile dipendano solo dal modulo della funzione d'onda e che funzioni d'onda differenti per una fase rappresentino in realtà lo stesso stato. Chiaramente questo è corretto, ma ciò non giustifica un "declassamento" dell'importanza della fase di una funzione d'onda ψ .

La fase di una funzione d'onda (come la funzione d'onda in generale) non è un'osservabile in senso stretto, non è descritta da nessun operatore lineare, ma è comunque osservabile nel senso che sono misurabili effetti dipendenti da essa. La fase inosservabile è infatti una fase globale che moltiplica lo stato, mentre, se si considerano due funzioni d'onda in sovrapposizione, è chiaro che le fasi relative siano visibili, perché da esse dipendono gli effetti di interferenza. Un caso in cui è fondamentale determinare la fase delle funzioni d'onda è quando lo spazio delle configurazioni del sistema non è semplice-

mente connesso, cioè quando esistono dei "buchi" che, dati due punti nello stato, rendono impossibile deformare con continuità l'uno nell'altro tutti i cammini che collegano i due punti senza uscire dallo spazio. Se lo spazio delle configurazioni non è semplicemente connesso si trovano funzioni d'onda con certe ambiguità e la quantizzazione del sistema non è immediata. Per studiare come sia possibile formulare la meccanica quantistica su questi spazi è necessario introdurre preliminarmente alcuni concetti di topologia algebrica, tramite i quali è possibile definire in modo preciso cosa si intenda per connessione semplice e descrivere "in che modo" certi spazi topologici siano non semplicemente connessi. Questo sarà fatto all'inizio del primo capitolo, che si concluderà con una descrizione del rivestimento universale di uno spazio topologico: nei casi di interesse fisico, è sempre possibile, appunto, "rivestire" uno spazio Q non semplicemente connesso con uno spazio \tilde{Q} che sia semplicemente connesso e che si proietti bene sullo spazio sottostante Q . La quantizzazione andrà eseguita passando per questo rivestimento, siccome è semplicemente connesso e farlo permette di evitare i problemi sopra citati.

Prima di spiegare la procedura di quantizzazione, nel secondo capitolo, sarà richiamata più in dettaglio la corrispondenza tra funzioni d'onda e stati in meccanica quantistica. Successivamente si discuterà il famoso effetto Aharonov - Bohm, che ha una grande importanza concettuale nella fisica. L'effetto venne presentato dagli autori in un articolo del '59 [1], in cui si discuteva del significato dei potenziali elettromagnetici nella meccanica quantistica. Gli autori mostrarono che, contrariamente a quanto accade nella fisica classica, è possibile avere effetti elettromagnetici anche ove i campi di forza sono nulli, purché siano non nulli i potenziali e questo dimostra che i potenziali elettrico e magnetico non siano puri strumenti matematici, ma oggetti dotati di significato fisico di per sé. Questo effetto si può vedere come legato alla non semplice connessione dello spazio delle configurazioni e verrà usato come esempio per introdurre alcune idee che saranno sviluppate nel seguito. Il secondo capitolo prosegue con la descrizione della procedura di quantizzazione su spazi non semplicemente connessi, che viene eseguita sul rivestimento universale dello spazio delle configurazioni e successivamente riproiettata sullo spazio fisico sottostante. Dovranno essere imposte opportune condizioni di proiettabilità sulle funzioni d'onda, al fine di ottenere una meccanica quantistica ragionevole sullo spazio di partenza e queste condizioni porteranno a funzioni d'onda polidrome. La possibilità di avere funzioni d'onda a più valori era già stata discussa sin dai primi tempi di sviluppo della meccanica quantistica (si veda, ad esempio, l'articolo di Pauli del '39 [15]) e si giunse alla conclusione che, appunto, esse sono fisicamente accettabili quando lo spazio delle configurazioni non è semplicemente connesso. Per lavorare con esse ci si deve restringere ai loro rami, in cui sono discontinue; vedremo che la

discontinuità impedisce di rimuovere la fase della funzione d'onda senza modificare caratteristiche misurabili del sistema, come ad esempio il suo spettro energetico.

Nell'ultima sezione del secondo capitolo e per tutto il terzo capitolo si presentano esempi della teoria esposta. Il primo esempio è il caso di una particella singola su una circonferenza, poi si descrivono i sistemi di particelle identiche in \mathbb{R}^n . Quello delle particelle identiche è un caso di particolare importanza in cui non è banale ottenere il corretto spazio delle configurazioni, il quale può non essere semplicemente connesso. Dall'applicazione della teoria vista si vede come emergano in modo naturale i concetti ben noti di bosone e fermione e si vedrà che, in due dimensioni, esiste una classe più generica di particelle chiamate anioni, le cui caratteristiche interpolano quelle di bosoni e fermioni. Il capitolo si conclude con la descrizione del caso delle due particelle identiche su una circonferenza.

Capitolo 1

Elementi di topologia algebrica

Come si è detto, per ottenere le informazioni necessarie per trattare fisicamente certi spazi è necessario introdurre alcuni concetti di topologia algebrica. In generale, la topologia algebrica si occupa di costruire, dato uno spazio topologico, immagini di esso su strutture algebriche (principalmente gruppi), tramite cui è possibile avere una descrizione dello spazio. Nel seguito, ci si concentrerà sulla teoria dell'omotopia (in particolare sul primo gruppo di omotopia degli spazi topologici) e sulla teoria dei rivestimenti. Si presenterà una introduzione ai suddetti concetti, concentrandosi sugli aspetti di più immediata applicazione fisica e sui risultati formali che servono per descriverli con sufficiente accuratezza. La prima sezione §1.1 introdurrà il concetto di omotopia, che formalizza la relazione tra cammini che possono essere deformati con continuità l'uno nell'altro, mentre la §1.2 continua la discussione della teoria dell'omotopia descrivendo le proprietà del gruppo fondamentale, il quale permette di classificare i cammini omotopi in un certo spazio. Le sezioni §1.3 e §1.4 sono una introduzione alla teoria dei rivestimenti, per cui bisogna passare per costruire la meccanica quantistica se lo spazio delle configurazioni non è semplicemente connesso.

1.1 Cammini e omotopia

Sia X uno spazio topologico, che assumiamo per semplicità connesso. Il primo gruppo di omotopia sarà definito tramite i loop di X e per arrivare ad esso è utile introdurre in generale una definizione dei cammini nello spazio, che, intuitivamente, si possono vedere come delle traiettoria $\gamma(t)$ nello spazio X che vengono percorse nel tempo, rappresentato dalla variabile t in un certo intervallo reale.

Definizione 1.1. Si dice **cammino** (oppure percorso) nello spazio X una applicazione continua γ dall'intervallo reale $[0, 1]$ allo spazio topologico:

$$\gamma : [0, 1] \rightarrow X. \quad (1.1.1)$$

Il cammino è chiamato *loop* (o anche cappio, laccio...) in x se $\gamma(0) = \gamma(1) = x$; il punto x è detto *punto base* del loop.

Ricordiamo che la continuità di γ è definita secondo le topologie di $[0, 1]$ e di X : la controimmagine di ogni aperto di X tramite γ dev'essere un aperto di $[0, 1]$. Si sottolinea che un cammino è individuato sia dai punti geometrici in X che vengono percorsi $\gamma([0, 1])$ (il *supporto* del cammino) sia da come vengono percorsi, cioè dalle "velocità" lungo il percorso. Un loop corrisponde ad una traiettoria che ritorna al punto di partenza, ed è, infatti, chiamato anche cammino chiuso. L'ipotesi che γ sia definita su $[0, 1]$ è solamente di comodo e la definizione si potrebbe dare su un arbitrario intervallo reale.

È possibile definire un concetto di traiettoria puramente geometrica "rimuovendo" le informazioni dinamiche su come venga percorsa. Per farlo si definisce una opportuna relazione di equivalenza.

Definizione 1.2. Un **cambiamento di parametrizzazione** è una funzione $\varphi : [0, 1] \rightarrow [0, 1]$ tale che sia continua, suriettiva e strettamente crescente (quindi chiaramente φ è invertibile).

Definizione 1.3. Definiamo **equivalenti** due cammini $\gamma(t)$ e $\tilde{\gamma}(s)$ che sono legati da un cambiamento di parametrizzazione, cioè per cui esiste un cambiamento di parametrizzazione φ tale che

$$\gamma(t) = \tilde{\gamma}(\varphi(t)) \quad \forall t \in [0, 1]. \quad (1.1.2)$$

Si può dimostrare che quella così definita è effettivamente una relazione di equivalenza. I cammini equivalenti hanno perso le informazioni su come la traiettoria è percorsa, siccome il cambiamento di parametrizzazione permette di percorrerla in ogni modo (sufficientemente regolare) possibile, purché sia sufficientemente regolare e non inverta l'ordine di percorrenza (φ è richiesta crescente); ciò che rimane ad accomunare gli elementi di una classe di equivalenza sono i punti di partenza e di arrivo e il supporto, dato che, se è soddisfatta la (1.1.2), $\gamma([0, 1]) = \tilde{\gamma}([0, 1])$. Gli elementi del quoziente dell'insieme dei cammini su X rispetto alla suddetta relazione di equivalenza sono le *curve* su X .

Un utile strumento per relazionare diversi cammini è il concetto di *omotopia*, che formalizza la possibilità di deformare in modo continuo un cammino in un altro.

Definizione 1.4. Si definisce **omotopia** una mappa $h : [0, 1] \times [0, 1] \rightarrow X$, $(\lambda, t) \mapsto h_\lambda(t)$ che sia continua e che per ogni λ definisca una famiglia $h_\lambda(t)$ di cammini in X .

Definizione 1.5. Siano γ e δ due cammini in X ; γ è detto **omotopo** a δ se esiste una omotopia h che li collega, cioè se

$$h_0(t) = \gamma(t) \quad (1.1.3a)$$

$$h_1(t) = \delta(t) \quad (1.1.3b)$$

per ogni $t \in [0, 1]$.

L'omotopia h , dunque, definisce una famiglia di cammini "intermedi" che realizzano la deformazione continua tra due cammini γ e δ omotopi. Quella così definita è anche detta *omotopia libera* perché non c'è alcun vincolo sui punti estremi $h_\lambda(0)$, $h_\lambda(1)$; è possibile dare una definizione più restrittiva (*omotopia con estremi fissati*), richiedendo che i punti iniziali e finali siano indipendenti dal parametro λ , cioè

$$h_\lambda(0) = x_0 \text{ e } h_\lambda(1) = x_1 \quad \forall \lambda \in [0, 1]. \quad (1.1.4)$$

I cammini della famiglia $h_\lambda(t)$ che lega due percorsi omotopi e che soddisfanno questa ulteriore condizione evidentemente non invertono mai il verso di percorrenza.

Interessa lavorare perlopiù con cammini γ chiusi, per i quali vale $\gamma(0) = \gamma(1) = x$, dove x è il punto base di γ ; unendo questa richiesta alla 1.1.4 ottengo una definizione di un tipo di omotopia più forte, tipico dei loop.

Definizione 1.6. Siano γ e δ due loop in \mathbb{X} con punto base x ; γ si dice **omotopo con punto base x** (o, più brevemente, omotopi in x) a δ se esiste un'omotopia libera $h : [0, 1] \times [0, 1] \rightarrow X$ che li collega tale che

$$h_\lambda(0) = h_\lambda(1) = x \quad \forall \lambda \in [0, 1], \quad (1.1.5)$$

cioè tutti i cammini della famiglia $h_\lambda(t)$ devono essere anch'essi loop con punto base x .

La Figura 1.1 mostra un esempio grafico di loop omotopi con punto base. Grazie alle definizioni 1.5 e 1.6 è possibile definire la *relazione di omotopia* \sim tra due cammini: $\gamma \sim \delta \Leftrightarrow \gamma$ è omotopo a δ . Analogamente si definisce la relazione di omotopia con punto base tra due loop e si indica allo stesso modo. È di grande importanza il seguente risultato:

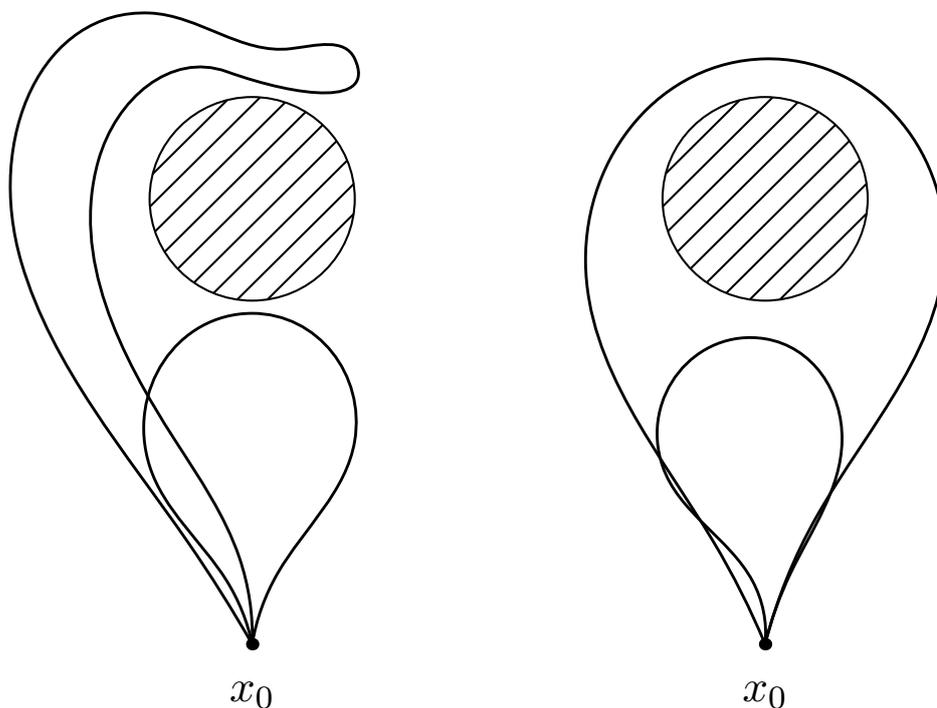


Figura 1.1: A sinistra, si ha un esempio di coppia di loop omotopi con punto base x_0 sullo spazio topologico rappresentato dal piano bucato dal cerchio evidenziato; a destra sono mostrati due loop non omotopi.

Teorema 1.1. *La relazione di omotopia (libera o con punto base) tra due cammini è una relazione di equivalenza.*

Dimostrazione. Sono da verificare le proprietà riflessiva, simmetrica e transitiva. La riflessiva è banale, perché chiaramente ogni cammino è omotopo a se stesso tramite $h = id_{[0,1] \times [0,1]}$; anche la verifica della proprietà simmetrica è immediata perché se γ è omotopo a δ secondo la famiglia di curve $h_\lambda(t)$, si ha un'omotopia di δ in γ realizzata da $h_{1-\lambda}(t)$, che è ancora continua (perché composizione di funzioni continue) e $1 - \lambda \in [0, 1]$. Infine, la transitività segue dal fatto che se γ è omotopo a δ tramite $h_\lambda(t)$ e δ è omotopo a ϵ tramite $\tilde{h}_\lambda(t)$, allora la famiglia

$$H_\lambda(t) = \begin{cases} h_{2\lambda}(t) & \text{se } 0 \leq \lambda < 1/2 \\ \tilde{h}_{2\lambda-1}(t) & \text{se } 1/2 \leq \lambda \leq 1 \end{cases} \quad (1.1.6)$$

realizza un'omotopia tra γ ed ϵ . La continuità di $H_\lambda(t)$ segue dalla ovvia continuità di $h_{2\lambda}(t)$ e di $\tilde{h}_{2\lambda-1}(t)$ sui loro domini di definizione nella (1.1.6) e

dalla continuità nel punto di raccordo $\lambda = 1/2$, in cui si riducono a $h_1(t) = \tilde{h}_0(t) = \delta(t)$. \square

Prima di procedere oltre, mostriamo un utile esempio di omotopia. Si considerino due cammini γ e δ con gli stessi punti iniziale e finale su un insieme convesso $X \subseteq \mathbb{R}^n$; tutti i cammini γ e δ di questo tipo sono (liberamente) omotopi in modo da soddisfare anche la richiesta (1.1.4). L'omotopia che lega i due cammini è individuata dalla combinazione lineare convessa

$$h_\lambda(t) = (1 - \lambda)\gamma(t) + \lambda\delta(t),$$

che definisce una famiglia di percorsi che si deforma con "velocità" costante da γ a δ , siccome per ogni $\bar{t} \in [0, 1]$, parametrizza il segmento da $\gamma(\bar{t})$ a $\delta(\bar{t})$, ed è chiaramente un'applicazione continua in entrambi gli argomenti. La famiglia $h_\lambda(t)$ è ben definita in X grazie alla richiesta di convessità, che assicura che tutti i segmenti $h_\lambda(\bar{t}) \forall \bar{t}$ siano interamente contenuti in X . La (1.1.4) è soddisfatta per via dell'ipotesi che fissa i punti iniziale e finale dei due cammini, in modo che i due segmenti $h_\lambda(0)$ e $h_\lambda(1)$ degenerino mappando tutti i λ nei due punti fissati. Specificando questo ragionamento per i cammini chiusi, trovo che tutti i loop con stesso punto base x_0 in uno spazio convesso sono omotopi in x_0 . Con un analogo ragionamento si dimostra anche che due cammini in un qualsiasi spazio X che differiscono solo per la parametrizzazione (cioè che corrispondono alla stessa curva) sono omotopi.

1.2 Il gruppo fondamentale

Le famiglie di curve omotope formano classi di equivalenza $[\gamma]$ dette *classi di omotopia*, indicate tramite un rappresentate della famiglia. Si considerino ora in particolare le classi di loop omotopi con punto base x_0 ; l'insieme di queste classi di equivalenza è il quoziente dell'insieme dei loop in $x_0 \in X$ rispetto alla relazione di omotopia con punto base e si vedrà che ha una naturale struttura di gruppo.

Definizione 1.7. L'insieme delle classi di omotopia con punto base $[\gamma]$ dei cammini chiusi $\gamma : [0, 1] \rightarrow X$ in $x_0 \in \mathbb{X}$ si indica con $\pi_1(\mathbb{X}, x_0)$ ed è chiamato **gruppo fondamentale in x_0** (o primo gruppo di omotopia in x_0) di X .

Questo gruppo è una delle immagini dello spazio X che la topologia algebrica costruisce ed è possibile grazie ad esso descrivere alcune importanti proprietà di \mathbb{X} che saranno utili nel seguito. È chiamato anche *primo* gruppo di omotopia perché è possibile costruire una sequenza di altri gruppi di omotopia $\pi_n(X, x_0)$, utilizzando al posto dei cammini $\gamma : [0, 1] \rightarrow X$ delle mappe $[0, 1]^n \rightarrow X$ (con in più tutte le condizioni del caso).

Per costruire la struttura di gruppo di $\pi_1(X, x_0)$ si definisce prima una operazione di composizione dei cammini.

Definizione 1.8. Siano γ, δ due cammini su X con un estremo in comune $\gamma(1) = \delta(0)$. Definisco il **prodotto** dei due cammini come:

$$\gamma \cdot \delta = \begin{cases} \gamma(2t) & \text{se } 0 \leq t < 1/2 \\ \delta(2t - 1) & \text{se } 1/2 \leq t \leq 1 \end{cases} . \quad (1.2.1)$$

Questa definizione corrisponde a percorrere in sequenza in due cammini, ciascuno a velocità doppia. La continuità della mappa prodotto è garantita dalla richiesta di avere un estremo in comune.

Assumiamo ora che γ e δ siano due loop con lo stesso punto base x_0 e consideriamo la relazione di omotopia con punto base; in questo caso la richiesta di continuità è naturalmente soddisfatta. Sullo spazio dei loop con punto base fissato non si riesce ad introdurre la struttura di gruppo con questa operazione perché, sebbene presenti naturali candidati al ruolo di loop inverso e loop neutro, essi non hanno buone proprietà e inoltre l'operazione risulta non associativa. I candidati sarebbero: $\gamma^{-1}(t) = \gamma(1 - t)$ (loop percorso in verso opposto) per l'elemento inverso e $\bar{e}(t) = x_0 \forall t$ (loop degenerare nel punto base) per l'elemento neutro.

I problemi sono i seguenti:

- componendo $\bar{e} \cdot \gamma$ non riottengo γ ma un cammino con lo stesso sostegno fermo in x_0 fino $t = 1/2$ e che percorre il resto a velocità doppia da $t = 1/2$ a $t = 1$;
- componendo $\gamma \cdot \gamma^{-1}$ non risulta \bar{e} ma si ottiene un percorso chiuso che ha un sostegno diverso dall'unico punto x_0 ;
- l'operazione non è associativa perché $(\gamma \cdot \delta) \cdot \epsilon$ e $\gamma \cdot (\delta \cdot \epsilon)$ sono parametrizzati diversamente, in modo simile al primo punto.

Si può passare oltre a questi problemi notando che, anche se $\gamma \cdot \gamma^{-1} \neq \bar{e}$, si ha $\gamma \cdot \gamma^{-1} \sim \bar{e}$ l'omotopia è mostrata nella Figura 1.2 ed è possibile grazie al fatto il passaggio intermedio per il punto base x_0 avviene per $t = 1/2$: non è un estremo ed è possibile spostarlo, non è più un punto fissato. Per quanto riguarda gli altri due punti, i risultati differiscono solo per un cambiamento di parametrizzazione da quelli cercati; allora, come si è detto, $\bar{e} \cdot \gamma \sim \gamma$ e $(\gamma \cdot \delta) \cdot \epsilon \sim \gamma \cdot (\delta \cdot \epsilon)$. Il prodotto definito definisce una struttura di gruppo a meno di omotopie, per cui sarà possibile dotare $\pi_1(X, x_0)$ delle proprietà di gruppo, siccome su $\pi_1(X, x_0)$ cammini differenti solo per omotopie sono identificati.

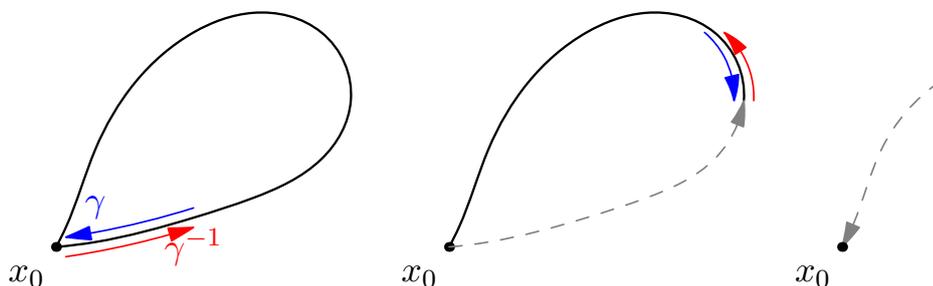


Figura 1.2: Rappresentazione dell'omotopia che rende equivalente ogni loop composto al suo inverso al loop banale.

Gli elementi di $\pi_1(X, x_0)$ sono classi di equivalenza di cammini chiusi; definiamo ora una operazione di composizione interna al gruppo fondamentale che moltiplichi le classi di omotopia.

Definizione 1.9. Siano $[\gamma], [\delta] \in \pi_1(X, x_0)$; definisco il **prodotto** delle due classi di omotopia come

$$[\gamma] \cdot [\delta] = [\gamma \cdot \delta], \quad (1.2.2)$$

cioè il prodotto di due classi è la classe di omotopia che ha per rappresentate il prodotto dei cammini rappresentanti delle due classi moltiplicate.

Per verificare che la definizione sia ben posta sulle classi di equivalenza, bisogna mostrare che è effettivamente indipendente dalla scelta del rappresentate. Siano allora i loop γ_1, γ_2 tali che $\gamma_1 \sim \gamma_2$ tramite l'omotopia $h_\lambda(t)$ e δ_1, δ_2 tali che $\delta_1 \sim \delta_2$ tramite l'omotopia $k_\lambda(t)$; allora $\gamma_1 \cdot \delta_1 \sim \gamma_2 \cdot \delta_2$. La famiglia di cammini chiusi che realizza questa omotopia è definita dal prodotto $h_\lambda(t) \cdot k_\lambda(t)$, che è continuo in entrambi gli argomenti e, per $\lambda = 0$, dà $h_0(t) \cdot k_0(t) = \gamma_1 \cdot \delta_1$ e, per $\lambda = 1$, $h_1(t) \cdot k_1(t) = \gamma_2 \cdot \delta_2$.

Teorema 1.2. $\pi_1(X, x_0)$ è un gruppo rispetto al prodotto di classi di omotopia $[\gamma] \cdot [\delta] = [\gamma \cdot \delta]$.

Dimostrazione. Si è appena visto che l'operazione è ben definita sull'insieme, dunque bisogna verificare solo l'esistenza dell'inverso per ogni classe, dell'elemento neutro e dell'associatività. L'**elemento neutro** è

$$e = [\bar{e}], \quad \text{dove } \bar{e}(t) = x_0 \quad \forall t \in [0, 1] \quad (1.2.3)$$

e lo è perché, come già osservato in precedenza, $\bar{e} \cdot \gamma, \gamma \cdot \bar{e}$ ed \bar{e} differiscono solo per la parametrizzazione, per cui sono omotopi. Per la stessa ragione l'operazione è associativa. Data una classe $[\gamma]$, il suo **inverso** è definito come

$$[\gamma]^{-1} = [\gamma^{-1}], \quad \text{dove } \gamma^{-1}(t) = \gamma(1 - t), \quad (1.2.4)$$

cioè è la classe associata al cammino percorso in verso opposto. Come si era mostrato intuitivamente nella figura 1.2, $\gamma^{-1} \cdot \gamma$ è omotopo ad \bar{e} e $[\gamma^{-1} \cdot \gamma] = e$. La famiglia cammini che realizza questa omotopia è $H_\lambda(t) = h_\lambda(t) \cdot h_\lambda^{-1}(t)$, dove $h_\lambda(t)$ è il percorso uguale a $\gamma(t)$ per $t \in [0, 1 - \lambda]$ ed è stazionario in $\gamma(1 - \lambda)$ se $t \in]1 - \lambda, 1]$ e l'altro è il suo percorso inverso. Per $\lambda = 0$ si ha $h_0(t) = \gamma(t) \forall t$, dunque $H_0(t)$ coincide con γ percorso una volta e poi in verso opposto, e per $\lambda = 1$ $h_1(t) = \gamma(t)$ per $t = 0$ e $h_1(t) = \gamma(1 - 1)$ per $t \in]0, 1]$, per cui $h_1(t) = x_0 \forall t$ e così anche $H_1 = \bar{e}$; infine, la famiglia è evidentemente continua e si è così verificata l'omotopia. \square

Fino ad ora il gruppo fondamentale è stato definito solo rispetto ad un particolare punto base x_0 fissato; è naturale chiedersi se esista una relazione tra gruppi fondamentali definiti su punti diversi dello spazio topologico. Siccome $\pi_1(X, x_0)$ è definito tramite i cammini in x_0 , è chiaro che esso sia legato solo alla componente di X connessa per archi con x_0 ; allora, per rispondere a questa domanda, supponiamo d'ora in avanti che lo spazio topologico X sia connesso per archi. Questa ipotesi non è particolarmente restrittiva, considerandola dal punto di vista delle applicazioni fisiche, in cui tipicamente X è lo spazio delle configurazioni di un qualche sistema.

Siano due punti distinti $x, y \in X$. Vedremo che è possibile costruire in modo naturale un *isomorfismo di cammini* tra $\pi_1(X, x)$ e $\pi_1(X, y)$, per cui il gruppo fondamentale di uno spazio connesso per cammini è sostanzialmente unico. Si consideri un loop γ in y e sia ρ un cammino che va da x a y (che esiste per l'ipotesi di connessione per archi), allora $\rho\gamma\rho^{-1}$ è un loop in x .¹ Questo fatto è evidente ed è mostrato nella Figura 1.3. In questo modo associo a ogni loop in y un loop in x e posso allora definire una corrispondenza tra le classi di omotopia in $\pi_1(X, y)$ e le classi in $\pi_1(X, x)$.

Definizione 1.10. La corrispondenza tra le classi di omotopia in y e le classi in x è realizzata dalla mappa

$$\begin{aligned} \phi_\rho : \pi_1(X, y) &\rightarrow \pi_1(X, x) \\ [\gamma] &\longmapsto \phi_\rho([\gamma]) = [\rho\gamma\rho^{-1}], \end{aligned} \tag{1.2.5}$$

dove ρ è un arbitrario cammino che va da x a y .

Per verificare che sia effettivamente una mappa ben definita tra le classi, bisogna mostrare che $[\rho\gamma\rho^{-1}]$ sia sempre la stessa classe di omotopia in x al variare della scelta del rappresentate γ di $[\gamma] \in \pi_1(X, y)$. Sia γ' un loop in y

¹Il prodotto di cammini non è associativo e andrebbe specificato un ordinamento $(\rho\gamma)\rho^{-1}$ o $\rho(\gamma\rho^{-1})$, ma non è necessario farlo perché in questo contesto si è interessati solo alle classi di omotopia.

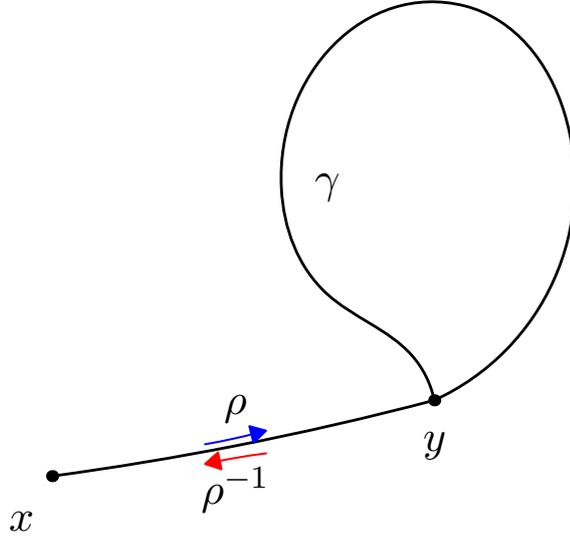


Figura 1.3: Costruzione della corrispondenza tra loop in y e loop in x .

tale che $\gamma \sim \gamma'$ tramite la famiglia di cammini h_λ . Allora i cammini $\rho h_\lambda \rho^{-1}$ realizzano una omotopia tra $\rho\gamma\rho^{-1}$ e $\rho\gamma'\rho^{-1}$, come si può immediatamente verificare; dunque la mappa (1.2.5) è ben definita. Possiamo enunciare ora il risultato cercato.

Teorema 1.3. *La mappa $\phi_\rho : \pi_1(X, y) \rightarrow \pi_1(X, x)$ è un isomorfismo tra i gruppi fondamentali in x e in $y \forall x, y \in X$ se X connesso per archi.*

Corollario. *Se X è uno spazio topologico connesso per archi, allora esiste un unico gruppo astratto associato ai gruppi fondamentali $\pi_1(X, x)$ nei suoi punti $x \in X$.*

Dimostrazione. Per verificare che ϕ_ρ sia un isomorfismo tra i due gruppi fondamentali, devo mostrare che sia un omomorfismo e che sia biunivoca. È un omomorfismo perché preserva le operazioni dei due gruppi:

$$\phi_\rho([\gamma][\delta]) = \phi_\rho([\gamma\delta]) = [\rho\gamma\delta\rho^{-1}] = [\rho\gamma\rho^{-1}][\rho\delta\rho^{-1}],$$

dove l'ultima uguaglianza è giustificata dal fatto che

$$(\rho\gamma\rho^{-1})(\rho\delta\rho^{-1}) \sim \rho\gamma(\rho^{-1}\rho)\delta\rho^{-1} \sim \rho(\gamma\delta)\rho^{-1}.$$

La mappa ϕ_ρ è invertibile con inversa $\phi_{\rho^{-1}}$, perché per ogni γ loop in y

$$\phi_\rho \circ \phi_{\rho^{-1}}([\gamma]) = [\rho\rho^{-1}\gamma\rho\rho^{-1}] = [\gamma],$$

$$\phi_{\rho^{-1}} \circ \phi_\rho([\gamma]) = [\rho^{-1}\rho\gamma\rho^{-1}\rho] = [\gamma].$$

Il corollario è banale dalla definizione di gruppo astratto come classe di equivalenza di gruppi isomorfi. \square

Definizione 1.11. Dato uno spazio topologico X connesso per archi, si definisce **gruppo fondamentale** di X e si indica con $\pi_1(X)$ l'unico gruppo astratto corrispondente ai gruppi fondamentali nei suoi punti $\pi_1(X, x) \forall x \in X$.

L'isomorfismo ϕ_ρ sopra definito non è in generale canonico, nel senso che la corrispondenza che realizza può dipendere dalla scelta del cammino ρ che collega x a y : ϕ_ρ si dice canonica se dati due percorsi ρ_1, ρ_2 da x a y , per ogni $[\gamma] \in \pi_1(X, y)$ $\phi_{\rho_1}([\gamma]) = \phi_{\rho_2}([\gamma])$. Si può dimostrare [11] che vale il seguente risultato:

Teorema 1.4. *L'isomorfismo $\phi_\rho : \pi_1(X, y) \rightarrow \pi_1(X, x)$ è canonico se e solo se il gruppo fondamentale $\pi_1(X)$ è abeliano.*

Nota. Si ricordi che, essendo isomorfi, $\pi_1(X, y)$ è abeliano se e solo se $\pi_1(X, x)$ lo è (e quindi anche il gruppo astratto associato $\pi_1(X)$).

Consideriamo ora un esempio che sarà utile nel seguito: il gruppo fondamentale della circonferenza $X = S^1$ è isomorfo al gruppo ciclico infinito: $\pi_1(S^1) = \mathbb{Z}$. La dimostrazione completa è laboriosa, e per questo sarà posticipata fino alla sezione 1.4 (in cui sarà accennata), ma il risultato in sé è estremamente intuitivo. Sia $x_0 \in S^1$ un qualsiasi punto della circonferenza e si considerino i percorsi chiusi con x_0 come punto base. A ciascun loop si può associare il numero di giri che il esso compie lungo la circonferenza, contando, per convenzione, con segno positivo i giri in verso antiorario e con segno negativo i giri in verso orario; questo numero è chiamato *indice di avvolgimento* (o, in inglese, "winding number") del loop ed è, chiaramente, un numero intero. È nell'esperienza di tutti che, dato un laccio chiuso appoggiato su una circonferenza, non sia possibile cambiare il numero di giri che questo laccio fa sulla circonferenza a meno di farlo uscire dalla circonferenza svincolandolo da essa, cioè sollevandolo o attraversando il buco. Questo si traduce matematicamente nel seguente fatto intuitivo:

è possibile deformare con continuità un cammino in un altro se e solo se essi hanno lo stesso indice di avvolgimento,

perché il numero di giri compiuti è un intero, non può variare in maniera continua, e dunque un'omotopia può collegare solo percorsi con stesso indice di avvolgimento. Per cambiare il numero di giri, la trasformazione dovrebbe essere discontinua o "attraversare" il buco della circonferenza ed entrambe le situazioni non sono lecite. Così, dato che chiaramente il loop banale $\bar{e}(t) = x_0$ ha indice di avvolgimento 0, la classe di loop *nullomotopi* (cioè omotopi ad un punto, che è il loop banale) $e = [e]$ è formata da tutti i loop che non compiono giri intorno alla circonferenza. In generale, due loop sono nella

stessa classe di omotopia se e solo se hanno lo stesso indice di avvolgimento, per cui posso enumerare in modo biunivoco tutte le classi di omotopia con un numero intero; questo suggerisce l'identificazione $\pi_1(S^1) = \mathbb{Z}$. Moltiplicando due loop, essi vengono percorsi entrambi in successione, per cui è chiaro che il giri percorsi si sommano e il prodotto dei loop avrà indice di avvolgimento pari alla somma di quelli dei due loop, cioè, chiamando $n([\gamma]) \in \mathbb{Z}$ l'indice di avvolgimento di $[\gamma] \in \pi_1(S^1, x_0)$, si ha

$$n([\gamma][\delta]) = n([\gamma]) + n([\delta]) \quad \forall [\gamma], [\delta] \in \pi_1(S^1, x_0).$$

Questo vuol dire che il morfismo $\pi_1(S^1, x_0) \rightarrow \mathbb{Z}$, $[\gamma] \mapsto n([\gamma]) \in \mathbb{Z}$ è un isomorfismo tra i due gruppi e \mathbb{Z} è effettivamente il gruppo fondamentale della circonferenza. Un ragionamento del tutto analogo si sarebbe potuto ripetere considerando come spazio topologico il piano senza un punto -ad esempio l'origine- trovando che $\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z}$.

Entrambi gli esempi presentati hanno una caratteristica: sono spazi con un buco unidimensionale che rende impossibile deformare con continuità un qualsiasi loop in un punto, cioè il cui gruppo fondamentale contiene altri elementi oltre all'elemento neutro, corrispondente alla classe dei cammini omotopi a quello banale. Il gruppo fondamentale è l'immagine algebrica di X che permette di "catturare" questa proprietà dello spazio topologico.

Definizione 1.12. Uno spazio topologico X è detto **semplicemente connesso** se il suo gruppo fondamentale è banale: $\pi_1(X) = \{e\}$, cioè se ogni loop in X è omotopo al loop banale.

Dunque S^1 e $\mathbb{R}^2 \setminus \{0\}$ sono due esempi di spazi non semplicemente connessi, mentre, ogni insieme convesso è semplicemente connesso (si era osservato in precedenza che in uno spazio convesso esisteva una omotopia lineare che rendeva tutti i cammini chiusi in un punto omotopi, ed essendo il loop banale un cammino chiuso, esiste solo la classe di omotopia dell'elemento neutro). Uno spazio topologico non semplicemente connesso è anche detto *multiplamente connesso*. Si sottolinea che l'esistenza di un buco non è necessariamente determinate per la connessione semplice: S^n , e analogamente $\mathbb{R}^{n+1} \setminus \{0\}$, risultano semplicemente connessi per $n \geq 2$, perché ogni cammino chiuso può essere contratto con continuità in un punto "aggirando" il buco, come si può vedere intuitivamente con S^2 e $\mathbb{R}^3 \setminus \{0\}$; in questi due spazi è evidente che non sia invece possibile deformare una "superficie chiusa" (una mappa $\Sigma : [0, 1] \times [0, 1] \rightarrow X$ tale che $\Sigma(\partial([0, 1] \times [0, 1])) = x_0 \in X$) in un punto, se questa racchiude l'origine. Per descrivere con completezza i buchi degli spazi topologici servirebbero anche i gruppi di omotopia di ordine superiore (o i gruppi di omologia), ma qui ci si concentrerà solo sul gruppo

fondamentale perché è questo quello di maggiore utilità nella fisica, siccome permette di classificare le traiettorie delle particelle.

Gli spazi semplicemente connessi sono così chiamati perché vale il seguente risultato:

Teorema 1.5. *Uno spazio topologico X è semplicemente connesso se e solo se per ogni coppia di punti in X esiste un'unica classe di cammini omotopi (con estremi fissati) che li congiunge.*

Dimostrazione. Dati due punti in X , di sicuro esiste almeno un cammino $[\gamma]$ che li congiunge per l'ipotesi di connessione per archi. Supponendo ne esista un secondo chiamato $[\delta]$, siccome $\pi_1(X) = e$ tutti i loop sono banali e $[\delta^{-1}\delta] = [\gamma\delta^{-1}] = e$. Allora $[\gamma] = [\gamma\delta^{-1}\delta] = [\delta]$. Abbiamo dimostrato l'implicazione (\implies); l'inversa (\impliedby) segue banalmente dal fatto che, considerando i due punti coincidenti, se si ha un'unica classe di cammini che congiunge un punto con se stesso si ha un'unica classe di cammini chiusi nel punto e allora ogni loop è omotopo al loop banale. \square

Un risultato utile e molto semplice per calcolare gruppi fondamentali è il fatto che la mappa che ad uno spazio topologico associa il suo gruppo fondamentale si comporta bene rispetto al prodotto cartesiano, cioè vale il seguente risultato (la cui dimostrazione segue direttamente dalla definizione di topologia prodotto):

Teorema 1.6. *Se X e Y sono spazi topologici connessi per archi, allora vale la relazione*

$$\pi_1(X \times Y) = \pi_1(X) \times \pi_1(Y).$$

Ad esempio, grazie a questo teorema, possiamo affermare che il gruppo fondamentale del toro T^2 è $\pi_1(T^2) = \pi_1(S^1 \times S^1) = \mathbb{Z} \times \mathbb{Z}$.

Grazie all'omotopia è possibile costruire una relazione di equivalenza tra due spazi topologici, e non solo tra mappe su questi spazi, detta *equivalenza di omotopia*.

Definizione 1.13. Siano X, Y due spazi topologici; X e Y sono detti **omotopicamente equivalenti** se esistono due mappe $f : X \rightarrow Y$ e $g : Y \rightarrow X$ continue e tali che le loro composizioni sono omotope all'identità:

$$f \circ g \sim id_X \tag{1.2.6a}$$

$$g \circ f \sim id_Y. \tag{1.2.6b}$$

Le mappe f, g sono dette *equivalenze di omotopia* e sono l'una l'*inversa omotopica* dell'altra.

Ad esempio, \mathbb{S}^1 e $\mathbb{R}^2 \setminus \{0\}$ sono omotopicamente equivalenti. Qui in realtà si è parlato di omotopia non di cammini ma di mappe su uno spazio $\varphi : X \rightarrow X$; questa generica omotopia è definita il modo analogo:

Definizione 1.14. Le mappe $\varphi_1, \varphi_2 : X \rightarrow Y$ sono **omotope** se esiste una applicazione $F : [0, 1] \times X \rightarrow Y$, $(\lambda, x) \mapsto F_\lambda(x)$ continua in entrambi gli argomenti e tale che la famiglia indotta di mappe $F_\lambda : X \rightarrow Y$ soddisfi

$$F_0 = \varphi_1, \quad F_1 = \varphi_2.$$

Concludiamo la sezione con il seguente importante risultato:

Teorema 1.7. *Siano X, Y due spazi topologici. Se X e Y sono omotopicamente equivalenti, allora $\pi_1(X) = \pi_1(Y)$. Inoltre, due spazi omeomorfi sono anche omotopicamente equivalenti, per cui hanno gruppi fondamentali isomorfi.*

1.3 I rivestimenti di uno spazio topologico

La teoria dei rivestimenti è una parte importante della topologia algebrica; i rivestimenti sono uno strumento di notevole utilità per calcolare i gruppi fondamentali di vari spazi e non solo: si vedrà che è fondamentale per la formulazione della meccanica quantistica su spazi non semplicemente connessi. L'idea sarà quella di -appunto- "rivestire" uno spazio delle configurazioni multiplamente connesso con uno spazio semplicemente connesso e costruire su quest'ultimo una meccanica quantistica da proiettare sullo spazio delle configurazioni iniziale. Prima di arrivare al rivestimento universale, che sarà quello utilizzato in pratica, è necessaria un'introduzione generale ai rivestimenti. Per questa sezione non si ha pretesa di completezza e molti risultati saranno presentati senza le dimostrazioni, che si possono trovare su [6], [10] e [18].

Definizione 1.15. Siano due spazi topologici X e \tilde{X} e sia una mappa continua tra essi detta *proiezione di rivestimento* $p : \tilde{X} \rightarrow X$; la coppia (\tilde{X}, p) è detta un **rivestimento** di X se esiste un ricoprimento aperto $\{U_\alpha\}$ di X tale che, per ogni α , $p^{-1}(U_\alpha)$ è un'unione disgiunta di aperti di \tilde{X} , ciascuno dei quali è omeomorfo ad U_α restringendo p .

L'idea è che un rivestimento sia uno spazio topologico che possa essere visto localmente come una collezione di copie dello spazio che riveste: nella definizione si hanno più aperti di \tilde{X} che sono omeomorfi ognuno allo stesso

aperto di X . Nel seguito indicheremo sinteticamente un rivestimento come la sola proiezione o come il solo spazio, quando non c'è rischio di confusione.

Se consideriamo come esempio $X = S^1$, un suo rivestimento è dato dai reali con la proiezione $p : \mathbb{R} \rightarrow S^1$, $p(t) = (\cos(2\pi t), \sin(2\pi t))$, dove il ricoprimento U_α è formato da due qualunque archi aperti che uniti danno la circonferenza e i $p^{-1}(U_\alpha)$ sono unioni di segmenti in \mathbb{R} . La Figura 1.4 dà una rappresentazione intuitiva di questo caso.



Figura 1.4: La retta reale vista come rivestimento di S^1 dato dalla proiezione può essere visualizzata identificandola con un'elica che si avvolge all'infinito sulla circonferenza (immagine presa da [6]).

Un altro rivestimento si può scrivere considerando la circonferenza nel piano complesso come $S^1 = \{x \in \mathbb{C} \mid |x| = 1\}$: definendo adesso $p : S^1 \rightarrow S^1$, $p(z) = z^n \forall n \in \mathbb{N}$ si ha una mappa che riveste tutta la circonferenza unitaria n volte. Ad esempio, per $n = 2$, ogni punto x della circonferenza corrisponde alle sue due radici complessi $z = \sqrt{x}$: si ha un rivestimento doppio di S^1 .

Come si vede anche in questo caso, tipicamente *gli insiemi $p^{-1}(x)$ sono discreti*, nel senso che se, preso un qualsiasi $\tilde{x} \in p^{-1}(x)$, posso sempre trovare un aperto sufficientemente piccolo della topologia di \tilde{X} tale che esso contenga solo \tilde{x} e nessun altro punto di $p^{-1}(x)$, cioè l'intersezione dell'aperto e di $p^{-1}(x)$ è solo \tilde{x} : non ci sono punti di $p^{-1}(x)$ arbitrariamente vicini. L'idea è che, se U_α è l'aperto contenente x del ricoprimento che realizza la definizione, esso è mandato tramite p^{-1} in una famiglia di aperti disgiunti \tilde{U}_{α_j} di \tilde{X} su cui, per ciascuno di essi, la restrizione p_j^{-1} è biunivoca, dunque manda x in un solo $\tilde{x}_j \in U_{\alpha_j}$ e non ce ne sono altri perché sono aperti disgiunti. Da questo esempio si nota anche che la cardinalità degli insiemi $p^{-1}(x) = \sqrt{x}$ è costante

al variare di x , siccome tutti i complessi (0 a parte) hanno due radici; questa proprietà vale in generale e segue da:

se $p : \tilde{X} \rightarrow X$ è un rivestimento di X , gli insiemi $p^{-1}(x)$ sono omeomorfi al variare di x su ogni componente connessa di X .

X è detto *spazio base* e gli insiemi $p^{-1}(x)$ sono chiamati *fibre*, perché si vedrà che è possibile dotare i rivestimenti della struttura di fibrato, e la cardinalità delle fibre (che è localmente costante e costante se X è connesso, usando l'omeomorfismo come biiezione che "conta" gli elementi) è detta *numero di fogli*. Dagli esempi mostrati si evince che per uno stesso spazio topologico possono esistere più rivestimenti; il rivestimento universale, che è quello a cui si è interessati, riveste anche tutti gli altri rivestimenti dello spazio.

Nel trattare i rivestimenti è importante studiare le relazioni indotte dalla proiezione tra le mappe sullo spazio base e quelle sul rivestimento. A tale scopo si dà la seguente definizione.

Definizione 1.16. Dati due spazi topologici X e Y , il rivestimento $p : \tilde{X} \rightarrow X$ e una applicazione $f : Y \rightarrow X$, un **sollevamento** di f è una applicazione $\tilde{f} : Y \rightarrow \tilde{X}$ tale che $p \circ \tilde{f} (\equiv p\tilde{f}) = f$.

È possibile dimostrare il seguente teorema, che spesso è chiamato **proprietà di sollevamento delle omotopie**. La situazione è descritta graficamente dal diagramma riportato nella Figura 1.5.

Teorema 1.8. *Siano uno spazio topologico X e un suo rivestimento $p : \tilde{X} \rightarrow X$; date una mappa $f_0 : Y \rightarrow X$, una sua omotopia $f_\lambda : Y \rightarrow X$ e un suo sollevamento $\tilde{f}_0 : Y \rightarrow \tilde{X}$ esiste una e una sola omotopia \tilde{f}_λ di \tilde{f}_0 che sia un sollevamento dell'omotopia f_λ di f_0 .*

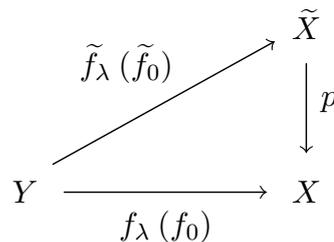


Figura 1.5

Considerando $Y = \{y_0\}$ come un punto, la mappa $f_0(y_0) = x_0$ identifica un punto di X e l'omotopia f_t descrive un "trascinamento continuo" del punto x_0 su X , cioè un cammino. In questo caso, la proprietà di sollevamento delle omotopie diventa la **proprietà di sollevamento dei cammini**:

Corollario. *Per ogni cammino $\gamma : [0, 1] \rightarrow X$ e per ogni punto \tilde{x}_0 che viene proiettato nel punto iniziale $\gamma(0) = x_0$ esiste un unico cammino $\tilde{\gamma} : [0, 1] \rightarrow \tilde{X}$ con punto iniziale \tilde{x}_0 che sia un sollevamento di γ su \tilde{X} .*

Per visualizzare quanto affermato si consideri l'esempio della circonferenza, intesa come sottoinsieme di \mathbb{R}^2 , con rivestimento \mathbb{R} e $p : \mathbb{R} \rightarrow S^1$, $p(t) = (\cos(2\pi t), \sin(2\pi t))$, e si faccia riferimento alla figura 1.4. Sia γ un cammino in S^1 con punto iniziale $\gamma(0) = x_0 = (1, 0)$; il punto iniziale è sollevato tramite p su $\tilde{X} = \mathbb{R}$ in $p^{-1}(x_0) = \{\tilde{x}_0 \in \mathbb{R} \mid 2\pi k, k \in \mathbb{Z}\}$ (questo insieme ha la cardinalità di \mathbb{Z} , dunque è un ricoprimento con un'infinità numerabile di fogli). Scegliendo uno qualunque dei punti che si proiettano in x_0 come punto di partenza, si ha un unico cammino $\tilde{\gamma}$ su \mathbb{R} che sia un sollevamento di γ (cioè per cui $p\tilde{\gamma} = \gamma$) e che cominci dal punto $\tilde{x}_0 = 2\pi k$ scelto. Vediamo inoltre che se il cammino è un loop sullo spazio $X = S^1$, non necessariamente il suo sollevamento è ancora un loop sul rivestimento $\tilde{X} = \mathbb{R}$; ad esempio il loop definito da $\gamma : [0, 1] \rightarrow S^1$, $\gamma(t) = (\cos(2\pi t), \sin(2\pi t))$ (è la stessa legge della proiezione, ma qui definisce un cammino), che ha punto base $\gamma(0) = \gamma(1) = (1, 0)$ è sollevato da p nell'identità $\tilde{\gamma} : [0, 1] \rightarrow \mathbb{R}$, $\tilde{\gamma}(t) = id_{\mathbb{R}}(t) = t$, perché $p \circ id_{\mathbb{R}} = \gamma$. Ciò accade proprio perché il rivestimento copre lo spazio base più volte: per tornare al punto iniziale anche sul rivestimento servono *più giri* sullo spazio base; in questo caso, \mathbb{R} ricopre infinite volte S^1 e non si tornerà mai nel punto di partenza anche sul rivestimento, a meno di considerare loop banali in S^1 , che non compiono giri.

Al contrario, dato uno spazio X e il rivestimento $p : \tilde{X} \rightarrow X$, ogni loop su \tilde{X} in \tilde{x}_0 , sollevamento di x_0 , è proiettato tramite p in un loop su X in x_0 . Ciò segue dalla definizione di funzione: se $\tilde{\gamma}(0) = \tilde{\gamma}(1)$, chiaramente $p\tilde{\gamma}(0) = p\tilde{\gamma}(1)$ e anche la proiezione è un loop.² Questa corrispondenza induce naturalmente una mappa tra i gruppi fondamentali $\pi_1(X, x_0)$ e $\pi_1(\tilde{X}, \tilde{x}_0)$, che si può definire come

$$\begin{aligned} p_* : \pi_1(\tilde{X}, \tilde{x}_0) &\rightarrow \pi_1(X, x_0) \\ [\tilde{\gamma}] &\longmapsto p_*([\tilde{\gamma}]) = [p \circ \tilde{\gamma}], \end{aligned} \tag{1.3.1}$$

cioè associando ad una classe di omotopia di \tilde{X} la classe di omotopia in X della proiezione di un cammino rappresentante della classe di \tilde{X} . Questa mappa è ben definita sulle classi di omotopia perché se $\tilde{\gamma}' \sim \tilde{\gamma}$ secondo l'omotopia \tilde{h}_λ , tramite $p : \tilde{X} \rightarrow X$ su X trovo che anche le proiezioni γ', γ sono omotope con omotopia h_λ . Dalle proprietà di sollevamento segue il prossimo risultato.

²Come si diceva sopra, γ loop $\not\Rightarrow$ $\tilde{\gamma}$ loop, perché p non è richiesta invertibile e p^{-1} può non essere una funzione ben definita sulla base

Teorema 1.9. *La mappa $p_* : \pi_1(\tilde{X}, \tilde{x}_0) \rightarrow \pi_1(X, x_0)$ indotta dal rivestimento $p : \tilde{X} \rightarrow X$ è un omomorfismo iniettivo. Il sottogruppo immagine $p_*(\pi_1(\tilde{X}, \tilde{x}_0)) \subseteq \pi_1(X, x_0)$ è l'insieme delle classi di omotopia in $\pi_1(X, x_0)$ il cui sollevamento in (\tilde{X}, \tilde{x}_0) è ancora un loop.*

Essendo p_* iniettiva, ha come nucleo solo l'elemento neutro: l'unica classe di loop in \tilde{X} che si proietta nei loop banali in X sono i loop banali in \tilde{X} . Come si era osservato prima, esistono loop in X che sollevati non danno loop in \tilde{X} : siccome $p_*(\pi_1(\tilde{X}, \tilde{x}_0)) \subseteq \pi_1(X, x_0)$, il numero di classi di loop che si possono costruire sul rivestimento generalmente diminuisce e ci si avvicina alla semplice connessione. Viene naturale porsi la seguente domanda: è sempre possibile trovare un rivestimento di spazio topologico che sia semplicemente connesso? La costruzione che ora sarà fatta mostra che la risposta è affermativa, entro moderate ipotesi di regolarità che non saranno qui approfondite: si richiede che lo spazio sia connesso per archi, localmente connesso per archi e semilocalmente semplicemente connesso (cioè, intuitivamente, che ci sia un limite inferiore alla dimensione dei buchi). A patto di restringersi alle componenti connesse dei domini in studio, sono ipotesi normalmente valide nei casi di interesse fisico.

L'idea alla base della costruzione è la seguente. Consideriamo un rivestimento di X con proiezione $p : (\tilde{X}, \tilde{x}_0) \rightarrow (X, x_0)$ che sia semplicemente connesso; $\forall \tilde{x} \in \tilde{X}$ esiste una sola classe di omotopia di cammini che congiungono \tilde{x} a \tilde{x}_0 per semplice connessione, dunque si possono identificare i punti di \tilde{X} con le classi di omotopia sul rivestimento; mediante la proiezione p , queste omotopie sono sollevamenti di omotopie su X ; questo indica di costruire il rivestimento semplicemente connesso partendo dai cammini su X .

Costruiamo allora lo spazio $\mathcal{P}X$ dei cammini su X , su esso prendiamo i cammini il punto iniziale x_0 fissato, lasciamo libero il punto di arrivo $x \in X$ e infine indentifichiamo tramite la relazione di omotopia \sim (con punti iniziale e finali fissati, secondo la (1.1.4)) i cammini equivalenti; siamo allora portati a definire

$$\tilde{X} = \mathcal{P}X / \sim = \{[\gamma] \mid \gamma \in \mathcal{P}X, \gamma(0) = x_0\} \quad (1.3.2)$$

e definiamo anche la mappa di proiezione come

$$\begin{aligned} p : \tilde{X} &\rightarrow X \\ \tilde{x} = [\gamma] &\mapsto p([\gamma]) = \gamma(1), \end{aligned} \quad (1.3.3)$$

cioè proietto ogni elemento del rivestimento, che è una classe di cammini omotopi su X , sul punto di arrivo in X del cammino. La funzione p è ben

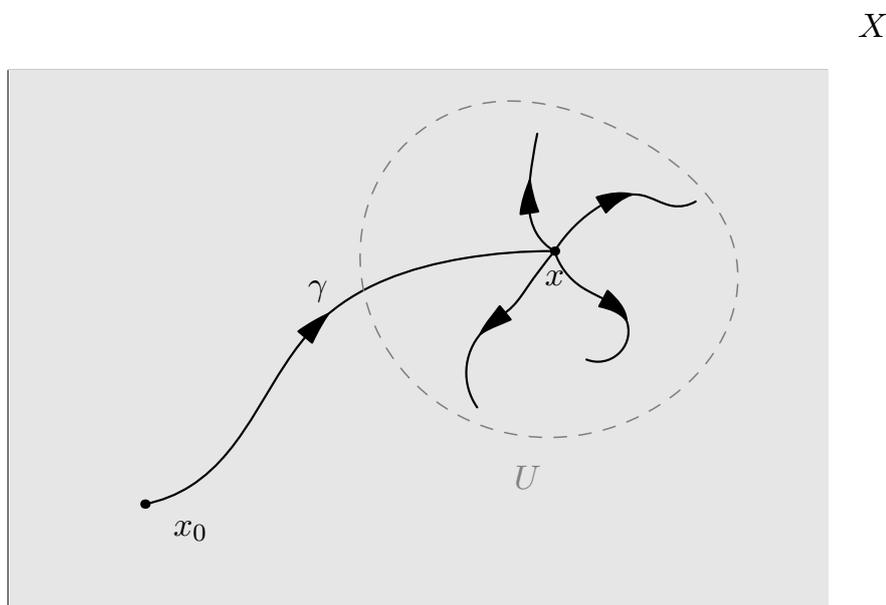


Figura 1.6: Rappresentazione di un intorno del punto $[\gamma]$ del rivestimento universale, costituito dalla classe della composizione di γ stesso e di tutti i cammini δ (di cui ne sono stati raffigurati sono alcuni) che cominciano dal punto finale di γ e terminano entro un certo aperto dello spazio base.

definita perché l'omotopia è richiesta con estremi fissati. Questa proiezione è suriettiva perché X è supposto connesso per archi e dunque è un rivestimento di tutto lo spazio iniziale.

A questo punto bisogna costruire la topologia del "candidato" rivestimento universale \tilde{X} . Sia \mathcal{U} la famiglia di sottoinsiemi $U \subset X$ connessi per archi tali per cui la mappa indotta dall'inclusione $\pi_1(U) \rightarrow \pi_1(X)$ è banale, cioè per cui ogni loop in U , visto come loop in X , è deformabile in un punto. Definiamo la topologia di \tilde{X} come generata dai seguenti aperti $U_{[\gamma]}$, dove $\gamma \in \mathcal{P}X$ è un cammino da $\gamma(0) = x_0$ a $\gamma(1) = x$, che è un punto in $U \in \mathcal{U}$:

$$U_{[\gamma]} = \{[\gamma \cdot \delta] \mid \delta \in \mathcal{P}U \text{ con punto iniziale } \delta(0) = x_0\}. \quad (1.3.4)$$

Un esempio di insieme $U_{[\gamma]}$ è rappresentato nella Figura 1.6. Si può dimostrare che con questa topologia p è continua e l'insieme \tilde{X} ha le proprietà richieste per essere un rivestimento. Il prossimo teorema enuncia un risultato centrale:

Teorema 1.10. *Sia X uno spazio topologico e sia $\tilde{X} = \mathcal{P}X / \sim$ definito come nella (1.3.2); il rivestimento \tilde{X} costruito è connesso per archi e semplicemente connesso.*

Dimostrazione. Mostriamo prima la connessione per archi. Sia $[\gamma] \in \tilde{X}$ un qualsiasi elemento del rivestimento; $\gamma : [0, 1] \rightarrow X$ è un cammino su X e definisco la sua restrizione γ_λ su $[0, \lambda]$. Ho dunque una famiglia di cammini³ su X $\lambda \mapsto \gamma_\lambda$ e posso definire questa corrispondenza sulle classi di omotopia come $\lambda \mapsto [\gamma_\lambda]$, ottenendo un cammino su \tilde{X} . Questo cammino nel rivestimento ha per punto iniziale $[\gamma_0]$, che è $[\gamma|_{[0,0]}] = [x_0]$, classe dei loop banali in x_0 , ed ha per punto finale $[\gamma_1] = [\gamma|_{[0,1]}] = [\gamma]$ e dunque è possibile congiungere ogni coppia di punti con cammini. Inoltre notiamo che questo cammino su \tilde{X} è un sollevamento di γ con la proiezione (1.3.3).

Per verificare la semplice connessione, si mostra che l'immagine tramite la proiezione sui gruppi fondamentali $p_*(\pi_1(\tilde{X}, \tilde{x}_0)) \subseteq \pi_1(X, x_0)$ è costituita dal solo elemento neutro, così, dovendo l'omomorfismo p_* essere per il teorema 1.9 iniettivo, anche il gruppo $\pi_1(\tilde{X}, \tilde{x}_0)$ sarà banale. Si era detto che $p_*(\pi_1(\tilde{X}, \tilde{x}_0))$ è costituita da tutti e soli i loop γ di X che sollevati restano loop di \tilde{X} . Si è detto che γ è sollevato in $\lambda \mapsto [\gamma_\lambda]$ e se il sollevamento è un cammino chiuso, allora $[\gamma_1] = [\gamma_0] \Rightarrow [\gamma] = [x_0]$, cioè γ è nullomotopo e $p_*(\pi_1(\tilde{X}, \tilde{x}_0))$ è banale. \square

Abbiamo verificato che, finché questa costruzione funziona, è possibile rivestire uno spazio topologico con uno spazio semplicemente connesso; diamo allora la seguente definizione:

Definizione 1.17. Un rivestimento (\tilde{X}, p) di uno spazio X è chiamato **rivestimento universale** di X se lo spazio \tilde{X} è semplicemente connesso.

Il nome è giustificato dal seguente fatto, che non dimostriamo [6].

Teorema 1.11. *Se (\tilde{X}, p) è un rivestimento fondamentale di uno spazio X , allora è possibile trovare una mappa di proiezione p' tale che (\tilde{X}, p') è un rivestimento di ogni altro rivestimento di X . Inoltre, tutti i rivestimenti fondamentali sono tra loro isomorfi,*

dove l'isomorfismo tra rivestimenti \tilde{X}_1, \tilde{X}_2 è definito come una mappa biunivoca ϕ tra i due spazi che conservi la struttura delle proiezioni, cioè tale che $p_1 = \phi \circ p_2$. Dunque il rivestimento che era stato costruito con i cammini su X non è solo un esempio di rivestimento fondamentale, ma può esserne visto come un modello generico.

Un esempio di rivestimento universale è proprio il caso per la circonferenza di \mathbb{R} con la proiezione $p : \mathbb{R} \rightarrow S^1$, $p(\tilde{x}) = (\cos(2\pi\tilde{x}), \sin(2\pi\tilde{x}))$: fissando, ad

³In precedenza i cammini erano stati definiti su $[0, 1]$, ma era una ipotesi soltanto di comodo ed è possibile farne a meno; anche nel seguito sarà spesso così.

esempio $x_0 = (1, 0)$, le classi di cammini omotopi su S^1 con punto iniziale x_0 e punto finale $x = (\cos(2\pi\bar{x}), \sin(2\pi\bar{x}))$ sono distinte dal numero n di giri che compiono (indice di avvolgimento) e hanno come rappresentanti i cammini definiti dal prodotto $\gamma \cdot \delta$, dove

$$\begin{aligned}\gamma : [0, 1] &\rightarrow S^1, \quad t \mapsto \gamma(t) = (\cos(2\pi nt), \sin(2\pi nt)), \\ \delta : [0, \bar{x}] &\rightarrow S^1, \quad t \mapsto \delta(t) = (\cos(2\pi t), \sin(2\pi t));\end{aligned}$$

qui si sono considerati solo i cammini in senso antiorario, ma contando anche quelli orari si ha chiaramente una corrispondenza biunivoca delle classi con \mathbb{R} (decomponendo $\mathbb{R} = \bigcup_{n \in \mathbb{Z}} [n, n + 1[$, il numero di giri -dato da γ - individua l'intervallo $[n, n + 1[$ e il punto di arrivo -dato da δ - il punto all'interno di questo intervallo).

L'idea per dimostrare che il gruppo fondamentale della circonferenza sia \mathbb{Z} è proprio partire da questo rivestimento. Se $\tilde{\gamma}_n$ è un cammino su \mathbb{R} da $\tilde{\gamma}(0) = 0$ a $\tilde{\gamma}(1) = n \in \mathbb{Z}$, la proiezione lo manda in un loop di S^1 , siccome $(\cos(2\pi n), \sin(2\pi n)) = (\cos(0), \sin(0))$, e questo loop ha indice di avvolgimento n . È possibile allora costruire la corrispondenza $\phi : \mathbb{Z} \rightarrow \pi_1(S^1)$, $\phi(n) = [p\tilde{\gamma}_n]$, che all'intero n associa la classe di omotopia del loop proiezione di $\tilde{\gamma}_n$. Questa proiezione è banalmente un omomorfismo tra i due gruppi: $\phi(n + m) = \phi(n) \cdot \phi(m)$, perché $n + m$ è mandato in un loop che compie $n + m$ giri, che è composizione di un loop che compie n giri e di uno che ne compie m . Tramite le proprietà di sollevamento enunciate si dimostra che è anche biunivoca e di conseguenze è un isomorfismo tra i due gruppi, per cui $\pi_1(S^1) = \mathbb{Z}$.

1.4 Fibrati e rivestimento universale

L'ultima sezione di questa introduzione matematica è dedicata alla struttura di fibrato principale del rivestimento universale di uno spazio topologico e alla descrizione dell'azione del gruppo fondamentale sul rivestimento universale. Per comodità, prima di ritornare allo studio della struttura dei rivestimenti, è utile richiamare alcune definizioni sui fibrati e alcune loro proprietà. L'idea che sta sotto alla definizione, è che i fibrati siano una sorta di generalizzazione del concetto del prodotto cartesiano di due spazio topologici $E = B \times F$, estendendolo in modo che lo spazio totale E sia localmente equivalente a $B \times F$, ma che possa differire da esso per proprietà topologiche globali. I fibrati saranno costruiti "incollando" ad ogni punto di uno spazio base B degli spazi, detti fibre, omeomorfi ma distinti in modo che sia definita in modo banale solo la proiezione sullo spazio base, mentre nei prodotti cartesiani

esistono le due proiezioni banali $proj_1$ e $proj_2$ su entrambe le componenti del prodotto.

Definizione 1.18. Un **fibrato** (E, π, B, F) è una struttura caratterizzata dai seguenti elementi:

- uno spazio topologico E detto *spazio totale*,
- uno spazio topologico B detto *spazio base*,
- uno spazio topologico F detto *fibra*,
- una *proiezione* $\pi : E \rightarrow B$ che sia una mappa suriettiva e continua,

dati in modo che la proiezione π sia *localmente banale*, cioè deve esistere un ricoprimento aperto $\{U_i\}$ della base B tale che per ogni U_i esista un omeomorfismo $\phi_i : \pi^{-1}(U_i) \rightarrow U_i \times F$ in modo che, se $\tilde{x} \in \pi^{-1}(U_i)$, $proj_1 \circ \phi_i(\tilde{x}) = \pi(\tilde{x})$ (dove $proj_1$ è l'ordinaria proiezione sul primo argomento di un prodotto cartesiano). Equivalentemente, il diagramma della Figura 1.7 dev'essere commutativo.

$$\begin{array}{ccc}
 \pi^{-1}(U_i) & \xrightarrow[\sim]{\phi_i} & U_i \times F \\
 \pi \downarrow & \swarrow proj_1 & \\
 U_i & &
 \end{array}$$

Figura 1.7

Questa definizione realizza l'idea di spazio E localmente equivalente a $B \times F$, appunto perché le mappe ϕ_j (chiamate *banalizzazioni locali* o *carte locali*) realizzano omeomorfismi tra aperti di E e il prodotto di un aperto di B per la fibra F .

Ogni punto $x \in B$, corrisponde tramite la proiezione inversa a più punti di E : $\pi^{-1}(x) = \{\tilde{x}_\alpha\}$. Per la richiesta di commutatività del diagramma della Figura 1.7, questi punti saranno mappati dalla banalizzazione locale ϕ_i in una coppia di punti in $U_i \times F$ per cui *tutti gli \tilde{x}_α corrispondono allo stesso punto della base*: $\phi_i(\tilde{x}_\alpha) = (x, \xi_\alpha)$. Dalla definizione segue allora che tutti gli insiemi $\pi^{-1}(x)$, con x sono omeomorfi a F , e per questo sono chiamati *fibra nel punto x* . La definizione è stata data con delle ipotesi minime, ma è chiaramente possibile richiedere che spazio base e fibra abbiano la struttura di varietà topologica/differenziabile; farlo darebbe anche allo spazio totale la struttura di varietà topologica/differenziabile e nella categoria delle varietà differenziabili gli omeomorfismi ϕ_j diventano diffeomorfismi.

Gli aperti U_i necessariamente si intersecano per ricoprire B . Sia una coppia di carte (U_i, ϕ_i) , (U_j, ϕ_j) intersecantesi, cioè con $U_i \cap U_j \neq \emptyset$; l'insieme $\pi^{-1}(U_i \cap U_j)$ potrà essere banalizzato sia da ϕ_i che da ϕ_j , e in generale le due mappe daranno coordinate locali diverse in $(U_i \cap U_j) \times F$ ai punti $\tilde{x} \in \pi^{-1}(U_i \cap U_j)$. Consideriamo le funzioni

$$\phi_{ij} = \phi_i \circ \phi_j^{-1} : (U_i \cap U_j) \times F \rightarrow (U_i \cap U_j) \times F.$$

Un punto $\tilde{x} \in E$ che si proietta in $x \in B$, tramite ϕ_j sarà mandato in una certa coppia (x, ξ_j) . Dunque, $\phi_{ij} = \phi_i \circ \phi_j^{-1}$: manda (x, ξ_j) prima in \tilde{x} e infine in (x, ξ_i) , realizzando un cambio di coordinate locali. Il punto dello spazio base resta lo stesso e viene cambiato solo il punto della fibra, dunque le ϕ_{ij} individuano degli omeomorfismi della fibra $t_{ij}(x)$ dati da

$$(x, \xi_j) \xrightarrow{\phi_{ij}} (x, \xi_i) = (x, t_{ij}(x) \cdot \xi_j).$$

L'invertibilità segue banalmente dal fatto che ϕ_i e ϕ_j sono omeomorfismi. Le $t_{ij}(x)$ sono chiamate *funzioni di transizione* e sono elementi di un gruppo \mathcal{G} (almeno $\mathcal{G} = \text{Homeo}(F)$) chiamato **gruppo di struttura** del fibrato, per cui è definita una opportuna azione a sinistra sulle fibre. Sia le ϕ_{ij} che le funzioni di transizione t_{ij} , per come sono definite, rispettano le seguenti *condizioni di cociclo*:

- (i) $\phi_{ii} = id_{U_i \times F} \iff t_{ii}(x) = id_F$;
- (ii) $\phi_{ij}^{-1} = \phi_{ji} \iff t_{ij}(x)^{-1} = t_{ji}(x)$;
- (iii) $\phi_{ij} \circ \phi_{jk} = \phi_{ik} \iff t_{ij}(x) \circ t_{jk}(x) = t_{ik}(x)$,

dove la (i) vale su tutti gli U_i , la (ii) sulle intersezioni di coppie di aperti per cui $U_i \cap U_j \neq \emptyset$ e la (iii) sulle intersezioni di tre aperti tali che $U_i \cap U_j \cap U_k \neq \emptyset$.

Un esempio ovvio è chiaramente il prodotto cartesiano $B \times F$ per qualsiasi coppia di spazi topologici con $\pi = proj_1$; tale fibrato è detto *fibrato banale*. Ora consideriamo una varietà differenziabile n -dimensionale \mathcal{M} . Ogni suo punto x ammette lo spazio tangente $T_x \mathcal{M}$, che per ogni x è isomorfo a \mathbb{R}^n ed è, dunque, anch'esso una varietà. Considerando come spazio totale

$$T\mathcal{M} = \bigcup_{x \in \mathcal{M}} \{x\} \times T_x \mathcal{M}$$

e dotandolo della proiezione $\pi : T\mathcal{M} \rightarrow \mathcal{M}$, $\pi(x, \vec{v}_x) = x$ (dove \vec{v}_x è un vettore tangente in $T_x \mathcal{M}$), si ottiene un fibrato definito sulla varietà che viene chiamato il *fibrato tangente* della varietà. Il fibrato tangente ha come fibre degli \mathbb{R}^n ed appartiene alla classe dei fibrati vettoriali: un **fibrato vettoriale**

ha per fibra uno spazio vettoriale, dunque $F \cong \mathbb{R}^n$ per un qualche n (purché sia a dimensione finita, come di solito accade). Spesso si fa la richiesta più stringente che le funzioni di transizione siano non solo omeomorfismi della fibra, ma anche automorfismi. Per un fibrato vettoriale la richiesta che siano automorfismi equivale a richiedere che siano applicazioni lineari invertibili (gli ordinari cambiamenti di base), per cui il gruppo di struttura di un fibrato vettoriale diventa $Aut(\mathbb{R}^n) = GL(n, \mathbb{R})$.

Prima di proseguire, ricordiamo che l'azione $X \times \mathcal{G} \rightarrow X$ di un gruppo \mathcal{G} su uno spazio X si dice *transitiva* se per ogni coppia di punti $x, x' \in X$ esiste un elemento del gruppo $g \in \mathcal{G}$ che li collega, cioè $x' = x \cdot g$, e che l'azione è detta *libera* se non esiste altro elemento g di \mathcal{G} oltre all'identità $e \in \mathcal{G}$ che lascia invariati tutti i punti di X rispetto alla sua azione, cioè $x \cdot g = x \forall x \Rightarrow x = e$. Definiamo ora una famiglia di fibrati di fondamentale importanza in fisica.

Definizione 1.19. Sia (E, π, B, F) un fibrato e sia \mathcal{G} un gruppo topologico. (E, π, B, F) è detto un **fibrato principale** con gruppo di struttura \mathcal{G} (o più brevemente un \mathcal{G} -fibrato) se:

1. è definita un'azione destra continua $E \times \mathcal{G} \rightarrow E$ che preservi le fibre (cioè $\tilde{x} \in \pi^{-1}(x) \Rightarrow \tilde{x} \cdot g \in \pi^{-1}(x) \forall g \in \mathcal{G}$) e che su esse sia
 - libera: $\tilde{x} \cdot g = \tilde{x} \forall \tilde{x} \in p^{-1}(x) \Rightarrow g = e$,
 - transitiva: $\forall \tilde{x}, \tilde{x}' \in p^{-1}(x) \exists g \in \mathcal{G} : \tilde{x}' = \tilde{x} \cdot g$.
2. il gruppo di struttura del fibrato è il gruppo \mathcal{G} e agisce su se stesso mediante traslazione sinistra.

Nell'ipotesi 1. della definizione, la richiesta di preservare le fibre può anche essere scritta come $\pi(x \cdot g) = \pi(x) \forall x, g$. La transitività dell'azione assicura di poter ricostruire interamente ogni fibra $\pi^{-1}(x)$, dato un punto in essa, tramite questa azione destra del gruppo. Vale infatti $\pi^{-1}(x) = \{\tilde{x}_0 \cdot g \mid g \in \mathcal{G}\}$, fissato un \tilde{x}_0 tale che $\pi(\tilde{x}_0) = x$. Il fatto che l'azione sia libera implica che per ogni punto \tilde{x} esista uno e un solo elemento $g \in \mathcal{G}$ che ricostruisce $\tilde{x} = \tilde{x}_0 \cdot g$, perché $\tilde{x} \cdot g = \tilde{x} \cdot h$ implica $\tilde{x} \cdot (gh^{-1}) = \tilde{x} \Rightarrow gh^{-1} = e$, cioè $g = h$. Dunque, fissato un qualsiasi $\tilde{x}_0 \in \pi^{-1}(x)$, posso mettere in corrispondenza biunivoca ciascun $\tilde{x} \in \pi^{-1}(x)$ con l'elemento del gruppo che lo ricostruisce e non ci sono elementi del gruppo che fanno uscire $\tilde{x}_0 \cdot g$ da $\pi^{-1}(x)$, siccome l'azione preserva le fibre. Dalla continuità dell'azione destra segue che la corrispondenza $p^{-1}(x) \cong \mathcal{G}$ è un omeomorfismo. Quindi un fibrato principale ha per fibra il gruppo \mathcal{G} e può essere pensato come localmente equivalente al prodotto $B \times \mathcal{G}$.

Le carte locali di un fibrato principale sono tipo $\phi_j : \pi^{-1}(U_j) \rightarrow U_j \times G$, $\tilde{x} \mapsto (x, g_j)$, per cui le funzioni di transizione realizzano la corrispondenza $(x, g_j) \mapsto (x, g_i) = (x, t_{ij}(x) \cdot g_j)$. Come nei fibrati vettoriali era naturale richiedere che fosse lineare la corrispondenza fra gli \mathbb{R}^n che costituivano le fibre, qui la naturale richiesta di regolarità è che l'azione che lega gli elementi del gruppo g_j e $g_i = t_{ij}(x) \cdot g_j$ corrisponda all'operazione interna al gruppo. Ciò equivale a chiedere che $t_{ij}(x) = h$ sia un elemento h del gruppo e che la mappa $\mathcal{G} \rightarrow \mathcal{G}$ dia $g_j \mapsto g_i = h \cdot g_j = hg_j$, dove l'azione del gruppo di struttura è indicata con \cdot e la sua operazione interna come monomio; questa è esattamente la definizione di traslazione sinistra. Riassumendo, *un \mathcal{G} -fibrato ha il gruppo \mathcal{G} sia come fibra che come gruppo di struttura.*

Prima di procedere a mostrare la struttura di fibrato dei rivestimenti, è utile soffermarsi un momento su concetto legato ai fibrati che torna spesso utile in fisica e sarà citato nei successivi capitoli. Il concetto è quello di sezione, che, tornando all'idea di fibrato come estensione del prodotto cartesiano, generalizza il concetto di grafico di una funzione. Se $E = B \times F$, una funzione $f : B \rightarrow F$ ha come grafico le coppie di punti $g(x) = (x, f(x)) \in B \times E$, ottenute associando ad ogni punto $x \in B$ un elemento $f(x)$ dello spazio F : al variare del punto x si sceglie un elemento della fibra "spostandosi" su di essa. Per il fibrato banale $B \times E$, si è detto che π è l'ordinaria proiezione sul primo argomento $proj_1$, per cui $(\pi \circ g)(x) = proj_1(x, f(x)) = x$. Le sezioni estendono direttamente questa idea alla struttura dei fibrati:

Definizione 1.20. Sia (E, π, B, F) un fibrato; una **sezione** di E è una funzione $\sigma : B \rightarrow E$ continua e tale che $\pi \circ \sigma = id_B$.

La richiesta $\pi \circ \sigma = id_B$ permette di formalizzare l'idea di associare ad ogni punto dello spazio base un elemento della fibra in quel punto, siccome proiettando si deve ritornare al punto di partenza. Oltre all'esempio banale già detto dei grafici delle funzioni, un altro esempio tipico di sezione sono i campi vettoriali (continui) su una varietà differenziabile \mathcal{M} , che in geometria differenziale sono definiti come sezioni del fibrato tangente $T\mathcal{M}$ (ad ogni punto $x \in \mathcal{M}$ associano un vettore in quel punto, cioè un elemento dello spazio tangente $T_x\mathcal{M}$ che costituisce la fibra).

Il problema dell'esistenza di una sezione per un fibrato non è banale, esistono casi in cui non sono ammessi. Ad esempio, il famoso *teorema della palla pelosa* [10] afferma che non esiste un campo vettoriale continuo tangente ad una sfera bidimensionale che sia ovunque non nullo. Detto in altro modo, se E è il fibrato ottenuto rimuovendo il vettore nullo alle varie fibre $T_x S^2$ del fibrato tangente alla sfera TS^2 , l'enunciato significa che E non ammette una sezione. Per questo è utile introdurre il concetto di *sezione locale*, una sezione

$\sigma : U (\subset B) \rightarrow E$ definita solo su un aperto dello spazio base, e riferirsi alla sezione definita precedentemente come sezione *globale*. I problemi nel definire sezioni globali sono legati a come due punti estremi si "incollano"; questi non possono presentarsi su aperti sufficientemente piccoli, per cui tutti i fibrati ammettono sezioni locali. In fisica si lavora molto con i fibrati principali, per cui vale il seguente teorema [19]:

Teorema 1.12. *Un fibrato principale ammette una sezione globale se e solo se è banale.*

Torniamo ora alla descrizione dei rivestimenti: il fatto che siano interpretabili come un generico fibrato è diretto dalle definizioni. Lo spazio totale è appunto il rivestimento $E = \tilde{X}$, lo spazio base è lo spazio topologico di partenza $B = X$ e la proiezione coincide con la proiezione di rivestimento $\pi = p : \tilde{X} \rightarrow X$. Concentrandosi come detto sui rivestimenti universali, la proiezione p è suriettiva (dimostrato nella costruzione di \tilde{X} con i cammini) come richiesto. Gli insiemi $p^{-1}(x)$, che si erano detti essere tutti omeomorfi (se X è connesso), sono le fibre nei vari punti e la fibra tipo F ne è un modello generico. Queste fibre sono insiemi discreti e, se $\{U_i\}$ è il ricoprimento di $B = X$ che realizza la definizione di fibrato, i prodotti $U_i \times F$ diventano una collezione di copie di U_i , con cardinalità pari a quella delle fibre. Quindi le proprietà che caratterizzano la struttura di fibrato si verificano usando come U_i gli U_α del ricoprimento che realizza definizione di rivestimento, in quanto anche $p^{-1}(U_\alpha)$ è una collezione di copie di U_α in numero pari alla cardinalità delle fibre e da questo segue $p^{-1}(U_\alpha) \cong U_\alpha \times F$. Pongo $p^{-1}(U_\alpha) = \bigcup_j \tilde{U}_{\alpha_j}$, dove gli \tilde{U}_{α_j} sono le copie di U_α che costituiscono $p^{-1}(U_\alpha)$. L'indice j conta le copie, che sono in numero pari alla cardinalità della fibra. Posso dunque associare ad ogni j in modo biunivoco un elemento della fibra tipo $\xi \in F$, $\xi = f(j)$; questa corrispondenza sarà banalmente continua (perché sono spazi discreti) e restano da mettere in corrispondenza gli \tilde{U}_{α_j} con U_α , ma questo è ovvio dalla definizione di rivestimento, per cui l'omeomorfismo è realizzato da $p|_{\tilde{U}_{\alpha_j}} \equiv p_{\alpha_j} : \tilde{U}_{\alpha_j} \rightarrow U_\alpha$. Dunque le banalizzazioni locali sono

$$\phi_\alpha(\tilde{x}) = (p_{\alpha_j}(\tilde{x}), f(j)),$$

scegliendo per il punto \tilde{x} l'opportuno j a cui corrisponde. La mappa costruita rispetta la condizione $proj_1 \circ \phi_\alpha = p$, perché, se $\tilde{x} \in p^{-1}(U_\alpha)$, $proj_1 \circ \phi_\alpha(\tilde{x}) = p_{\alpha_j}(\tilde{x}) = x$ e chiaramente è uguale a $p(\tilde{x}) = x$, essendone $p_{\alpha_j}(\tilde{x})$ una restrizione.

Abbiamo verificato che il rivestimento \tilde{X} è un fibrato sullo spazio base X , ma si può dire di più: il rivestimento universale ha infatti la struttura di $\pi_1(X)$ -fibrato principale, cioè la fibra di \tilde{X} può essere interpretata come il

gruppo fondamentale dello spazio base e anche il gruppo di struttura di questo fibrato è il gruppo fondamentale $\pi_1(X)$.⁴ Per verificarlo, mostriamo prima che le fibre sono omeomorfe al gruppo fondamentale e per farlo è necessario definire un'azione destra di $\pi_1 X$ che sia continua, libera, transitiva e che preservi le fibre.

Consideriamo allora un punto $\tilde{x} \in \tilde{X}$; questo punto apparterrà ad una certa fibra $p^{-1}(x) \subset \tilde{X}$, insieme dei punti che si proiettano sullo stesso x ; su questi punti agirà il gruppo tramite i suoi elementi $[\gamma] \in \pi_1(X)$. Per la proprietà di sollevamento dei cammini, ogni loop γ (con punto base x) è sollevato in un'unico cammino $\tilde{\gamma}$ su \tilde{X} se ne viene fissato il punto iniziale; per continuità della proiezione, è preservata la struttura delle classi di omotopia e classi di loop su X sono mandate in univoche classi di cammini su \tilde{X} , fissando il punto iniziale. Scelgo allora come punto iniziale del sollevamento il punto $\tilde{x} \in p^{-1}(x)$ su cui voglio fare agire la trasformazione. Definisco allora questa azione del gruppo sul rivestimento come l'applicazione che al punto iniziale $\tilde{\gamma}(0) = \tilde{x}$ associa il punto finale del cammino sollevato $\tilde{\gamma}(1)$. Essendo γ un loop, $\gamma(1) = \gamma(0)$ e il punto di arrivo è ancora $\tilde{\gamma}(1) \in p^{-1}(x)$, dunque è un'azione definita internamente a $p^{-1}(x)$ e *preserva le fibre*.

Definizione 1.21. Sia X uno spazio topologico con gruppo fondamentale $\pi_1(X)$ e sia $p : \tilde{X} \rightarrow X$ un suo rivestimento. L'**azione (destra) di $\pi_1(X)$ sulla fibra $p^{-1}(x)$** è definita da

$$\begin{aligned} \star : p^{-1}(x) \times \pi_1(X, x) &\rightarrow p^{-1}(x) \\ (\tilde{x}, [\gamma]) &\longmapsto \tilde{x} \star [\gamma] = \tilde{\gamma}(1), \end{aligned} \tag{1.4.1}$$

dove $\tilde{\gamma}$ è il sollevamento di γ con punto iniziale \tilde{x} .

Questa mappa è ben definita e dipende effettivamente solo dalla classe di omotopia del loop, perché, come si è detto, scegliendo come rappresentante un loop omotopo γ' , il sollevamento $\tilde{\gamma}'$ è omotopo con estremi fissati a $\tilde{\gamma}$ e non cambia il punto di arrivo. L'azione è stata definita considerando per ogni punto il gruppo fondamentale *in quel punto* in modo da avere loop che visibilmente cominciano "nel punto giusto", ma in termini di azione astratta non è fondamentale, essendo tutti i gruppi fondamentali nei vari punti isomorfi.

Conseguenza di questa definizione è che le fibre hanno la struttura di $\pi_1(x)$ -insieme, cioè su di esse è definita l'azione \star del un gruppo $\mathcal{G} = \pi_1(X)$ tale che:

⁴Anche ad un rivestimento più generico si può dare la struttura di fibrato principale, ma con un gruppo di struttura contenuto nel gruppo fondamentale.

- (i) $\tilde{x} \star e = \tilde{x} \quad \forall \tilde{x} \in p^{-1}(x)$ (con e elemento neutro di $\mathcal{G} = \pi_1(X)$),
(ii) $(\tilde{x} \star g) \star h = \tilde{x} \star (g \cdot h) \quad \forall g, h \in \mathcal{G} = \pi_1(X), \forall \tilde{x} \in p^{-1}(x)$;

la dimostrazione di queste proprietà è una verifica immediata. Queste sono le usuali proprietà che si richiedono affinché una azione sia considerabile ben definita su un certo insieme. In questo caso, dotando il gruppo fondamentale della topologia discreta, l'azione è banalmente anche continua e la struttura diventa di $\pi_1(X)$ -spazio.

Pensando al rivestimento universale come nella costruzione in cui ogni punto è il punto finale di una classe di cammini omotopi è evidente che l'azione sia *libera*, cioè, $\tilde{x} \star [\gamma] = \tilde{x} \quad \forall \tilde{x} \Rightarrow [\gamma] = e$. Questo perché se $\tilde{x} \star [\gamma] = \tilde{x}$ per ogni \tilde{x} , γ dev'essere per forza nullomotopo, altrimenti il suo punto d'arrivo corrisponderebbe per costruzione ad un punto diverso del rivestimento universale. L'azione definita è chiaramente anche *transitiva*, cioè esiste un elemento del gruppo $[\gamma]$ che collega ogni coppia di punti \tilde{x}, \tilde{x}' proprio perché è così che si è costruito il modello di rivestimento universale, a cui ogni altro è isomorfo.

Dunque l'azione è interna alle fibre, libera e transitiva e questo permette di affermare che le orbite $\tilde{x} \star \pi_1(X) = \{\tilde{x} \star [\gamma] \mid [\gamma] \in \pi_1(X)\}$ siano tutte in corrispondenza biunivoca con le fibre $p^{-1}(x)$ e la continuità dell'azione le rende omeomorfe. Prima di terminare la costruzione del rivestimento come fibrato, ricordiamo la decomposizione di \tilde{X} in domini fondamentali. Ogni punto di \tilde{X} appartiene ad un'orbita e posso costruire una relazione di equivalenza su \tilde{X} identificando due punti \tilde{x}, \tilde{x}' se stanno in una stessa orbita. L'insieme di queste classi di equivalenza è il quoziente dello spazio rispetto all'azione $\tilde{X}/\pi_1(X)$ e queste classi partizionano \tilde{X} . Scegliere un punto di ogni classe equivale a scegliere un punto di ogni fibra e si ottiene chiaramente un insieme in corrispondenza biunivoca con X e si dimostra [17] possibile fare questa scelta in modo da dividere \tilde{X} in una unione di copie \tilde{X}_j omeomorfe (o diffeomorfe, considerando anche la struttura differenziabile) di X chiamate **domini fondamentali**. L'azione destra di $\pi_1(X)$ su \tilde{X} collega tutti i domini fondamentali, siccome permette di passare tra punti diversi della stessa fibra.

Resta da vedere che il gruppo fondamentale sia anche identificabile con il gruppo di struttura del fibrato. Se $x \in U_\alpha \cap U_\beta$, gli $\tilde{x} \in p^{-1}(x)$ hanno due diverse banalizzazioni locali ϕ_β e ϕ_α , che danno a \tilde{x} le coordinate (x, ξ_β) e $(x, \xi_\alpha) = \phi_\alpha \circ \phi_\beta^{-1}(x, \xi_\beta)$ e questo definisce le funzioni di transizione $t_{\alpha\beta}(x)$ $\xi_\alpha = t_{\alpha\beta}(x) \cdot \xi_\beta$. Le $t_{\alpha\beta}(x)$ si richiedono appartenenti a un gruppo, che è il gruppo di struttura del fibrato. Richiediamo che

$$\phi_{\alpha\beta} = \phi_\alpha \circ \phi_\beta^{-1} \in \text{Aut}(\tilde{X}),$$

cioè che siano isomorfismi (con lo stesso significato che si usa nel Teorema 1.11) del rivestimento \tilde{X} in sé. È naturale porre tale richiesta, siccome sono queste le biiezioni che conservano la struttura di un rivestimento. Il gruppo degli automorfismi di un rivestimento $Aut(\tilde{X})$ è chiamato spesso anche gruppo delle *trasformazioni di rivestimento*. Considerando ad esempio il rivestimento di S^1 dato da $p : \mathbb{R} \rightarrow S^1$, $p(t) = (\cos 2\pi t, \sin 2\pi t)$ e identificando $\tilde{S}^1 = \mathbb{R}$ con l'elica che si arrotola su S^1 , come nella figura 1.4, gli automorfismi di rivestimento sono identificabili con le traslazioni verticali dell'elica.

Definizione 1.22. Un rivestimento $p : \tilde{X} \rightarrow X$ si dice **regolare** se per ogni $x \in X$ e per ogni $\tilde{x}, \tilde{x}' \in p^{-1}(x)$ esiste una trasformazione di rivestimento che manda \tilde{x} in \tilde{x}' .

Ricordiamo ora che dato un gruppo G e un suo sottogruppo H , H è detto un sottogruppo normale di G se è invariante sotto coniugio, cioè se per ogni $h \in H$ gli elementi ghg^{-1} sono ancora in H per ogni $g \in G$. Non tutti i sottogruppi sono normali; un esempio banale di sottogruppo normale è appunto il sottogruppo di G costituito dalla sola identità, perché chiaramente $geg^{-1} = e$. Oppure un altro esempio è un qualsiasi sottogruppo di un gruppo abeliano, che è normale perché $gh = hg \Rightarrow ghg^{-1} = h \in H$. Per un sottogruppo generico H si può definire il suo normalizzatore $N(H)$, che è il più piccolo sottogruppo di G di cui H è sottogruppo normale. Se H è normale chiaramente $N(H) = G$. Enunciamo ora il seguente teorema senza dimostrazione, per la quale si può fare riferimento a [6].

Teorema 1.13. Sia $p : (\tilde{X}, \tilde{x}_0) \rightarrow (X, x_0)$ un rivestimento connesso per archi di X connesso per archi e localmente connesso per archi. Allora:

- a) \tilde{X} è regolare se e solo se $p_*(\pi_1(\tilde{X}, \tilde{x}_0)) \subset \pi_1(X, x_0)$ ne è un sottogruppo normale
- b) $Aut(\tilde{X})$ è isomorfo al quoziente $N(H)/H$, con $H = p_*(\pi_1(\tilde{X}, \tilde{x}_0))$.

Il rivestimento universale soddisfa le ipotesi del teorema ed è normale per il punto a), perché $H = p_*(\pi_1(\tilde{X}, \tilde{x}_0)) = \{e\} \in \pi_1(X, x_0)$, tutti i loop in esso si proiettano in loop banali di X , e chiaramente il gruppo banale è un sottogruppo normale di qualsiasi gruppo. Siccome H è normale $N(H) = \pi_1(X, x_0)$ e per il punto b) $Aut(\tilde{X}) \cong \pi_1(X, x_0)/\{e\} = \pi_1(X, x_0)$. Allora il gruppo di automorfismi del rivestimento universale \tilde{X} è isomorfo al gruppo fondamentale $\pi_1(X)$. Tornando alla costruzione di \tilde{X} come fibrato principale, si è detto che $Aut(\tilde{X})$ è il gruppo in cui risiedono le trasformazioni $\phi_{\alpha\beta}$ ed è banale costruire un isomorfismo tra esso e il gruppo di struttura, a cui appartengono

le funzioni di transizione. Questo completa la costruzione: *il rivestimento universale di uno spazio topologico è un $\pi_1(X)$ -fibrato principale che ha per base lo spazio che per riveste.*

Capitolo 2

Spazi multiplamente connessi in meccanica quantistica

La dinamica classica di un sistema avente spazio delle configurazioni Q è descritta dalle sue equazioni del moto, che tipicamente sono equazioni differenziali del secondo ordine rispetto al tempo t nella posizione $\vec{q}(t)$. Essendo di secondo ordine, è sufficiente dare posizione e velocità $\vec{q}(t_0), \dot{\vec{q}}(t_0)$ ad un certo istante t_0 per determinare univocamente l'evoluzione temporale successiva del sistema; tutte le informazioni sullo *stato classico* di un sistema in un certo istante sono contenute nella posizione e nella velocità in quell'istante. In modo equivalente si può passare ad usare l'impulso $\vec{p} = m\dot{\vec{q}}$ al posto della velocità $\dot{\vec{q}}$: posizione e impulso (\vec{q}, \vec{p}) possono essere viste come le coordinate di uno spazio, che viene chiamato spazio delle fasi, i cui punti corrispondono agli stati di un sistema classico.

Quando il sistema classico viene quantizzato, si perde la chiarezza di questa corrispondenza tra grandezze cinematiche e stati del sistema. In particolare, per ottenere una trattazione coerente, diventa necessario identificare in modo astratto lo *stato quantistico* di un sistema; un modo per farlo è usare le funzioni d'onda. Una funzione d'onda, però, ha gradi di libertà ridondanti nella sua fase ed è un punto importante chiedersi se sia possibile fissarli in qualche modo. Nella sezione §2.1 si analizza la corrispondenza tra stati quantistici e funzione d'onda, nella §2.2 è portato l'esempio dell'effetto Aharonov-Bohm, che mostra alcuni problemi legati alla fase. Successivamente, nella §2.3 si daranno alcune indicazioni su come costruire la meccanica quantistica quando lo spazio delle configurazioni del sistema non è semplicemente connesso, perché è questo il caso che porta ai problemi in questione sulla fase. Infine, nella sezione §2.4 si farà un primo esempio di applicazione della teoria descritta, studiando il problema di una particella singola vincolata ad una circonferenza.

2.1 Stati quantistici e fasi

Nella meccanica quantistica la corrispondenza tra lo stato e l'oggetto matematico che lo descrive non è così semplice come nel caso classico, in cui c'è una corrispondenza biunivoca tra punti (\vec{q}, \vec{p}) dello spazio delle fasi e stati. Quantizzando un sistema classico, posizione ed impulso da funzioni vengono promossi ai due operatori autoaggiunti \hat{q} e \hat{p} e non determinano più necessariamente lo stato del sistema, ma sono due osservabili come tante altre: possono essere misurate, ma il loro valore non può identificare lo stato, perché esistono stati su cui le misure di posizione e impulso possono dare uno spettro di diversi risultati. Anzi, si può affermare che non esista nessuno stato identificabile da posizione e impulso, proprio perché vale il principio di indeterminazione di Heisenberg: non esistono stati in cui sia posizione che impulso assumono entrambi valori ben definiti.

È evidente che si debba cambiare approccio e lo si fa postulando che lo stato quantistico di un sistema sia codificato in un oggetto astratto, che è il **vettore (ket) di stato** $|\psi\rangle$. Questi oggetti astratti devono essere elementi di uno spazio vettoriale perché gli esperimenti suggeriscono (ad esempio per gli effetti di interferenza) che debba valere una forma di principio di sovrapposizione degli stati, il quale viene tradotto in una combinazione lineare dei corrispondenti ket. Si richiede che lo spazio vettoriale abbia ulteriore struttura: esso dev'essere uno *spazio di Hilbert* \mathcal{H} , cioè in esso dev'essere definito un prodotto interno hermitiano tale che la metrica che questo prodotto induce renda \mathcal{H} completo. È una richiesta necessaria anche per ragioni tecniche, siccome la teoria spettrale si intreccia ampiamente con la meccanica quantistica. I vettori di stato possono essere rappresentati tramite le loro coordinate in una opportuna base; una scelta comune è quella di considerare la base ortonormale di autovettori $|\vec{q}\rangle$ (autostati) che diagonalizzano l'operatore posizione \hat{q} . In questo modo, un vettore di stato è identificato dal suo insieme rappresentativo (nome che le coordinate in una certa base prendono in questo contesto), che viene indicato come

$$\langle \vec{q} | \psi \rangle \equiv \psi(\vec{q}), \quad \vec{q} \in Q.$$

L'oggetto $\psi(\vec{q})$ è una funzione a valori complessi definita sullo spazio delle configurazioni Q ed è la **funzione d'onda** del sistema (nella rappresentazione delle posizioni o di Schrödinger) e contiene tutte le informazioni necessarie per identificare lo stato quantistico del sistema, esattamente come il ket di stato, la cui espansione nella rappresentazione è

$$|\psi\rangle = \int_Q d^3q |\vec{q}\rangle \langle \vec{q} | \psi \rangle. \quad (2.1.1)$$

In questa trattazione ci si concentrerà sulla meccanica quantistica nella rappresentazione di Schrödinger e in un contesto non relativistico. Si lavorerà soprattutto con sistemi senza spin, in cui lo stato quantistico è descritto da una funzione d'onda scalare; se il sistema avesse spin s sarebbe necessaria una funzione d'onda vettoriale a $g_s = 2s + 1$ componenti e molte affermazioni si generalizzano in modo diretto. La funzione d'onda in sé non ha immediato significato fisico, se non appunto il fatto che individui lo stato. Un motivo per cui ci si concentra sulla rappresentazione delle posizioni è che classicamente le coordinate spaziali hanno *de facto* un ruolo privilegiato ed è più naturale partire da essa per quantizzare il sistema. Un'altra ragione è perché l'ordinaria interpretazione di Born è caratteristica di questa rappresentazione; essa vuole che tutta l'informazione fisica sia contenuta nel modulo quadrato $|\psi(\vec{q})|^2$, il quale deve rappresentare la densità di probabilità di localizzare una particella nello spazio. Essendo una densità di probabilità, è naturale richiedere che valga la condizione di normalizzazione

$$\langle \psi, \psi \rangle_{L^2} = \int_Q |\psi(\vec{q})|^2 d^3q = 1,$$

ma spesso è comodo indebolirla e lavorare solo con funzioni d'onda normalizzabili $\psi \in L^2(Q)$, dato che basterà dividere tale funzione per la sua norma per renderla interpretabile come densità di probabilità. Per coerenza, lo stesso deve valere con i ket astratti $|\psi\rangle$: come la richiesta di normalizzazione porta naturalmente a definire le funzioni d'onda di Schrödinger come elementi dello spazio di Hilbert $L^2(Q)$ su cui è definito il corrispondente prodotto interno, anche per gli stati astratti dovrà esistere un opportuno prodotto interno rispetto cui i ket direttamente interpretabili fisicamente sono normalizzati. La richiesta di normalizzazione su un ket di stato è

$$\langle \psi | \psi \rangle = 1,$$

la quale, tramite la relazione di Parseval $\langle \varphi | \psi \rangle = \int_Q d^3q \langle \varphi | \vec{q} \rangle \langle \vec{q} | \psi \rangle$, equivale alla condizione di normalizzazione sulle funzioni d'onda. I ket non normalizzati si usano nella pratica, ma fisicamente rappresentano lo stesso stato del corrispondente ket a norma unitaria.

Lo spazio di Hilbert introdotto dà una descrizione ridondante degli stati: più punti danno lo stesso stato, ed è naturale identificare le funzioni d'onda che danno lo stesso stato. Non si può restringere lo spazio in cui si lavora perché in generale non è detto che resti uno spazio di Hilbert, ma vista la ridondanza si pone l'equivalenza $\psi \sim \psi'$ se

$$\frac{\psi}{\langle \psi, \psi \rangle} = \frac{\psi'}{\langle \psi', \psi' \rangle}$$

(facendo cadere nella notazione il pedice L^2 al prodotto interno), cioè se $\exists c > 0$ tale che $\psi' = c\psi$. Si era detto però che tutte le informazioni fisiche sono contenute in $|\psi(\vec{q})|^2$ e questo rende necessaria una ulteriore identificazione. Se la funzione d'onda è moltiplicata per una fase $e^{i\theta} \in U(1)$, essendo a modulo unitario, essa si perde in $|\psi(\vec{q})|^2$, per cui un cambiamento della **fase globale** (o complessiva) della funzione d'onda

$$\psi(\vec{q}) \mapsto \psi'(\vec{q}) = e^{i\theta}\psi(\vec{q})$$

risulta inosservabile. Essendo identificabili due stati che differiscono sia per un numero reale positivo (la norma) che per una fase qualsiasi, complessivamente si stanno identificando due stati che differiscono per un qualsiasi numero complesso: $\psi(\vec{q}) \sim \psi'(\vec{q}) \iff \exists \lambda \in \mathbb{C}$ tale che $\psi(\vec{q}) = \lambda\psi'(\vec{q})$. Dunque gli stati quantistici non corrispondono in modo biunivoco ai vettori dello spazio di Hilbert ma ai suoi sottospazi unidimensionali, detti *raggi*, che sono gli elementi dello spazio proiettivo (di Hilbert) che queste identificazioni individuano. La fase può essere pure una funzione dipendente dal punto e l'identificazione resta valida, l'unico vincolo è che appunto il cambiamento sia globale: se si ha una sovrapposizione di stati $\psi(\vec{q}) + \psi'(\vec{q})$, è inosservabile la differenza da $e^{i\theta(\vec{q})} \cdot (\psi(\vec{q}) + \psi'(\vec{q}))$, mentre è diverso lo stato $\psi(\vec{q}) + e^{i\theta}(\vec{q}) \cdot \psi'(\vec{q})$, siccome una variazione di fase locale porta a diversi effetti di interferenza.

La fase globale è un grado di libertà ridondante delle funzioni d'onda ed è utile poterlo fissare in modo opportuno. Tuttavia, non sempre è possibile farlo: esistono casi in cui la fase presenta patologie, è ambigua, e questa ambiguità impedisce di eliminarla in modo "innocuo", determinando conseguenze misurabili sul sistema. Nella successiva sezione si descrive l'esempio dell'effetto Aharonov - Bohm, un caso tipico in cui la fase della funzione d'onda del sistema diventa polidroma e ciò causa effetti di interferenza.

2.2 L'effetto Aharonov-Bohm

Consideriamo un solenoide cilindrico infinito per il momento spento, che occupi la regione \mathcal{C} contenuta in \mathbb{R}^3 . Prendiamo un punto base \vec{x}_0 lontano dal cilindro come sorgente di elettroni che vengono inviate verso il cilindro stesso (senza poterlo attraversare), descritte ciascuna dalla funzione d'onda $\psi(\vec{x})$ e poniamo uno schermo piano che riveli la posizione delle particelle al di là del cilindro, in modo che sia posizionato parallelamente all'asse del cilindro e opposto ad \vec{x}_0 . La situazione, schematizzata dal punto di vista di una base del cilindro nella Figura 2.1, ha evidente simmetria cilindrica; è quindi comodo passare a descriverla in coordinate cilindriche $\vec{x} = (z, r, \theta)$,

dove z è lungo l'asse del cilindro e r è centrato nella base. Per simmetria possiamo studiare solo ciò che accade su una sezione $(z_0, r, \theta) \sim (r, \theta)$ con z_0 qualsiasi valore fissato, siccome il comportamento sarà indipendente da z_0 .

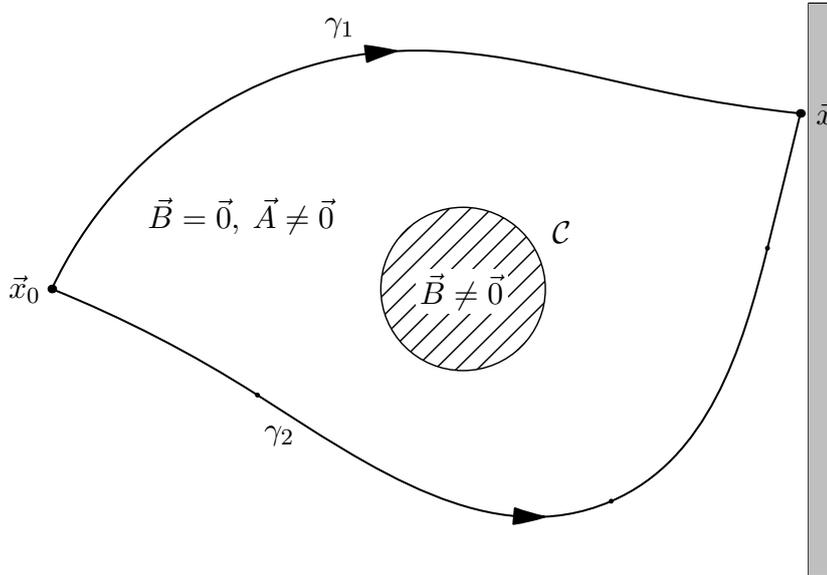


Figura 2.1: Schema di un apparato ideale per l'osservazione dell'effetto Aharonov - Bohm. Considerando \vec{x}_0 come punto base, le particelle inviate possono seguire i cammini γ_1 e γ_2 prima di ricongiungersi nel punto di osservazione \vec{x} ; la non omotopia di γ_1 e γ_2 causa uno sfasamento che sposta la figura di interferenza che si avrebbe se \vec{A} fosse nullo.

La situazione è evidentemente analoga all'esperimento della doppia fenditura: detta in breve, il cilindro, al posto del setto tra le fenditure, "separa" in due la funzione d'onda, per poi farla ricongiungere ed interferire. Facendo misure di posizione appena dopo il cilindro, si può considerare l'osservabile "la particella è passata a destra o a sinistra del cilindro". Questa osservabile può assumere solo due valori e di conseguenza lo stato preparato inizialmente ψ ha autostati ψ_1 e ψ_2 in cui può collassare, dunque ψ si potrà scrivere come loro sovrapposizione: $\psi = \psi_1 + \psi_2$. L'esistenza di queste due componenti in cui si può dividere la funzione d'onda genera un fenomeno d'interferenza, alla stessa maniera in cui ciò accade nell'esperimento della doppia fenditura.

Accendiamo ora il campo magnetico \vec{B} del solenoide: esso sarà costante all'interno e di valore nullo al di fuori. Per dare una descrizione quantistica del fenomeno, è necessario quantizzarlo partendo dalla formulazione canonica del problema. Come è noto, un campo magnetico non è associato a un potenziale

scalare ma entra nella dinamica mediante un potenziale vettore, che porta a ridefinire l'impulso. L'hamiltoniana classica del sistema è

$$H(\vec{q}, \vec{p}) = \frac{1}{2m} (\vec{p} - e\vec{A}(\vec{q}))^2,$$

dove \vec{A} è il potenziale vettore associato al campo magnetico, che soddisfa la relazione $\vec{B} = \vec{\nabla} \times \vec{A}$. Quantizzando in modo ordinario [5], l'equazione di Schrödinger è

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle = \frac{1}{2m} (\hat{\mathbf{p}} - e\vec{A}(\hat{\mathbf{q}}))^2 |\psi(t)\rangle;$$

scrivendo gli oggetti nella rappresentazione delle posizioni, l'impulso diventa l'operatore differenziale $-i\hbar\vec{\nabla}$ e il potenziale vettore l'operatore moltiplicativo $\vec{A}(\vec{q})$, per cui l'equazione di Schrödinger diventa

$$i\hbar \frac{\partial \psi}{\partial t}(\vec{x}, t) = \frac{1}{2m} (i\hbar\vec{\nabla} + e\vec{A}(\vec{q}))^2 \psi(\vec{x}, t). \quad (2.2.1)$$

Prima di risolverla, bisogna determinare la forma funzionale del potenziale vettore nella regione in cui è ammesso il moto della particella. Al di fuori del solenoide il campo magnetico è nullo, ma questo in generale non implica che sia nullo anche il potenziale vettore:

$$\vec{B} = \vec{\nabla} \times \vec{A} = 0 \implies \vec{A} = \vec{\nabla}\varphi \neq \vec{0}.$$

Questa espressione merita ulteriori commenti. È ben noto che un campo irrotazionale abbia un potenziale regolare, ma questo vale solo se il dominio della relazione è semplicemente connesso (così il campo diventa conservativo). In questo caso la regione \mathcal{C} è interdotta ed è possibile scrivere ancora $\vec{A} = \vec{\nabla}\varphi$, purché si ammetta φ come funzione polidroma. Questo perché il "candidato" campo potenziale φ dev'essere della forma

$$\varphi(\vec{x}) = \int_{\gamma(\vec{x})} \vec{A} \cdot d\vec{s},$$

dove $\gamma(\vec{x})$ è un cammino dal punto base \vec{x}_0 al punto in cui il campo è valutato. Tale integrale di linea è l'integrale di una 1-forma chiusa ($\vec{A} \cdot d\vec{s}$ è chiusa ove \vec{A} è irrotazionale) e ha lo stesso valore quando è valutato lungo cammini omotopi. Se lo spazio fosse semplicemente connesso esisterebbe un'unica classe di cammini congiungenti ogni coppia di punti in esso, cioè tutti i possibili cammini $\gamma(\vec{x})$ sarebbero omotopi, e $\varphi(\vec{x})$ sarebbe una funzione univoca ben definita. Invece lo spazio in questione non è semplicemente connesso,

perché, siccome è possibile scrivere $\mathcal{C} = \mathcal{B} \times \mathbb{R}$, dove \mathcal{B} è la sezione circolare del cilindro), si ha che

$$\pi_1(\mathbb{R}^3 \setminus \mathcal{C}) = \pi_1(\mathbb{R} \times (\mathbb{R}^2 \setminus \mathcal{B})) = \pi_1(\mathbb{R}) \times \pi_1(\mathbb{R}^2 \setminus \mathcal{B}) = \{e\} \times \mathbb{Z} = \mathbb{Z}.$$

Dunque, esistono più classi distinte di cammini omotopi e, integrando lungo cammini appartenenti a classi diverse, si ottengono valori diversi di $\varphi(\vec{x})$ per ogni \vec{x} , che diventa così una funzione polidroma.

Prendiamo ad esempio i due cammini γ_1 e γ_2 indicati nella figura 2.1. Essi evidentemente non sono omotopi, perché per deformatli con continuità bisognerebbe attraversare il buco circolare \mathcal{B} e questo non è ammesso, dunque porteranno a valori diversi $\varphi(\vec{x})_1 \neq \varphi(\vec{x})_2$, la cui differenza è pari a

$$\varphi(\vec{x})_2 - \varphi(\vec{x})_1 = \oint_{\gamma_2 \cdot \gamma_1^{-1}} \vec{A} \cdot d\vec{s} \stackrel{!}{=} \int_{\Sigma} (\nabla \times \vec{A}) \cdot d\vec{\sigma} = \int_{\Sigma} \vec{B} \cdot d\vec{\sigma} = \Phi,$$

che è il flusso del campo magnetico attraverso il solenoide, e dipende da γ_1 e γ_2 solo tramite la loro classe di omotopia, siccome il campo magnetico è non nullo solo nel solenoide. L'uguaglianza contrassegnata da $\stackrel{!}{=}$ segue dall'applicazione del teorema del rotore, se Σ è una superficie avente per bordo (orientato) $\partial\Sigma = \gamma_2 \cdot \gamma_1^{-1}$.

Assumendo, come è d'uso fare, che il potenziale vettore sia della forma $\vec{A} = A\hat{e}_z$, cioè solo lungo l'asse del cilindro, la relazione precedente determina il valore del potenziale vettore:

$$\vec{A}(r, \theta) = \frac{\Phi}{2\pi r} \hat{e}_z. \quad (2.2.2)$$

Inserendo la (2.2.2) nella (2.2.1) ho esplicitamente l'equazione di Schrödinger per questo problema e si potrebbe procedere a risolverla. Per vedere quello che ci interessa esiste una alternativa più conveniente rispetto alla soluzione diretta. Sfruttando la libertà di gauge del potenziale vettore, è noto che si possa compiere la trasformazione

$$\vec{A} \mapsto \vec{A}' = \vec{A} + \nabla\Lambda,$$

senza modificare i campi elettrico e magnetico (e dunque anche la dinamica del sistema), se Λ è una funzione regolare della sola posizione. Sotto a questa trasformazione, la funzione d'onda cambia secondo

$$\psi(\vec{x}, t) \mapsto \psi'(\vec{x}, t) = e^{\frac{ie}{\hbar}\Lambda(\vec{x})}\psi(\vec{x}, t), \quad (2.2.3)$$

che è unitaria e non modifica lo stato, cambia solo la sua rappresentazione (per una dimostrazione si veda, ad esempio, [16]). In questo caso, per eliminare il potenziale vettore verrebbe naturale prendere $\Lambda = -\varphi$. Il problema è che, strettamente parlando, non sarebbe corretto, perché solitamente si ammettono solo trasformazioni di gauge con funzione generatrice della trasformazione Λ sufficientemente regolare e univoca, mentre qui sarebbe polidroma (quindi una funzione discontinua, ristretta ai suoi rami). In questo caso possiamo comunque procedere formalmente ed eliminare il potenziale vettore, perché, come si vedrà, questa trasformazione non rimuove effetti fisici (in generale eliminare un potenziale potrebbe cambiare il problema, ad esempio eliminando cariche elettriche...) ma li "sposta" sulla funzione d'onda.

La trasformazione di gauge porta alla funzione d'onda modificata

$$\psi'(\vec{x}) = e^{-\frac{ie}{\hbar}\varphi(\vec{x})}\psi(\vec{x}), \quad (2.2.4)$$

che, come si può anche direttamente verificare sostituendo nella (2.2.1), soddisfa l'equazione di Schrödinger libera (in cui, esattamente come si voleva, è stato eliminato il potenziale vettore)

$$i\hbar\frac{\partial\psi'}{\partial t}(\vec{x}, t) = -\frac{\hbar^2}{2m}\nabla^2\psi'(\vec{x}, t),$$

se $\psi(\vec{x}, t)$ soddisfa l'equazione di Schrödinger completa (2.2.1). Dunque la funzione d'onda che descrive lo stato di una particella in questa configurazione sperimentale è descritto da

$$\psi(\vec{x}) = e^{\frac{ie}{\hbar}\varphi(\vec{x})}\psi'(\vec{x}) = e^{i\frac{e}{\hbar}\int_{\gamma(\vec{x})}\vec{A}\cdot d\vec{s}}\psi'(\vec{x}), \quad (2.2.5)$$

quando $\psi'(\vec{x})$ è una soluzione dell'equazione di Schrödinger libera. Una traccia del potenziale vettore rimane nel fatto che, nella soluzione (2.2.5), la fase $e^{\frac{ie}{\hbar}\varphi(\vec{x})}\psi'(\vec{x})$ è una funzione polidroma e, in generale, il valore della funzione d'onda dipende dal cammino percorso.

Esattamente come nel caso della doppia fenditura, supponiamo che la funzione d'onda sia sovrapposizione di ψ_1 e ψ_2 , corrispondenti ai due cammini (a destra o a sinistra) che percorre per superare il solenoide, e che oltre al solenoide queste due componenti interferiscano. Inoltre, assumiamo che per calcolare ψ_1 e ψ_2 il cammino da usare nell'integrale che definisce $\varphi(\vec{x})$ sia quello che segue la particella, cioè, rispettivamente, γ_1 e γ_2 .¹ Quando le due

¹Questi argomenti troverebbero in realtà una più naturale collocazione nella formulazione della meccanica quantistica con gli integrali sui cammini, perché, per calcolare la funzione d'onda propagata, si "sommano" i contributi di tutti i cammini e la fase $e^{iS[\gamma]/\hbar}$ è il peso del cammino γ . Questa fase è la stessa che compare qui, perché l'azione che genera il campo magnetico è proprio $e\int dt\vec{v}\cdot\vec{A} = e\int d\vec{s}\cdot\vec{A}$.

funzioni si ricongiungono in un punto \vec{x} sullo schermo, esse hanno accumulato due fasi diverse, per cui si ha una differenza di fase pari a

$$\Delta\left(\frac{e\varphi}{\hbar}\right) = \frac{e}{\hbar}(\varphi(x)_2 - \varphi(x)_1) = \frac{e\Phi}{\hbar}.$$

Questa differenza di fase ulteriore è indipendente dalla posizione del punto \vec{x} sullo schermo (per omotopia dei cammini) e va aggiunta alle differenze di fase "ordinarie" di propagazione delle onde, provocando così una modifica alla figura di interferenza al variare del flusso magnetico.

Quanto visto merita un commento ulteriore perché è di grande importanza concettuale nella fisica. Esso mostra come sia possibile avere effetti elettromagnetici anche in assenza di campi ed è un fatto che apparirebbe assurdo nell'impostazione della fisica classica, in cui \vec{E} e \vec{B} hanno un ruolo privilegiato, mentre i loro potenziali sono visti quasi alla stregua di un puro oggetto matematico. L'effetto Aharonov - Bohm ribalta questa prospettiva, mostrando come anche i potenziali abbiano di per sé un ruolo fisicamente determinante e, se non nulli, portino a effetti misurabili, anche dove i campi sono nulli. È mantenuta comunque l'invarianza di gauge dell'elettromagnetismo, perché ciò che è misurabile è il flusso magnetico e non il potenziale vettore in sé.

2.3 Quantizzazione su spazi non semplicemente connessi

L'esempio descritto nella sezione precedente ha mostrato quanto possa essere determinante il ruolo della fase della funzione d'onda di un sistema: quando essa diventa polidroma non è più in generale eliminabile senza perdere effetti fisici osservabili. Questa è una situazione tipica di quando il sistema ha lo spazio delle configurazioni che non è semplicemente connesso, esattamente com'era nel caso precedente, in cui $Q = \mathbb{R}^3 \setminus \mathcal{C}$. In questo caso, dati due punti nello spazio, non si ha un'unica classe di cammini omotopi che li congiunge. Lungo cammini topologicamente equivalenti la fase varia allo stesso modo, ma cambiando classe di omotopia cambia anche la variazione della fase, per cui lo stesso punto può corrispondere a più fasi diverse. È così anche in generale e non solo per il caso elettromagnetico visto con l'effetto Aharonov - Bohm, perché le interazioni fisiche sono descritte con potenziali che tipicamente sono forme differenziali chiuse, in assenza delle cariche sensibili all'interazione in esame, e integrarli su spazi non semplicemente connessi porta a primitive polidrome e ad effetti visibili, anche se non ci sono cariche.

Abbiamo discusso fino ad ora l'importanza della fase di una funzione d'onda, nonostante non compaia nella densità di probabilità $|\psi|^2$, ed è quindi evidente che nella trattazione vada inclusa un'identificazione della fase. Per affrontare il problema, è comodo [7] ridefinire le funzioni d'onda su uno spazio delle configurazioni "allargato" ad un fibrato principale $U(1)$ con spazio base Q , che chiamiamo \hat{Q} . Intuitivamente, questo fibrato è ottenuto associando una circonferenza, rappresentativa della fase, ad ogni punto dello spazio delle configurazioni; in coordinate locali, i punti questo fibrato saranno tipo $(\vec{q}, e^{i\theta})$. Questa costruzione è comoda perché permette di descrivere assieme tutte le classi di funzioni d'onda che differiscono per delle fasi: prima si fissa il modulo della funzione d'onda come funzione di un punto, successivamente bisogna determinare il valore della fase *in quel punto*, scegliendo un punto della circonferenza che è la fibra $e^{i\theta(\vec{q})}$ incollata a quel punto dello spazio base Q . Usando le coordinate locali del fibrato, posso indicare la funzione d'onda sul fibrato come $\psi(\hat{q}) = \psi(\vec{q}, e^{i\theta})$. L'azione di un elemento $g = e^{i\alpha}$ del gruppo di struttura $U(1)$ sul fibrato è definita come

$$\hat{q} = (\vec{q}, e^{i\theta}) \xrightarrow{g} \hat{q} e^{i\alpha} = (\vec{q}, e^{i(\theta+\alpha)}), \quad (2.3.1)$$

ed è un automorfismo che permette di "muoversi" sulla fibra. Vogliamo che questa funzione d'onda definita sul fibrato sia coerente con la richiesta che le osservabili dipendano solo da $|\psi|^2$, dunque un cambiamento della fase corrispondente all'azione di $U(1)$ non deve produrre effetti osservabili. Si richiede allora la condizione di proiettività

$$\psi(\hat{q}e^{i\alpha}) = e^{in\alpha} \cdot \psi(\hat{q}), \quad n \in \mathbb{Z}. \quad (2.3.2)$$

È chiaro che per l'invarianza di $|\psi|^2$ l'azione debba essere la moltiplicazione per una fase, ma si può vedere anche che questa azione debba essere una rappresentazione unitaria di $U(1)$ sulle funzioni d'onda, in modo del tutto simile a come sarà fatto esplicitamente più avanti (nella discussione che precede il Teorema (2.1)). Rimangono soltanto le fasi etichettate dall'intero n perché sono queste tutte le possibili rappresentazioni unitarie di $U(1)$.

Una discussione più approfondita (che si può trovare in [21]), mostra che le funzioni d'onda possano essere viste come sezioni di un fibrato vettoriale associato a \hat{Q} ²; in questo modo, il problema di trovare una funzione d'onda ben definita diventa equivalente a quello della costruzione di una sezione globale di un fibrato. Tale sezione esiste solo quando il fibrato \hat{Q} è banale, cioè globalmente diffeomorfo a $Q \times U(1)$, e ciò non si verifica quando Q

²Un fibrato vettoriale associato ad un certo fibrato E si ottiene "cambiando" la fibra di E con un opportuno spazio vettoriale e mantenendone il gruppo di struttura. Per dettagli si veda ad esempio [19].

non è semplicemente connesso, perché non si riesce a definire la fase come funzione continua univoca. Se \hat{Q} non è banale ammette solo sezioni locali, e, corrispondentemente, la fase delle funzioni d'onda si riesce a definire univocamente solo su aperti. I problemi di discontinuità dei rami della funzione polidroma diventano equivalenti a quelli del raccordo degli estremi della sezione del fibrato, che, quando è "avvolto", non si riescono ad incollare.

Studiamo dunque qual è la procedura per descrivere un sistema quantistico che abbia spazio delle configurazioni Q non semplicemente connesso. Nella sezione §1.3 abbiamo visto che, a meno di casi patologici, per ogni spazio topologico è possibile costruirne il rivestimento universale (\tilde{Q}, p) , passando per le classi di cammini omotopi tra due punti dello spazio Q . Sul rivestimento \tilde{Q} si potrà costruire una meccanica quantistica in modo ordinario, perché la sua semplice connessione garantisce che le funzioni d'onda siano ben definite globalmente ad un solo valore; il punto è costruire questa meccanica quantistica in modo che sia coerente con quanto deve accadere su Q , cioè deve essere *proiettabile* su Q .

Innanzitutto, il nuovo stato sarà rappresentato da una funzione d'onda sollevata $\tilde{\psi}(\tilde{q})$ ad un solo valore. Nella rappresentazione di Schrödinger le osservabili sono descritte da operatori definiti sulle funzioni d'onda su Q

$$\hat{A} : L^2(Q) \longrightarrow L^2(Q), \quad \psi(\vec{q}) \longmapsto \hat{A}\psi(\vec{q})$$

e vanno sollevati in operatori che agiscono sulle funzioni d'onda su \tilde{Q}

$$\tilde{A} : L^2(\tilde{Q}) \longrightarrow L^2(\tilde{Q}), \quad \tilde{\psi}(\tilde{q}) \longmapsto \tilde{A}\tilde{\psi}(\tilde{q}).$$

Questi operatori dovranno essere sollevati ed è possibile farlo senza ambiguità perché sono tipicamente operatori differenziali, o comunque operatori locali, ed è possibile ricoprire il loro dominio di definizione con un solo aperto del ricoprimento che realizza la definizione di rivestimento. Gli operatori sollevati avranno un loro dominio di autoaggiunzione, corrispondente alle condizioni al contorno che si impongono sulle funzioni d'onda, che sarà fondamentale nel determinare la soluzione al problema. Vogliamo trovare le **condizioni di proiettabilità** di $\tilde{\psi}$, cioè, la domanda a cui bisogna rispondere è la seguente:

quali sono le condizioni al contorno che bisogna imporre sulle funzioni d'onda $\tilde{\psi}$ su \tilde{Q} in modo che esse siano proiettabili su funzioni d'onda ψ accettabili su Q ?

Notiamo che il sollevamento della funzione d'onda non è diretto, va definito in modo diverso rispetto a com'era stato detto nella definizione (1.16), perché non combaciano i domini, come si può vedere dal diagramma della

Figura 2.2. Qui $\psi : Q \rightarrow \mathbb{C}$, $\tilde{\psi} : \tilde{Q} \rightarrow \mathbb{C}$ e $p : \tilde{Q} \rightarrow Q$, per cui la composizione $p\psi$ non è ben definita.

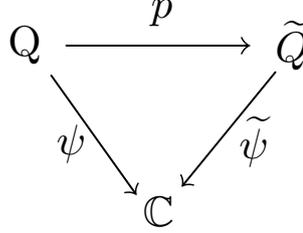


Figura 2.2

Una strada naturale potrebbe essere usare l'inverso della proiezione $p^{-1} : Q \rightarrow \tilde{Q}$ e richiedere che

$$\psi = \tilde{\psi} \circ p^{-1}, \tag{2.3.3}$$

ma questa condizione non è affatto banale perché tipicamente la proiezione di rivestimento p non è biunivoca e l'inversa non è una funzione ben definita. Schematizzando, la situazione è la seguente:

$$\vec{q} \xrightarrow{p^{-1}} p^{-1}(\vec{q}) = \{ \tilde{q}, \tilde{q}', \dots \} \xrightarrow{\tilde{\psi}} \{ \tilde{\psi}(\tilde{q}), \tilde{\psi}(\tilde{q}'), \dots \}.$$

Dunque, il sollevamento (2.3.3), a causa della non univocità di p^{-1} , porta naturalmente ad una funzione d'onda $\psi(\vec{q})$ polidroma. Esistono sistemi fisici descritti da tali funzioni, come si è visto per l'effetto Aharonov - Bohm, dunque è possibile accettare questa polidromia. Vedremo che per accettarla vanno imposti dei vincoli alla funzione d'onda sul rivestimento; questi vincoli si tradurranno in condizioni al contorno sui rami della funzione d'onda polidroma su Q e dunque in condizioni da aggiungere ai domini degli operatori.

Per comodità, richiamiamo le proprietà dei rivestimenti che mostrano come "coprano più volte" lo spazio topologico sottostante. Sia $\tilde{q} \in p^{-1}(\vec{q})$ è un punto che solleva, \vec{q} e sia γ un cammino chiuso su Q con punto base \vec{q} . Per la proprietà di sollevamento dei cammini, è unico il cammino $\tilde{\gamma}$ su \tilde{Q} che solleva γ , una volta fissato il sollevamento $\tilde{q} \in p^{-1}(\vec{q})$ da cui si vuole che cominci. Si era detto anche che questo sollevamento $\tilde{\gamma}$ tipicamente non è un loop basato in \tilde{q} , ma ha per punto finale un altro $\tilde{q}' \in p^{-1}(\vec{q})$ appartenente ad un diverso dominio fondamentale e questo punto, per la definizione (1.21), si può scrivere come $\tilde{q}' = \tilde{q} \star [\gamma]$.

Questa azione permetteva di ricostruire le fibre:

*compiendo cammini chiusi in \vec{q} nello spazio delle configurazioni Q ,
si ottengono cammini sul rivestimento universale \tilde{Q} che portano a
tutti i punti distinti $\tilde{q}_j = \tilde{q} \star [\gamma]_j$ del rivestimento che si proiettano
nel punto base dei loop su Q ,*

dove j indicizza le classi di omotopia $[\gamma]_j \in \pi_1(Q)$.

Sappiamo dunque la relazione tra i vari punti $\tilde{q}, \tilde{q}', \dots \in p^{-1}(\vec{q})$ e a questo punto ci chiediamo: qual è la corrispondente relazione tra le loro immagini $\tilde{\psi}(\tilde{q}), \tilde{\psi}(\tilde{q}'), \dots \in \tilde{\psi}(p^{-1}(\vec{q})) = \{\tilde{\psi}(\tilde{q} \star [\gamma]_j)\}_{\pi_1(Q)}$?

I vari punti $\tilde{q} \in p^{-1}(\vec{q})$ corrispondono tutti alla stessa posizione. Vogliamo mantenere l'interpretazione probabilistica: le misure di posizione hanno come densità di probabilità associata $|\psi(\vec{q})|^2$ e la probabilità di misurare la posizione di una particella si vuole che resti calcolabile senza ambiguità, dunque $|\psi(\vec{q})|^2$ dev'essere ben definita come funzione a un solo valore. Questo ci porta a chiedere che il modulo quadro della funzione d'onda sia ben proiettabile, cioè che valga

$$|\tilde{\psi}(\tilde{q})|^2 = |\tilde{\psi}(\tilde{q}')|^2 \quad \forall \tilde{q}, \tilde{q}' \in p^{-1}(\vec{q});$$

allora, cambiando il punto appartenente alla stessa fibra $p^{-1}(\vec{q})$, la funzione d'onda sollevata deve differire al più per una fase. Ogni coppia di punti \tilde{q}, \tilde{q}' è collegata in generale da più di un cammino chiuso in Q , perché tutti i loop omotopi in \vec{q} si sollevano in cammini omotopi con estremi \tilde{q} e \tilde{q}' fissati. Q è semplicemente connesso e $\tilde{\psi}$ dev'essere a un solo valore, per cui questa fase non può dipendere in generale dal cammino seguito. Si è appena detto però che i loop omotopi in \vec{q} collegano gli stessi punti di \tilde{Q} , dunque $\tilde{\psi}$ è a un solo valore se si chiede che la fase che acquisisce dipenda solo dalla classe di omotopia $[\gamma] \in \pi_1(Q)$ del loop percorso. Siamo giunti alla seguente condizione che vincola la polidromia della ψ conseguente dalla (2.3.3) [11]:

*la funzione d'onda $\tilde{\psi}$ costruita sul rivestimento universale \tilde{Q} è
proiettabile su Q se soddisfa la condizione*

$$\tilde{\psi}(\tilde{q} \star [\gamma]) = a([\gamma])\tilde{\psi}(\tilde{q}) \quad \forall \tilde{q} \in \tilde{Q}, \quad (2.3.4)$$

dove a è una fase, per cui $|a([\gamma])| = 1 \quad \forall [\gamma] \in \pi_1(Q)$.

In questo modo, proiettando su Q , si ottiene una funzione d'onda ψ polidroma, i cui valori nello stesso punto differiscono solo per una fase e il modulo quadro è una funzione monodroma. Ricordando la decomposizione del rivestimento \tilde{Q} in domini fondamentali, sappiamo che in ogni dominio fondamentale \tilde{Q}_j esiste uno e un solo valore $\tilde{q}_j \in p^{-1}(\vec{q})$ per ogni $\vec{q} \in Q$,

cioè ogni dominio fondamentale "copre Q una sola volta". Ottengo quindi i rami di ψ , che chiamo ψ_j , proiettando su Q soltanto la restrizione al dominio fondamentale $\tilde{\psi} |_{\tilde{Q}_j} \equiv \psi_j$ della funzione d'onda sul rivestimento.

I rami ψ_j sono funzioni univoche ben definite in $L^2(Q)$, ma si differenziano dalle comuni funzioni d'onda che si hanno su spazi semplicemente connessi perché sono discontinue nel punto in cui si raccorderebbero tra loro; la continuità si recupera solo considerando assieme tutti i rami della funzione polidroma. Per passare da un ramo all'altro si deve compiere in un cammino chiuso in un punto su Q , siccome è questa l'azione che porta da un dominio fondamentale all'altro.

Chiamando $\psi_{[\gamma]j}$ il ramo in cui ψ_j va dopo aver compiuto il loop $[\gamma]$, si ha una azione di $\pi_1(Q)$ sui rami delle funzioni d'onda che lega i diversi valori nello stesso punto di ψ

$$\begin{aligned} a_{[\gamma]} : L^2(Q) &\longrightarrow L^2(Q) \\ \psi_j &\longmapsto a_{[\gamma]}\psi_j \equiv \psi_{[\gamma]j} = a([\gamma]) \cdot \psi_j \end{aligned} \quad (2.3.5)$$

e questa azione induce la mappa

$$\begin{aligned} \rho_a : \pi_1(Q) &\longrightarrow \text{Aut}(L^2(Q)) \\ [\gamma] &\longmapsto \rho_a([\gamma]) = a_{[\gamma]}. \end{aligned} \quad (2.3.6)$$

Possiamo dire di più e caratterizzare ulteriormente la fase $a([\gamma])$ sfruttando la mappa ρ_a appena definita. Siano due classi $[\gamma], [\gamma'] \in \pi_1(Q)$ e tre punti $\tilde{q}, \tilde{q}', \tilde{q}'' \in \tilde{Q}$ tali che

$$\tilde{q}' = \tilde{q} \star [\gamma] ; \quad \tilde{q}'' = \tilde{q}' \star [\gamma'],$$

per cui vale anche $\tilde{q}'' = (\tilde{q} \star [\gamma]) \star [\gamma'] = \tilde{q} \star ([\gamma] \cdot [\gamma']) (= \tilde{q} \star [\gamma \cdot \gamma'])$. Applicando $\tilde{\psi}$ si ha:

$$\tilde{\psi}(\tilde{q}'') = \tilde{\psi}(\tilde{q}' \star [\gamma']) = a([\gamma'])\tilde{\psi}(\tilde{q}') = a([\gamma'])a([\gamma])\tilde{\psi}(\tilde{q}),$$

ma anche

$$\tilde{\psi}(\tilde{q}'') = \tilde{\psi}(\tilde{q} \star [\gamma] \cdot [\gamma']) = a([\gamma'] \cdot [\gamma])\tilde{\psi}(\tilde{q}).$$

Ma $\tilde{\psi}$ è monodroma, per cui deve valere $a([\gamma'] \cdot [\gamma]) = a([\gamma'])a([\gamma])$: l'applicazione $a : \pi_1(Q) \rightarrow U(1)$ che alla classe di cammini associa la fase con cui la funzione d'onda si trasforma è un omomorfismo.

Per definizione vale $\rho_a([\gamma] \cdot [\gamma']) = a_{[\gamma] \cdot [\gamma']}$ e di conseguenza anche

$$a_{[\gamma] \cdot [\gamma']}\psi_j = a([\gamma] \cdot [\gamma'])\psi_j \stackrel{!}{=} a([\gamma]) \cdot a([\gamma'])\psi_j = a_{[\gamma]}a_{[\gamma']}\psi_j,$$

dove in $\stackrel{!}{=}$ si è usato che $a : \pi_1(Q) \rightarrow U(1)$ è un omomorfismo. Si è dimostrato che anche ρ_a è un omomorfismo tra il gruppo $\pi_1(Q)$ e lo spazio degli automorfismi delle funzioni d'onda, per cui ρ_a è una rappresentazione unitaria di $\pi_1(Q)$ sulle funzioni d'onda. L'unitarietà è ovvia, perché moltiplica i rami per delle fasi. Enunciamo ora come teorema un riassunto dei risultati trovati in questa sezione.

Teorema 2.1. *La meccanica quantistica su uno spazio non semplicemente connesso Q si costruisce proiettandola dalla meccanica quantistica sul rivestimento universale \tilde{Q} , imponendo sulla funzione d'onda sollevata la condizione $\tilde{\psi}(\tilde{q} \star [\gamma]) = a([\gamma])\tilde{\psi}(\tilde{q})$. Questa condizione porta a una funzione d'onda su Q polidroma, per cui portare un punto \vec{q} intorno ad un loop γ su Q fa passare tra i vari rami di ψ , che differiscono tra loro per l'azione di una rappresentazione unitaria di $\pi_1(Q)$.*

Quanto visto si estende direttamente al caso con spin, in cui si hanno funzioni d'onda non scalari ma vettoriali: l'unica modifica è che le rappresentazioni unitarie con cui le funzioni d'onda si trasformano non sono più unidimensionali ma le fasi $a([\gamma])$ sono sostituite da matrici unitarie.

Consideriamo ora due diverse funzioni d'onda ψ_1, ψ_2 (ora il pedice non indica il ramo, ma le due distinte funzioni d'onda polidrome) che, dopo lo stesso cammino chiuso γ , prendono due fasi diverse:

$$\psi_1(\vec{q}) \xrightarrow{\gamma} a_1([\gamma])\psi_1(\vec{q}) \quad \text{e} \quad \psi_2(\vec{q}) \xrightarrow{\gamma} a_2([\gamma])\psi_2(\vec{q}).$$

Non esiste nessuna trasformazione unitaria sulle ψ che mandi una nell'altra come funzioni polidrome, dunque le due funzioni d'onda non rappresentano lo stesso stato ma sono due distinte quantizzazioni inequivalenti. Partendo da un sistema classico con spazio delle configurazioni non semplicemente connesso, esistono più possibilità distinte per quantizzare il sistema o, detta in altro modo, *esistono più sistemi quantistici corrispondenti allo stesso sistema classico*.

Il fatto che funzioni d'onda che prendono fasi diverse descrivano sistemi quantistici diversi è sottolineato ulteriormente dal fatto per esse non può valere il principio di sovrapposizione. Definendo infatti $\psi = \psi_1 + \psi_2$, se si compie un cammino chiuso, ψ_1 e ψ_2 cambiano di fasi diverse e si produce una fase relativa, che causa interferenze osservabili. A causa di ciò, dopo il giro γ , diventa

$$\psi(\vec{q}) \xrightarrow{\gamma} a_1 \{ \psi_1(\vec{q}) + a_1^* a_2 \cdot \psi_2(\vec{q}) \},$$

e i due valori non differiscono solo per una fase. Dunque $|\tilde{\psi}|^2$ non è da considerarsi proiettabile e *il principio di sovrapposizione vale solo per funzioni*

d'onda che vengono moltiplicate per la stessa fase, se trasportate lungo un loop di Q .

Emerge dunque una struttura decomposta dello spazio di Hilbert degli stati quantistici corrispondenti allo stesso sistema classico: esso si suddivide in una somma diretta di **settori di superselezione**, caratterizzati ciascuno dalla rappresentazione di $\pi_1(Q)$ che agisce sugli stati in esso. La regola che impone di usare il principio di sovrapposizione solo per gli stati nello stesso settore è un esempio di regola di superselezione.

Consideriamo ora gli operatori, che vanno sollevati da $L^2(Q)$ a $L^2(\tilde{Q})$ al fine di costruire la meccanica quantistica su \tilde{Q} . Concentriamoci sugli operatori funzione della posizione $\hat{A} = A(\hat{\mathbf{q}})$, essendo quelli che vengono direttamente toccati dalla presenza dei buchi nel dominio; chiaramente sarà possibile realizzare in modo consistente anche le variabili legate all'impulso [3]. Siccome lo spazio delle configurazioni fisico è Q , le osservabili legate alla posizione dovranno essere funzioni ben definite su Q e non su \tilde{Q} . Questi operatori funzione della posizione agiscono moltiplicativamente sulle funzioni d'onda ordinarie $\psi \in L^2(Q)$ ed è naturale assumere che l'azione resti dello stesso tipo. Allora, si soddisfano questi requisiti ponendo la regola

$$(\hat{A}\tilde{\psi})(\tilde{q}) = A(p(\tilde{q})) \cdot \tilde{\psi}(\tilde{q}). \quad (2.3.7)$$

L'equazione (2.3.5) definiva un'azione a destra del gruppo fondamentale $\pi_1(Q)$ sulle funzioni d'onda e questa azione corrispondeva a trasportare il punto in cui la funzione era calcolata lungo un cammino chiuso. Coerentemente, posso definire l'azione di $[\gamma] \in \pi_1(Q)$ sulle funzioni d'onda sollevate $\tilde{\psi}$ come l'effetto che ha su esse portare il punto in cui sono calcolate lungo il cammino su \tilde{Q} in cui si solleva γ , cioè:

$$([\hat{\gamma}]\tilde{\psi})(\tilde{q}) = \tilde{\psi}(\tilde{q} \star [\gamma]) \quad (2.3.8)$$

che, chiaramente, per la (2.3.4) è $a([\gamma])\tilde{\psi}(\tilde{q})$. Si è indicata l'azione col cappuccio $\hat{\cdot}$, come un operatore, per sottolineare il fatto che questa azione sia lineare. Vediamo ora che \hat{A} e $[\hat{\gamma}]$ commutano:

$$\begin{aligned} (\hat{A}[\hat{\gamma}])\tilde{\psi} &= A(p(\tilde{q}))([\hat{\gamma}]\tilde{\psi})(\tilde{q}) \\ &= A(p(\tilde{q}))\tilde{\psi}(\tilde{q} \star [\gamma]); \\ ([\hat{\gamma}]\hat{A})\tilde{\psi} &= (\hat{A}\tilde{\psi})(\tilde{q} \star [\gamma]) \\ &= A(p(\tilde{q} \star [\gamma]))\tilde{\psi}(\tilde{q} \star [\gamma]) \\ &\stackrel{!}{=} A(p(\tilde{q}))\tilde{\psi}(\tilde{q} \star [\gamma]) \\ &= (\hat{A}[\hat{\gamma}])\tilde{\psi}, \end{aligned}$$

dove $\stackrel{!}{=}$ vale perché $p(\tilde{q} \star [\gamma]) = p(\tilde{q})$, siccome l'azione di $\pi_1(Q)$ sul rivestimento \tilde{Q} "permuta" i punti all'interno della stessa fibra. Per il lemma di Schur, ogni operatore che commuta con tutti gli operatori di una rappresentazione irriducibile dev'essere un multiplo dell'identità. Se lo spazio non è semplicemente connesso, in generale $[\tilde{\gamma}]$ non è l'identità perché esistono loop non banali, dunque la rappresentazione degli osservabili sulle funzioni d'onda definita dalla (2.3.7) non è irriducibile. Per ridurla, almeno parzialmente, consideriamo le rappresentazioni irriducibili del gruppo fondamentale $\pi_1(Q)$, che determinano la decomposizione dello spazio di Hilbert $\mathcal{H} = L^2(Q)$ (in rappresentazione di Schrödinger)

$$\mathcal{H} = \bigoplus_j \mathcal{H}^{(j)},$$

in cui gli $\mathcal{H}^{(j)}$ sono i sottospazi di \mathcal{H} che si trasformano secondo la j -esima rappresentazione di $\pi_1(Q)$, che indico con Γ_j . Siccome \hat{A} commuta con $[\hat{\gamma}]$, esso non può mappare un elemento di \mathcal{H} che si trasforma con Γ_j in uno che si trasforma con Γ_l , $l \neq j$ (si ricordi che si ha assunto di avere \hat{A} che commuta con operatori autoaggiunti, per cui i sottospazi invarianti coincidono con gli autospazi ed essendo commutanti hanno gli stessi autospazi), per cui l'immagine soddisfa

$$\hat{A}\mathcal{H}^{(j)} \subseteq \mathcal{H}^{(j)}.$$

Questa significa che l'azione degli operatori su un certo sottospazio dello spazio Hilbert in cui le funzioni d'onda si trasformano tutte secondo la stessa rappresentazione del gruppo fondamentale dà ancora funzioni d'onda che si trasformano allo stesso modo, cioè *l'azione degli operatori su un certo settore di superselezione preserva il settore di superselezione*. Non c'è modo di passare da un settore all'altro tramite l'azione degli operatori fisicamente significativi e si era detto che le ψ in diversi settori non potevano neanche mescolarsi con il principio di sovrapposizione: è evidente che siano **quantizzazioni inequivalenti** dello stesso sistema classico. Concludiamo la sezione enunciando il seguente teorema, già di fatto discusso.

Teorema 2.2. *Le quantizzazioni inequivalenti di un sistema sono tante almeno quante sono le rappresentazioni irriducibili inequivalenti di $\pi_1(Q)$ sullo spazio delle funzioni d'onda.*

2.4 La particella singola sulla circonferenza

Come applicazione della teoria vista nella sezione precedente, studiamo il caso in cui una particella singola senza spin (quindi con funzione d'onda scalare)

è lasciata libera di muoversi su una circonferenza. La circonferenza S^1 è un tipico esempio di spazio connesso ma non semplicemente connesso e ha come gruppo fondamentale $\pi_1(S^1) = \mathbb{Z}$. L'esistenza di più classi distinte di cammini omotopi che congiungono le varie coppie di punti della circonferenza rende ammissibili funzioni d'onda poldrome e per trattarle, come si è detto, è utile costruire una meccanica quantistica sul rivestimento universale che sia poi proiettabile su S^1 .

Richiamiamo la struttura del rivestimento universale di S^1 , già descritta nel primo capitolo. Il rivestimento universale di S^1 è \mathbb{R} , con la proiezione $p(\varphi) = (\cos \varphi, \sin \varphi)$. Nel primo capitolo si era usata come proiezione $(\cos 2\pi\varphi, \sin 2\pi\varphi)$ perché per comodità teorica i cammini erano stati definiti su $[0, 1]$, ma è una richiesta a cui di fatto si può fare a meno e d'ora in poi così sarà. In questo modo, per ogni φ , tutti i punti $\varphi + 2\pi k$, $k \in \mathbb{Z}$ sono proiettati nello stesso punto $x \in S^1$ e le fibre sono questi insiemi discreti $\{\varphi + 2\pi k \mid k \in \mathbb{Z}\}$, ciascuno omeomorfo a $\pi_1(S^1) = \mathbb{Z}$ (per una traslazione), come dev'essere essendo il rivestimento un $\pi_1(S^1)$ -fibrato. I domini fondamentali sono identificabili, ad esempio, con i periodi $\dots [-2\pi, 0], [0, 2\pi], [2\pi, 4\pi], \dots$: ciascuno di essi è omeomorfo a S^1 (servono entrambi gli estremi inclusi e identificati, altrimenti non vale l'omeomorfismo). L'azione di $\pi_1(S^1) = \mathbb{Z}$ su \mathbb{R} è definita come lo spostamento di un punto $\varphi \in \mathbb{R}$ lungo il sollevamento di un loop in S^1 con punto base $x = p(\varphi)$. Se il loop ha indice di avvolgimento N , esso si solleva in un segmento di lunghezza $2\pi N$, dunque l'azione risultante è $\varphi \mapsto \varphi' = \varphi + 2\pi N$.

Le funzioni d'onda sul rivestimento non saranno generiche funzioni $\psi : \mathbb{R} \mapsto \mathbb{C}$, ma vanno imposte le condizioni di proiettabilità (2.3.4): l'azione del gruppo fondamentale sulla funzione d'onda - cioè trasportare il punto φ in cui è calcolata lungo un loop - deve modificare ψ per una rappresentazione unitaria di $\pi_1(S^1) = \mathbb{Z}$. La funzione d'onda è assunta scalare per cui la rappresentazione diventa la moltiplicazione per una fase. Con l'azione descritta sopra, questo vincolo si esprime come

$$\psi(\varphi + 2\pi N) = e^{i\theta(N)}\psi(\varphi),$$

dove la fase $a([\gamma])$ è diventata $e^{i\theta(N)}$, con N l'indice di avvolgimento che etichetta i loop. Bisogna trovare ora quali siano queste fasi, cioè le possibili rappresentazioni di \mathbb{Z} , usando il fatto che $e^{i\theta(N)}$ dev'essere un omomorfismo da \mathbb{Z} a $U(1)$. Il gruppo additivo \mathbb{Z} è identificabile col gruppo ciclico infinito e ha come solo generatore $N = 1$, corrispondente ai loop che compiono un unico giro in senso antiorario sulla circonferenza. Chiamiamo $\theta = \theta(1)$ l'angolo che determina la fase associata a questa classe di loop e assumiamo $\theta \in [0, 2\pi[$,

per evitare ridondanze. L'inverso del generatore è $N = -1$, ed essendo un omomorfismo la corrispondenza che associa la fase si ha che

$$1 \in \mathbb{Z} \mapsto e^{i\theta} \in U(1) \implies -1 \in \mathbb{Z} \mapsto (e^{i\theta})^{-1} = e^{i(-\theta)} \in U(1),$$

l'inverso è associato all'angolo opposto. Essendo \mathbb{Z} generato da 1, ogni altro suo elemento N sarà ottenibile come

$$N = \underbrace{1 + 1 + \dots + 1}_n + \underbrace{(-1) + (-1) + \dots + (-1)}_m;$$

e, sempre per le proprietà di omomorfismo della rappresentazione, sarà

$$\theta(N) = (n - m) \cdot \theta = N \cdot \theta.$$

Questo conclude l'identificazione delle rappresentazioni: si hanno infiniti settori di superselezione e dunque infinite quantizzazioni inequivalenti del problema classico di una particella vincolata ad una circonferenza, anche trascurando le diverse possibili simmetrie interne della particella quantistica. Le funzioni d'onda ottenute hanno una fase polidroma, la quale soddisfa la condizione

$$\psi(\varphi + 2\pi N) = e^{iN\theta} \psi(\varphi), \quad \theta \in [0, 2\pi[\quad (2.4.1)$$

e ogni angolo θ identifica una diversa quantizzazione, in cui la funzione d'onda prende la fase $e^{i\theta}$ prende dopo un cammino chiuso con indice di avvolgimento unitario sulla circonferenza.

Imponendo anche la condizione (2.4.1), una funzione ψ è a quadrato sommabile su \mathbb{R} se e solo se lo è su $[0, 2\pi] \cong S^1$ e si ha una naturale corrispondenza fra lo spazio di Hilbert che si costruirebbe sul rivestimento e quello fisicamente significato per il problema

$$\mathcal{H} = \left\{ \psi : [0, 2\pi] \rightarrow \mathbb{C} \mid \langle \psi, \psi \rangle < +\infty \right\}, \quad \langle \psi, \phi \rangle = \int_0^{2\pi} (\psi^* \phi)(\varphi) d\varphi.$$

Grazie a questa corrispondenza è possibile restringersi senza problemi a lavorare solo su un dominio fondamentale del rivestimento.

È intuitivo che debba essere soddisfatta la (2.4.1), perché i valori misurabili delle osservabili devono essere funzioni di S^1 e funzioni del rivestimento ed essi dipendono solo da $\psi^* \psi$. Ponendo che non ci siano ambiguità delle ampiezze di probabilità nei punti di raccordo dei domini fondamentali equivale a richiedere che

$$(\psi^* \psi)(2\pi N^+) = (\psi^* \psi)(2\pi(N+1)^-) \quad \forall N \in \mathbb{Z},$$

e soddisfarla porta a imporre, appunto, che le ψ possano al massimo variare di una fase. Questa condizione è dunque da aggiungere al dominio di tutti gli operatori e operatori non definibili compatibilmente con questa richiesta non saranno da considerare osservabili su S^1 . Lavorare solo su un dominio fondamentale del rivestimento porta a considerare un unico punto di raccordo, ad esempio $0^+ \sim 2\pi^-$, e i due corrispondenti punti del rivestimento sono legati dall'azione di un loop con indice di avvolgimento uno. Sul dominio $[0, 2\pi]$ la condizione di proiettabilità è

$$\psi(2\pi^-) = e^{i\theta}\psi(0^+), \quad (2.4.2)$$

e sarà da aggiungere al dominio degli operatori. Otterrò domini \mathcal{D}_θ , che sono dipendenti dall'angolo e dunque differenti in ogni settore di superselezione. Ad esempio, l'hamiltoniana della particella libera è $\hat{H} = \hat{\mathbf{p}}^2/2m$; in rappresentazione di Schrödinger $\hat{\mathbf{p}} = -i\hbar \partial/\partial\varphi$ (assumendo unitario il raggio della circonferenza) e diventa

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial\varphi^2}.$$

Nel dominio \mathcal{D}_θ dell'operatore hamiltoniano andrà posta la richiesta (2.4.2), dunque sarà $\mathcal{D}_\theta \subset \mathcal{H}$: l'operatore non può essere definito su tutto lo spazio di Hilbert. La (2.4.2) basterebbe a rendere l'operatore impulso autoaggiunto, ma non è sufficiente per l'autoaggiunzione di \hat{H} , per cui bisogna anche specificare il comportamento della derivata prima. Si considerino per ogni θ gli insiemi

$$\mathcal{D}_\theta = \left\{ \psi \in \mathcal{H} \mid \partial_\varphi^2 \psi \in L^2([0, 2\pi[), \psi(2\pi^-) = e^{i\theta}\psi(0^+), \right. \\ \left. \partial_\varphi \psi(2\pi^-) = e^{i\theta} \partial_\varphi \psi(0^+) \right\}, \quad (2.4.3)$$

in cui si è abbreviato $\partial_\varphi \equiv \partial/\partial\varphi$. Essi sono sottospazi di \mathcal{H} densamente definiti, per cui è possibile definire l'operatore aggiunto \hat{H}^\dagger , e su questi domini $\hat{H} = \hat{H}^\dagger$, come si può verificare facilmente con una doppia integrazione per parti: sono questi i domini di autoaggiunzione dell'operatore hamiltoniano coerenti con i vincoli sulla fase per la proiettabilità su S^1 . Esistono infiniti stati in $\mathcal{H} \setminus \mathcal{D}_\theta$, ma si può dimostrare [13] che su essi il valore atteso di $\hat{\mathbf{p}}^2$ è infinito: per eccitarli servirebbe un'energia infinita e non possono essere osservabili su una circonferenza. Questo mostra come la condizione (2.4.2) permetta di distinguere la meccanica quantistica sulla circonferenza da quella sul segmento. [4]

L'angolo θ , che etichetta i settori di superselezione, interviene negli autovalori delle osservabili, siccome determina il dominio degli operatori associati. Dunque è un angolo che si può misurare: dato un sistema certo sistema, è

misurabile quale sia la diversa quantizzazione dello stesso sistema classico che in esso si realizza. Per calcolare lo spettro si risolve l'equazione agli autovalori $\hat{H}\psi = E\psi$ associata all'operatore hamiltoniano, che è l'equazione di Schrödinger indipendente dal tempo

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial \varphi^2}(\varphi) = E\psi(\varphi), \quad \varphi \in [0, 2\pi];$$

la soluzione generale dell'equazione è banale ed è

$$\psi(\varphi) = A \cdot \exp\left(i \frac{\sqrt{2mE}}{\hbar} \varphi\right) + B \cdot \exp\left(-i \frac{\sqrt{2mE}}{\hbar} \varphi\right).$$

Impongo ora le condizioni al contorno, descritte nel dominio (2.4.3), sulla funzione d'onda ψ e sulla sua derivata $\partial_\varphi \psi$

$$\psi(2\pi) = e^{i\theta} \psi(0), \quad \frac{\partial \psi}{\partial \varphi}(2\pi) = e^{i\theta} \frac{\partial \psi}{\partial \varphi}(0);$$

queste condizioni portano al sistema lineare

$$\begin{cases} A + B = \exp\{-i(\theta - 2\pi\Omega)\}A + \exp\{-i(\theta + 2\pi\Omega)\}B \\ A - B = \exp\{-i(\theta - 2\pi\Omega)\}A - \exp\{-i(\theta + 2\pi\Omega)\}B \end{cases},$$

in cui si è posto per semplificare la notazione $\Omega = \frac{\sqrt{2mE}}{\hbar}$. Il sistema è risolto se sono soddisfatte le relazioni

$$A = A \exp\{-i(\theta - 2\pi\Omega)\} \quad \text{e} \quad B = B \exp\{-i(\theta + 2\pi\Omega)\},$$

le quali, ponendo uguali a uno gli esponenziali, portano alle seguenti condizioni di quantizzazione su Ω (e di conseguenza sull'energia):

$$\Omega_n = n + \frac{\theta}{2\pi}, \quad \Omega_{n'} = -n' - \frac{\theta}{2\pi},$$

dove n, n' sono numeri interi. Ricordando che $\theta \in [0, 2\pi[$, è chiaro che la prima possa essere risolta solo per $n = 0, 1, 2, 3, \dots$ e la seconda solo per $n' = \dots, -3, -2, -1,$, visto che $\Omega \geq 0$. Si possono allora indicizzare le autofunzioni con un solo numero intero n che varia su tutto \mathbb{Z} , ottenendo così, a parte la normalizzazione,

$$\psi(\varphi) = C \cdot \exp\left(i\left(n + \frac{\theta}{2\pi}\right)\varphi\right), \quad n \in \mathbb{Z}. \quad (2.4.4)$$

Siccome $n + \frac{\theta}{2\pi} = \Omega_n = \frac{\sqrt{2mE_n}}{\hbar}$, gli autovalori dell'energia sono

$$E_n = \frac{\hbar^2}{2m} \left(n + \frac{\theta}{2\pi} \right)^2, \quad n \in \mathbb{Z}. \quad (2.4.5)$$

Lo spettro risulta dipendente da θ e dunque l'angolo è misurabile, come era stato precedentemente affermato.

Per vedere i problemi della fase caratteristici dell'effetto Aharonov - Bohm, si era utilizzata una trasformazione di gauge singolare che aveva ricondotto il problema originario con potenziale vettore ad uno equivalente senza potenziale vettore ma con fase polidroma. È interessante vedere che questo problema possa essere condotto ad una sorta di effetto Aharonov-Bohm sulla circonferenza, usando una trasformazione di gauge analoga alla (2.2.4):

$$\psi(\varphi) \mapsto \bar{\psi}(\varphi) = e^{-\frac{i\theta}{2\pi}\varphi} \psi(\varphi).$$

Questa nuova funzione d'onda soddisfa non più la condizione al contorno $\psi(2\pi) = e^{i\theta}\psi(0)$, ma è diventata una funzione continua, perché

$$\bar{\psi}(2\pi) = e^{-\frac{i\theta}{2\pi}\varphi} \psi(\varphi) |_{\varphi=2\pi} = e^{-i\theta} \psi(2\pi) = e^{-i\theta} e^{i\theta} \psi(0) = \psi(0) = \bar{\psi}(0).$$

Questa trasformazione è della forma della (2.2.4), con $\frac{e}{\hbar}\Lambda(\varphi) = -\frac{\theta}{2\pi}\varphi$, per cui induce una trasformazione del potenziale vettore data dalla formula (2.2.3):

$$\bar{A} = A + \frac{\partial\Lambda}{\partial\varphi}, \quad \Lambda = -\frac{\theta\hbar}{2\pi e}\varphi \Rightarrow \bar{A} = -\frac{\theta\hbar}{2\pi e}.$$

Partendo da un problema con condizioni al contorno di discontinuità (corrispondenti ad una funzione d'onda naturalmente polidroma) e senza interazione, lo si è ricondotto ad un problema con condizioni al contorno di continuità ma con una interazione descritta da un potenziale vettore. Per un'altra situazione la conclusione sarebbe analoga: tutti i problemi con fase polidroma sono riconducibili a un problema simile all'effetto Aharonov-Bohm, in cui la particella porta una carica e e si accoppia con un potenziale vettore \vec{A} , con flusso magnetico non banale.

Capitolo 3

Particelle identiche

Nel capitolo precedente è stata descritta la procedura per sviluppare una teoria quantistica dei sistemi il cui spazio delle configurazioni non è semplicemente connesso e si è mostrato come primo esempio di applicazione la particella senza spin su una circonferenza. Questo capitolo è dedicato ad un caso di grande importanza, perché si presenta sistematicamente in meccanica quantistica: quando si hanno particelle identiche, gli stati collettivi delle particelle dovranno tenere conto di questa identità e andranno fatte opportune identificazioni che, spesso, portano a uno spazio delle configurazioni non semplicemente connesso. Discuteremo prima come costruire il corretto spazio delle configurazioni nella sezione §3.1 e successivamente si faranno alcuni esempi. Nella §3.2 studieremo come emerge la separazione delle particelle in bosoni e fermioni, e vedremo che esiste una ulteriore classe di particelle, dette anioni, le quali seguono statistiche che interpolano quelle bosoniche e fermioniche; discuteremo poi nella sezione §3.3 il caso di due particelle sulla circonferenza.

3.1 Lo spazio delle configurazioni per particelle indistinguibili

Nella fisica classica, il concetto di identità degli oggetti è una approssimazione ideale, perché tutti i corpi macroscopici sono diversi per almeno qualche piccola caratteristica. Spesso diversi oggetti del sistema in esame sono considerati identici per trascurare queste differenze, talmente piccole che si suppone abbiano un ruolo trascurabile nella dinamica. È sufficiente pensare al concetto di punto materiale, per vedere quanto siano utilizzate le particelle identiche in fisica classica. Tuttavia, anche da identiche, le particelle classiche rimangono sempre distinguibili, grazie all'esistenza del concetto di

traiettoria. Siccome l'evoluzione è univoca fissate posizione e velocità di una particella, posso etichettare le particelle tramite posizione e velocità di ciascuna a un qualche istante di tempo; negli istanti successivi, una misura di posizione e di velocità di una particella permetterà sempre di riconoscere con quale delle particelle etichettate inizialmente si ha a che fare.

Nella meccanica quantistica, l'identità delle particelle non è più una comoda approssimazione ma è un'evidenza sperimentale. Non sono mai state effettuate misure capaci di osservare differenze tra due elettroni e similmente per tutte gli altri tipi di particelle. Oltre ad emergere il concetto di identità totale, si perde anche la possibilità di distinguere particelle identiche. Un suggerimento a questo è dato dalle indeterminazioni quantistiche, che distruggono il concetto di traiettoria e più in generale di univocità del risultato di una misura su uno stato che evolve nel tempo; ciò impedisce di realizzare un'etichettatura coerente nel tempo come si faceva nella meccanica classica. Lo sviluppo della teoria quantistica ha mostrato come sia necessario assumere l'**indistinguibilità** delle particelle identiche, pena la deduzione di risultati assurdi. Un esempio è il famoso paradosso di Gibbs (per dettagli si veda ad esempio [8]), che mostra come la distinguibilità delle particelle identiche di un gas porterebbe ad un aumento di entropia, tipico di un processo irreversibile, in un certo processo che dovrebbe essere invece perfettamente reversibile. La soluzione al paradosso di Gibbs è proprio supporre le particelle quantistiche indistinguibili: questo porta a recuperare la reversibilità di tale processo. È l'ipotesi di indistinguibilità delle particelle identiche che conduce alla separazione delle particelle in bosoni e fermioni, per cui le innumerevoli osservazioni di queste due classi di particelle hanno ampiamente confermato la correttezza dell'ipotesi: l'indistinguibilità delle particelle identiche diventa un postulato della meccanica quantistica.

Avendo sottolineato l'importanza del concetto di indistinguibilità in meccanica quantistica, diventa chiara la necessità di tenerne conto per la costruzione di corrette funzioni d'onda che descrivano un sistema di particelle identiche. Storicamente si postulava che le funzioni d'onda fisicamente significative dovessero essere simmetriche o totalmente antisimmetriche rispetto allo scambio di particelle, ma, seguendo l'approccio topologico introdotto negli anni '70 da Leinaas e Myrheim in [9], si vede che, operando opportunamente sullo spazio delle configurazioni, non è necessario alcun postulato per ricostruire la separazione delle particelle in bosoni e fermioni. In più questo approccio permette di notare come la necessità di simmetria o antisimmetria delle funzione d'onda sia legata esclusivamente al fatto che il nostro mondo sia tridimensionale e che in due dimensioni questo vincolo scompaia. Consideriamo allora un sistema di N particelle identiche che possono muoversi

in un insieme $\mathcal{D} \subseteq \mathbb{R}^d$. Lo spazio delle configurazioni Q_N del sistema non è semplicemente \mathcal{D}^N , perché contiene N-uple "sovrabbondanti" di coordinate $\mathbf{q} = (\vec{q}_1, \dots, \vec{q}_N)$: siccome le particelle sono indistinguibili, permutarle dovrebbe dare lo stesso stato, descritto dalle stesse coordinate, mentre una N-upla $\mathbf{q} \in \mathcal{D}^N$ distingue fra le due configurazioni. Bisogna lavorare su Q_N in modo che sia contenuta già in esso l'identità di alcune configurazioni.

Essendo le particelle indistinguibili, un qualsiasi scambio delle loro posizioni deve essere la stessa configurazione e quindi produrre lo stesso stato quantistico: data una N-upla $(\vec{q}_1, \dots, \vec{q}_N)$. devo identificare con una relazione di equivalenza tutte le N-uple che differiscono per permutazioni. Formalmente la relazione sulle N-uple è:

$$\mathbf{q} \sim \mathbf{q}' \iff \exists \sigma \in S_N : \mathbf{q}' = \sigma \mathbf{q},$$

dove S_N è il gruppo delle permutazioni, o *gruppo simmetrico*, e l'azione di un suo elemento $\sigma \in S_N$, $\sigma : \{1, 2, \dots, N\} \rightarrow \{1, 2, \dots, N\}$ su una N-upla $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_N) \in \mathcal{D}^N$ è $\sigma \mathbf{q} = (\mathbf{q}_{\sigma(1)}, \dots, \mathbf{q}_{\sigma(N)})$, corrispondente a un generico scambio di posizioni. Questa identificazione ci suggerisce di usare come spazio delle configurazioni $Q_N = \mathcal{D}^N / \sim = \mathcal{D}^N / S_N$; fisicamente questo potrebbe già essere lo spazio corretto, ma può presentare complicazioni tecniche che portano spesso ad una ulteriore modifica anche fisicamente ragionevole. Vediamo prima un esempio.

Consideriamo il caso di due particelle in $\mathcal{D} = \mathbb{R}$. In questo caso $\mathbf{q} = (q_1, q_2) \in \mathbb{R}^2$, $S_2 = \{e, \sigma\}$ (con $e = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ e $\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$); due coppie \mathbf{q}, \mathbf{q}' sono equivalenti con \sim se sono uguali a meno di uno scambio delle due componenti e $Q_2 = \mathbb{R}^2 / S_2$ si ottiene identificando i punti di \mathbb{R}^2 simmetrici rispetto alla bisettrice di primo e terzo quadrante, che è la diagonale $\Delta = \{(q_1, q_2) \in \mathbb{R}^2 \mid q_1 = q_2\}$. Allora lo spazio delle configurazioni può essere visto come uno dei due semipiani di cui Δ divide \mathbb{R}^2 . Lo spazio delle configurazioni costruito, partendo da \mathbb{R}^2 , che non ha frontiera, ha "sviluppato" una frontiera Δ non nulla, costituita dai punti fissi della trasformazione di simmetria in cui si è tradotta la relazione di equivalenza. Questa è una situazione tipica dell'identificazione fatta e le frontiere di punti uniti che si sviluppano possono risultare patologiche, per cui tipicamente vengono escluse. La rimozione fatta ha ragioni sia di tipo matematico che fisico. Matematicamente è un problema perché danno regioni singolari di dimensione inferiore che sono tecnicamente complicate da trattare. La ragione fisica che porta alla loro esclusione è che i punti uniti sotto l'azione delle permutazioni sono i punti in cui ci sono particelle la cui posizione è esattamente coincidente, e tendenzialmente si esclude il fatto che due particelle distinte possano "attraversarsi", perché ci sarà una qualche interazione ad impedire tale scenario.

Gli argomenti descritti ci portano ad identificare il corretto *spazio delle configurazioni per N particelle indistinguibili* confinate in $\mathcal{D} \subseteq \mathbb{R}^d$ con

$$Q_N = (\mathcal{D}^N \setminus \Delta_{d,N})/S_N, \quad (3.1.1)$$

dove la diagonale $\Delta_{d,N}$ generalizza la Δ vista nel caso precedente come

$$\Delta_{d,N} = \{(\vec{q}_1, \dots, \vec{q}_N); \vec{q}_i \in \mathbb{R}^d \forall i \mid \exists i, j; i \neq j \text{ per cui } \vec{q}_i = \vec{q}_j\}.$$

Questo spazio delle configurazioni in generale non è semplicemente connesso e per descrivere la meccanica quantistica su di esso è necessario ricorrere alla trattazione descritta nel capitolo precedente.

3.2 Le statistiche su \mathbb{R}^d : bosoni, fermioni e anioni

Vogliamo discutere come si ricavi la funzione d'onda che descriva correttamente un sistema di N particelle libere di muoversi in \mathbb{R}^d . La (3.1.1) ci dice che lo spazio delle configurazioni del sistema è $Q_N = (\mathbb{R}^{d \cdot N} \setminus \Delta_{d,N})/S_N$ e il primo passo per costruire una teoria quantistica su di esso è identificarne il gruppo fondamentale. Se il gruppo fondamentale è banale, lo spazio è semplicemente connesso e la quantizzazione procede con funzioni d'onda monodrome e con le ordinarie condizioni al contorno. Se $\pi_1(Q_N) \neq \{e\}$, invece, si devono usare funzioni d'onda definite sul rivestimento universale di Q_N , imponendo su esse le condizioni (2.3.4), che vincolano la polidromia della fase della funzione d'onda proiettata sullo spazio delle configurazioni effettivo Q_N .

Prima di procedere con l'identificazione dei gruppi fondamentali di Q_N (non è detto sia lo stesso per ogni N) è utile richiamare le proprietà del gruppo delle permutazioni e del gruppo delle trecce. Come già detto, il gruppo delle permutazioni di N oggetti S_N è un gruppo finito di $N!$ elementi, ognuno dei quali corrisponde ad una permutazione, che è una biiezione sull'insieme dei primi N naturali $\sigma: \{1, 2, \dots, N\} \rightarrow \{1, 2, \dots, N\}$, cioè un qualsiasi scambio dell'ordine di N oggetti. Sono spesso indicate con $\binom{1, 2, \dots, N}{\sigma(1), \sigma(2), \dots, \sigma(N)}$, dove la prima riga indica l'ordine di partenza e la seconda quello finale. L'operazione interna è la composizione di funzioni, l'identità è $e = \binom{1, 2, \dots, N}{1, 2, \dots, N}$ e chiaramente ogni permutazione ammette inversa. Un esempio importante di permutazione sono le *trasposizioni*, cioè le permutazioni che scambiano solamente due posizioni, lasciando le altre inalterate. Com'è intuitivo, si può dimostrare che con opportuni scambi di due oggetti alla volta è possibile ottenere uno scambio qualsiasi di N oggetti: *ogni permutazione si può esprimere come*

composizione di trasposizioni. Questa proprietà vale ancora se ci si restringe alle trasposizioni σ_i che scambiano l' i -esimo oggetto con l' $i+1$ -esimo oggetto. Volendo dare una presentazione astratta del gruppo, consideriamo gli N oggetti $\{e, \sigma_1, \dots, \dots, \sigma_{N-1}\}$, che composti arbitrariamente danno sì tutte le permutazioni, ma è necessario introdurre alcune identificazioni, perché altrimenti gli oggetti generati sarebbero sovrabbondanti (darebbe un gruppo libero). Si può dimostrare [13] che le relazioni sui generatori σ_i necessarie e sufficienti per ottenere il gruppo delle permutazioni di N oggetti sono:

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}, \quad (3.2.1a)$$

$$\sigma_i \sigma_j = \sigma_j \sigma_i \text{ se } |i - j| \geq 2, \quad (3.2.1b)$$

$$\sigma_i^2 = e. \quad (3.2.1c)$$

Queste relazioni hanno una interpretazione immediata in termini di scambi di oggetti; in particolare, la seconda significa che due trasposizioni di oggetti continui σ_i sono indipendenti e commutano se non c'è un oggetto in comune tra le due e la terza, che vale in generale per tutte le trasposizioni, dice che ogni trasposizione è l'inversa di se stessa. Ponendo queste relazioni sul gruppo libero generato da $\{e, \sigma_1, \dots, \dots, \sigma_{N-1}\}$ ottengo il gruppo simmetrico S_N , la cui *presentazione* è scritta come

$$S_N = \langle e, \sigma_1, \dots, \sigma_{N-1} \mid \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}; \\ \sigma_i \sigma_j = \sigma_j \sigma_i \text{ se } |i - j| \geq 2; \sigma_i^2 = e \rangle \quad (3.2.2)$$

In particolare, è la relazione $\sigma_i^2 = e$ che rende il gruppo simmetrico un gruppo finito. Le catene del gruppo libero generato da $\{e, \sigma_1, \dots, \dots, \sigma_{N-1}\}$ sarebbero composizione di oggetti tipo σ_i^m , $m \in \mathbb{Z}$ e tutte differenti, $\sigma_i^m \neq \sigma_i^{m'} \forall m' \neq m$, infatti i gruppi liberi sono infiniti; la (3.2.1c) rende invece σ_i^m soltanto uguale a σ_i o all'identità e , a seconda della parità di m , e il gruppo delle permutazioni diventa finito.

Rimuovere la condizione $\sigma_i^2 = e$ dalla presentazione (3.2.2) fa passare dal gruppo delle permutazioni di N oggetti a un gruppo infinito, che è il gruppo delle trecce di N stringhe B_N :

$$B_N = \langle e, \sigma_1, \dots, \sigma_{N-1} \mid \sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1}; \\ \sigma_i \sigma_j = \sigma_j \sigma_i \text{ se } |i - j| \geq 2 \rangle. \quad (3.2.3)$$

Questo gruppo ha le stesse proprietà di indipendenza del risultato dall'ordine degli scambi che ha il gruppo delle permutazioni, ma $\sigma_i^2 \neq e$, $\sigma_i^3 \neq \sigma_i$ e così via. Ciascun elemento di B_N può essere visualizzato mediante, appunto, una treccia formata da N corde. Una treccia non è univocamente determinata

dagli ordinamenti iniziale e finale degli estremi, com'era per le permutazioni, ma dipende anche dalle trasformazioni eseguite per mandare la configurazione di partenza in quella di arrivo: la treccia "ricorda" la storia degli scambi eseguiti. La Figura 3.1 riporta un esempio grafico di quanto appena detto.

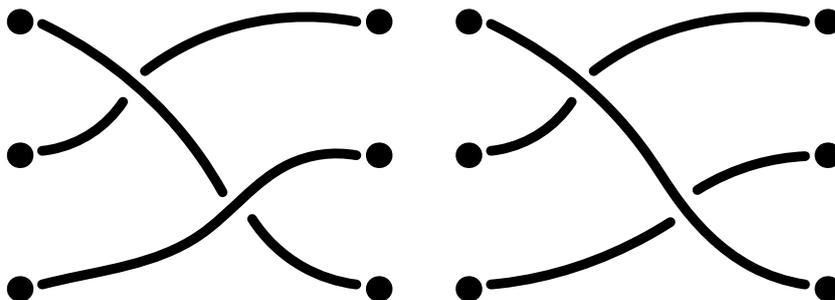


Figura 3.1: Rappresentazione di due trecce di $N=3$ stringhe con stessa modifica dell'ordinamento dei punti iniziali, ma non equivalenti perché ottenute con scambi fatti accavallando le stringhe in modo diverso.

Su B_N è possibile identificare σ_i con l'operazione di scambio della stringa i -esima con la stringa $i+1$ -esima *facendo passare sopra* (o anche sotto, basta fissare una scelta) la i -esima sulla $i+1$ -esima e σ_i^{-1} diventa scambiare la stringa i -esima con la stringa $i+1$ -esima facendo passare sotto (o sopra) la i -esima; $\sigma_i \neq \sigma_i^{-1}$ perché sono scambi realizzati sovrapponendo le corde in modo opposto. A questo punto, possiamo enunciare il risultato fondamentale di questa sezione.

Teorema 3.1. *Lo spazio delle configurazioni $Q_N = (\mathbb{R}^{d \cdot N} \setminus \Delta) / S_N$ di un sistema di N particelle identiche in \mathbb{R}^d ha per gruppo fondamentale*

$$\pi_1(Q_N) = S_N \quad \text{se } d \geq 3, \quad (3.2.4a)$$

$$\pi_1(Q_N) = B_N \quad \text{se } d = 2. \quad (3.2.4b)$$

Dimostrazione. Non si darà qui una dimostrazione completa del risultato, ma piuttosto un'idea intuitiva del perché sia così. Un loop su Q_N è ad esempio una traiettoria in cui tutte le particelle ritornano al punto di partenza in \mathbb{R}^d , ma non è l'unica possibilità: siccome lo spazio è quotientato rispetto al gruppo simmetrico, sono identificate tutte le configurazioni che differiscono per permutazioni e si ritorna al punto di partenza in Q_N anche se le particelle si scambiano di posto in qualsiasi modo. I cammini chiusi corrispondono a diversi scambi di particelle; un loop in cui le particelle non vengono scambiate è identificabile con l'azione di e , mentre un loop in cui la i -esima particella è scambiata con la $i+1$ -esima è identificabile con l'azione di σ_i . Queste

identificazioni sono sulle classi di omotopia di questi cammini, perché traiettorie con stessi punti iniziali e finali di \mathbb{R}^n sono deformabili con continuità l'una nell'altra. I loop "elementari", che componendoli permettono di dare tutti gli altri, saranno naturalmente quelli identificati con e con i σ_i , visto che i cammini chiusi sono direttamente identificabili con scambi di particelle. Verifichiamo prima la proprietà (3.2.1a), indicando con \oplus, \otimes, \odot le particelle i -esima, $i+1$ -esima e $i+2$ -esima, che intervengono nel considerare l'azione di $\sigma_i\sigma_{i+1}\sigma_i$ e $\sigma_{i+1}\sigma_i\sigma_{i+1}$.

$$\sigma_i\sigma_{i+1}\sigma_i : (\oplus, \otimes, \odot) \xrightarrow{\sigma_i} (\otimes, \oplus, \odot) \xrightarrow{\sigma_{i+1}} (\otimes, \odot, \oplus) \xrightarrow{\sigma_i} (\odot, \otimes, \oplus),$$

mentre

$$\sigma_{i+1}\sigma_i\sigma_{i+1} : (\oplus, \otimes, \odot) \xrightarrow{\sigma_{i+1}} (\oplus, \odot, \otimes) \xrightarrow{\sigma_i} (\odot, \oplus, \otimes) \xrightarrow{\sigma_{i+1}} (\odot, \otimes, \oplus);$$

com'è evidente danno lo stesso risultato. In modo del tutto analogo è possibile verificare anche la (3.2.1b). Valgono entrambe (indipendentemente da d) proprio perché i loop sono stati identificati con degli scambi, resta da verificare la (3.2.1c), che permette di distinguere se siano scambi che dipendono solo dalla configurazione finale o scambi che "ricordano" anche la traiettoria fatta per arrivare alla configurazione finale. Consideriamo un qualsiasi loop etichettato da σ_i , che scambia la particella i -esima e la particella $i+1$ -esima dopo aver percorso una certa traiettoria in \mathbb{R}^d ; un loop etichettato da σ_i^{-1} scambia le stesse particelle, ma percorrendo la stessa traiettoria in verso opposto. Questi loop possono essere rappresentati, a meno di omotopie, come nella Figura 3.2, in cui sono raffigurati giacenti su un piano. Se $d \geq 3$, il piano su cui giacciono i loop è immerso in uno spazio con almeno una dimensione in più ed è possibile trasformare uno nell'altro tramite una rotazione che ha per asse la congiungente delle posizioni iniziali e finali delle due particelle. Questa rotazione è chiaramente un'omotopia, per cui $\sigma_i \sim \sigma_i^{-1} \Rightarrow \sigma_i^2 \sim e$, per cui sono verificate tutte le relazioni tra i generatori della (3.2.2) e il gruppo fondamentale è S_N . L'argomentazione non vale più in due dimensioni: manca $\sigma_i^2 \sim e$, per cui il gruppo fondamentale non è più il gruppo simmetrico ma il gruppo delle trecce. □

Nel teorema non si è fatto riferimento al caso $d = 1$, perché non è particolarmente interessante in questo contesto: lo spazio delle configurazioni di un sistema di N particelle vincolate su \mathbb{R} è effettivamente semplicemente connesso e in esso non si presentano problemi di questo tipo. È uno spazio semplicemente connesso perché i loop non banali su Q_N corrispondono agli scambi di particelle, ma su una retta non è possibile scambiare due particelle

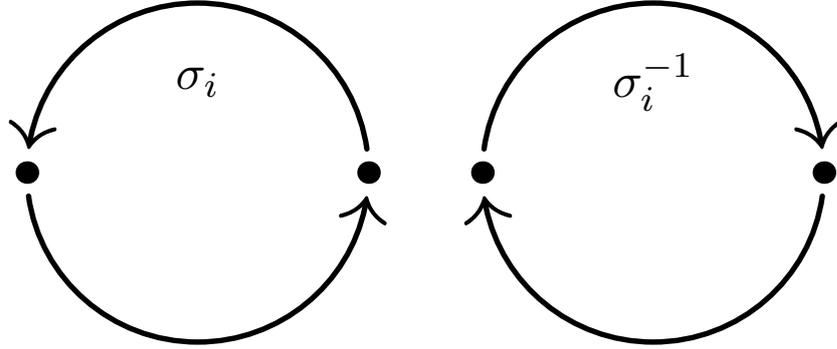


Figura 3.2: A sinistra il loop corrispondente al generatore di S_n σ_i , che scambia in verso antiorario la particella i -esima con la $i+1$ -esima, e a destra il loop corrispondente all'inverso del generatore σ_i^{-1} .

a meno che esse non si attraversino. Questo rende differente la trattazione di particelle identiche in una dimensione.

Per trovare la corretta funzione d'onda e studiare la dinamica di particelle identiche è quindi necessario separare i casi $d \geq 3$ e $d = 2$, siccome sono diversi i gruppi fondamentali. Come si era visto precedentemente, la strada da seguire è costruire la meccanica quantistica sul rivestimento universale, che in questo caso è $\mathbb{R}^{d \cdot N}$, ottenendo una funzione d'onda "estesa" $\tilde{\psi}(\tilde{q})$ su cui imporre le condizioni di proiettabilità sullo spazio delle configurazioni effettivo per rappresentare la funzione d'onda effettiva $\psi(q)$. Queste condizioni vogliono che $\tilde{\psi}(\tilde{q})$ cambi per azione di una rappresentazione unidimensionale unitaria $a([\gamma])$ del gruppo fondamentale quando si compie un loop su Q_N (cioè quando si scambiano particelle), per cui vanno identificate le possibili $a([\gamma])$. Ogni elemento del gruppo fondamentale è esprimibile come un opportuno prodotto dei suoi generatori e dei loro inversi e le rappresentazioni sono omomorfismi (soddisfano $a([\gamma] \cdot [\delta]) = a([\gamma])a([\delta])$), per cui è sufficiente trovare le immagini dei generatori σ_i e dei loro inversi. Per la condizione di proiettabilità $a([\gamma])$ dev'essere un fattore di fase $U(1)$, allora $a([\gamma]^{-1}) = a([\gamma])^{-1} = a([\gamma])^*$ e restano da trovare solo le immagini dei generatori e per farlo si traducono in vincoli sulle fasi le relazioni delle presentazioni dei due gruppi (3.2.2) e (3.2.3), sfruttando il fatto che la mappa $[\gamma] \mapsto a([\gamma])$ è un omomorfismo. Ponendo l'angolo θ_j tale che $a(\sigma_j) = e^{-i\theta_j}$, dalla relazione (3.2.1a) si ha che

$$\sigma_j \sigma_{j+1} \sigma_j = \sigma_{j+1} \sigma_j \sigma_{j+1} \implies e^{-i\theta_j} e^{-i\theta_{j+1}} e^{-i\theta_j} = e^{-i\theta_{j+1}} e^{-i\theta_j} e^{-i\theta_{j+1}},$$

cioè $e^{-i(2\theta_j + \theta_{j+1})} = e^{-i(2\theta_{j+1} + \theta_j)} \iff e^{-i\theta_j} = e^{-i\theta_{j+1}}$, da cui otteniamo che $\theta_j = \theta_{j+1} \pmod{2\pi}$, per cui, assumendo di lavorare con angoli in $[0, 2\pi[$, si

può utilizzare un unico angolo θ tale che $a(\sigma_j) = e^{-i\theta}$, indipendentemente da j . La relazione (3.2.1b) non pone nessun vincolo, perché il gruppo $U(1)$ è di per sé abeliano. Resta da considerare la (3.2.1c), che distingue S_N da B_N :

$$\sigma_j^2 = e \implies a(\sigma_j^2) = a(\sigma_j)^2 = e^{i2\theta} = a(e) = 1 ,$$

per cui restano possibili solamente $\theta = 0$ o $\theta = \pi$ se $d \geq 3$, mentre se $d = 2$ non si ha questo vincolo e sono ammessi tutti i $\theta \in [0, 2\pi[$, purché indipendenti dal generatore. Quelle trovate sono le immagini dei generatori, da cui è possibile dedurre quelle di qualsiasi elemento del gruppo. Per S_N ogni permutazione è combinazione delle trasposizioni σ_j e ottengo due casi. Il primo si ha per $\theta = 0$, per cui qualsiasi composizione continua a dare $a(p) = e^{i \cdot 0} = 1$. Se $\theta = \pi$, ogni scambio elementare σ_j porta un fattore $e^{i\pi} = (-1)^{n(p)}$, dove $n(p)$ è il numero di trasposizioni necessario per comporre un generico elemento $p \in S_N$. In realtà $n(p)$ non è unico, ma tutti i suoi possibili valori hanno la stessa parità, per questo $a(p) = (-1)^{n(p)}$ è ben definito per ogni p , per cui vale $+1$ o -1 . Questo numero è anche detto *segno* della permutazione. Riassumendo, in uno spazio tridimensionale (o più) una traiettoria corrispondente ad una permutazione p delle particelle identiche modifica la funzione d'onda in uno dei due seguenti modi:

$$\psi \xrightarrow{p} a(p)\psi = \psi \quad \forall p , \text{ oppure} \quad (3.2.5a)$$

$$\psi \xrightarrow{p} a(p)\psi = \text{sgn}(p) \cdot \psi \quad (3.2.5b)$$

In uno spazio bidimensionale, invece, una traiettoria non corrisponde più ad una permutazione ma ad una treccia. Una treccia elementare σ_j può dare un qualsiasi θ , per cui $\psi \xrightarrow{\sigma_j} e^{-i\theta}\psi$; l'inversa di una treccia elementare corrisponde dà $a(\sigma_i^{-1}) = a(\sigma_i)^{-1} = e^{i\theta}$. Allora, una traiettoria su Q_N corrispondente ad una generica treccia $b \in B_N$ modifica la funzione d'onda del sistema di particelle identiche come

$$\psi \xrightarrow{b} a(b)\psi = e^{i \sum(\pm\theta)}\psi = e^{im(b)\theta}\psi , \quad (3.2.6)$$

dove $m(b)$ è la differenza tra il numero di trecce elementari tipo σ_j (che danno θ) e il numero di trecce elementari inverse tipo σ_j^{-1} (che danno $-\theta$) che servono a comporre b .

Ricapitolando, in $d \geq 3$ dimensioni si è trovato che esistono soltanto due possibili rappresentazioni unitarie di $\pi_1(Q_N)$, per cui ci possono essere solo due settori di superselezione in cui le particelle sono distinte. Quantizzando come nel settore con $\theta = 0$, un qualsiasi scambio delle particelle identiche non produce alcun effetto sulla funzione d'onda, mentre, quantizzando come

nel secondo settore con $\theta = \pi$, uno scambio elementare cambia il segno della funzione d'onda. Si è ritrovato un risultato è ben noto: queste due classi di particelle corrispondono chiaramente ai **bosoni** e ai **fermioni**, che sono le classi ordinarie emergenti dalla quantizzazione di un sistema di particelle identiche.

La sorpresa si ha nello spazio a due dimensioni: esistono infiniti settori di superselezione e infinite quantizzazioni inequivalenti, etichettate dall'angolo θ , per cui le statistiche bosoniche e fermioniche sono interpolate da infinite altre classi di particelle intermedie, la cui funzione d'onda può prendere una fasi qualsiasi fissata sotto uno scambio. Queste particelle sono i cosiddetti **anioni**, parola riadattata dall'inglese "anyons", per sottolineare il fatto che sia ammessa una qualsiasi (appunto "any", in inglese) fase. Siccome la fase caratteristica del settore di superselezione interviene nella dinamica, gli anioni sono anche dinamicamente distinguibili da ordinari bosoni o fermioni. Oltre a questo, differiscono anche per le simmetrie che rispettano. L'inversione di parità P porta ad invertire il verso in cui due particelle si intrecciano, per cui manda un loop corrispondente alla treccia b in un loop corrispondente alla treccia b^{-1} e inverte l'angolo statistico θ . Allo stesso modo, anche l'inversione temporale T cambia il verso di intrecciamento delle particelle (siccome cambia la freccia del tempo il moto è percorso in verso opposto) e si inverte di nuovo l'angolo statistico. Queste inversioni non sono osservabili soltanto se le particelle sono bosoni o fermioni, perché $\theta = 0 \xrightarrow{P,T} \theta = 0$ e $\theta = \pi \xrightarrow{P,T} \theta = -\pi = \pi \pmod{2\pi}$. Dunque, in generale gli anioni rompono le simmetrie P e T , mentre solo bosoni e fermioni le rispettano; la simmetria combinata PT è chiaramente rispettata anche dagli anioni, siccome le due inversioni si annullano fra loro.

La trattazione di quanto segue esula dagli scopi di questa discussione, tuttavia, vale la pena soffermarsi un momento sull'esistenza degli anioni. Il nostro mondo ha tre dimensioni spaziali e per questo osserviamo ordinariamente solo bosoni e fermioni. Non possiamo osservare una particella ordinaria che si comporti come anione perché non si tratta di "schiacciare" il più possibile il sistema fino a renderne trascurabile lo spessore, lo si dovrebbe realmente ridurre a bidimensionale, non in approssimazione. Per questo, inizialmente, gli anioni erano visti più che altro come un interessante risultato teorico, ma lontano dalla realtà. Ciononostante, degli anioni sono stati osservati nel contesto dell'*effetto Hall quantistico* [20]. È possibile realizzare gas di elettroni che si comportano come bidimensionali: in essi gli elettroni sono liberi di muoversi in due dimensioni, ma hanno livelli energetici associati al moto nella terza dimensione che sono quantizzati discretamente; in determinati contesti, le eccitazioni discrete risultano inaccessibili, per cui

il moto è, di fatto, confinato a due dimensioni. In sistemi di questo tipo, se soggetti a basse temperature e forti campi magnetici, è possibile osservare un analogo quantistico dell'effetto Hall, in cui la resistenza di Hall R_H diventa quantizzata e assume valori universali dati da

$$R_H = \frac{1}{\nu} \frac{h}{e^2},$$

dove il rapporto tra costanti fondamentali h/e^2 diventa un quanto di resistenza. Il numero quantico ν può assumere numeri interi (effetto Hall quantistico intero) o valori razionali, tipicamente con denominatore piccolo e dispari (effetto Hall quantistico frazionario). Nell'effetto Hall quantistico frazionario sono stati effettivamente osservati anioni, ma come *quasiparticelle*, cioè come stati emergenti localizzati legati ad eccitazioni del sistema interagente.

3.3 Le particelle identiche sulla circonferenza

Nella sezione §2.4 si era già studiato il caso di una particella singola su S^1 , per mostrare un esempio di sistema che presentasse le peculiarità descritte sulla quantizzazione in spazi non semplicemente connessi. In questo capitolo si è invece descritto come l'aggiunta di particelle identiche sul dominio di confinamento costringa a modificare lo spazio delle configurazioni per tenere conto della loro indistinguibilità. Vediamo ora il risultato dell'unione della multipla connessione data da un dominio in partenza "bucato" alla multipla connessione portata dalle particelle identiche, nel caso particolare in cui si considerano più particelle sulla circonferenza S^1 . Assumeremo ancora di lavorare con particelle senza spin descritte da funzioni d'onda ψ scalari.

Secondo quanto detto nella (3.1.1), lo spazio delle configurazioni è

$$Q_N = \underbrace{(S^1 \times \cdots \times S^1)}_N \setminus \Delta_N / S_N,$$

dove $\Delta_N = \{(x_1, \dots, x_N) \in (S^1)^N \mid \exists i \neq j : x_i = x_j\}$ è la diagonale. La rimozione di Δ_N impedisce che possano avvenire loop in cui due particelle per scambiarsi si attraversano, dunque gli unici cammini continui chiusi rimanenti che siano non banali sono quelli in cui le N particelle si scambiano di posto ruotando in un ciclo. È evidente che possano essere tutti generati da un ciclo elementare in cui ogni particella si scambia con quella alla sua sinistra (quindi in senso antiorario), che assume il ruolo di generatore. Permutando ciclicamente di un posto in verso orario le posizioni delle particelle ottengo il loop inverso e qualsiasi altro cammino chiuso si può scrivere come composizione di questi. È dunque evidente la struttura delle classi di loop omotopi:

esse formano un gruppo isomorfo al gruppo libero ad un generatore, cioè il gruppo ciclico \mathbb{Z} . Più avanti ci concentreremo sul caso di $N = 2$ particelle sulla circonferenza e per questo caso particolare ritroveremo in modo più formale il risultato. Notiamo che questo non è cambiato rispetto a quando si era descritta la particella singola sulla circonferenza: vale $\pi_1(Q_N) = \mathbb{Z} \quad \forall N$, se lo spazio sottostante è S^1 .

Come nei casi precedenti, la descrizione va costruita sul rivestimento universale dello spazio delle configurazioni. Il rivestimento di Q_N sarà identificabile con una certa striscia N -dimensionale di \mathbb{R}^N , di cui più punti si proietteranno nello stesso punto di Q_N , dando origine a funzioni d'onda proiettate poldrome. Proiettando le restrizioni della funzione d'onda ai domini fondamentali si ottengono i rami della funzione d'onda su Q_N , si passa da un ramo all'altro mediante l'azione di un loop. Si è detto più volte che un loop corrisponde a uno scambio di particelle (e nel caso della circonferenza le particelle singole non possono compiere loop non banali a causa della rimozione di Δ_N). Diventa evidente, allora, che si possano identificare i domini fondamentali con regioni in cui le particelle hanno un ordinamento fissato e che il passaggio tra i domini fondamentali sia realizzato le permutazioni cicliche delle particelle.

Come preannunciato, limitiamoci per semplicità al caso con due particelle indistinguibili vincolate alla circonferenza. Per prima cosa, bisogna identificare correttamente lo spazio delle configurazioni

$$Q_2 = [(S^1 \times S^1) \setminus \Delta_2] / S_2.$$

Lo spazio $S^1 \times S^1$ è chiaramente il toro bidimensionale T^2 : ogni punto del toro corrisponde a una coppia di punti su una circonferenza. Per per capire che spazio si ottenga identificando $(x, y) \sim (y, x)$, con $x, y \in S^1$, conviene vedere la circonferenza come un segmento reale con estremi identificati, ad esempio $[0, 1]$, $0 \sim 1$. Nelle sezioni precedenti si era usato $[0, 2\pi]$, ma chiaramente per la topologia non cambia nulla. Così, il toro diventa omeomorfo al quadrato con i lati opposti identificati rappresentato nella Figura 3.3(a). Chiamando x la coordinata orizzontale e y la verticale, si ha che $(0, y) \sim (1, y) \forall y$, siccome $x = 0 \sim x = 1$. Questo significa che i due lati verticali sono identificati; con analogo ragionamento si identificano i lati orizzontali. I lati identificati sono contrassegnati in figura dallo stesso tipo di freccia, la cui punta indica il verso in cui i due lati sono da intendere incollati. I punti eliminati dal quadrato sono la diagonale, tratteggiata, e i vertici (perché $(0, 1) \sim (0, 0) \in \Delta_2$). A questo punto bisogna quozientare rispetto al gruppo delle permutazioni, cioè identificare i punti simmetrici rispetto alla diagonale del quadrato $(x, y) \sim (y, x)$.

Questa identificazione si realizza "piegando" il quadrato sulla sua diagonale e il risultato è rappresentato nella Figura 3.3(b). La piegatura identifica i lati verticali con gli adiacenti lati orizzontali, per cui ottengo tutti i lati equivalenti, dato che i lati opposti erano già identificati. I due lati del quadrato rimanenti sono identificati e vanno incollati, rispettando il verso delle frecce. Farlo non è immediato, perché incollarli direttamente, come in un cono, non rispetterebbe i versi dell'identificazione. Conviene operare un taglio preliminare lungo l'altra diagonale del quadrato, rappresentato nella Figura 3.3(c), che sarà ricucito successivamente. Incollando i due lati adiacenti del triangolo, si ottiene la striscia rappresentata nella Figura 3.3(d).

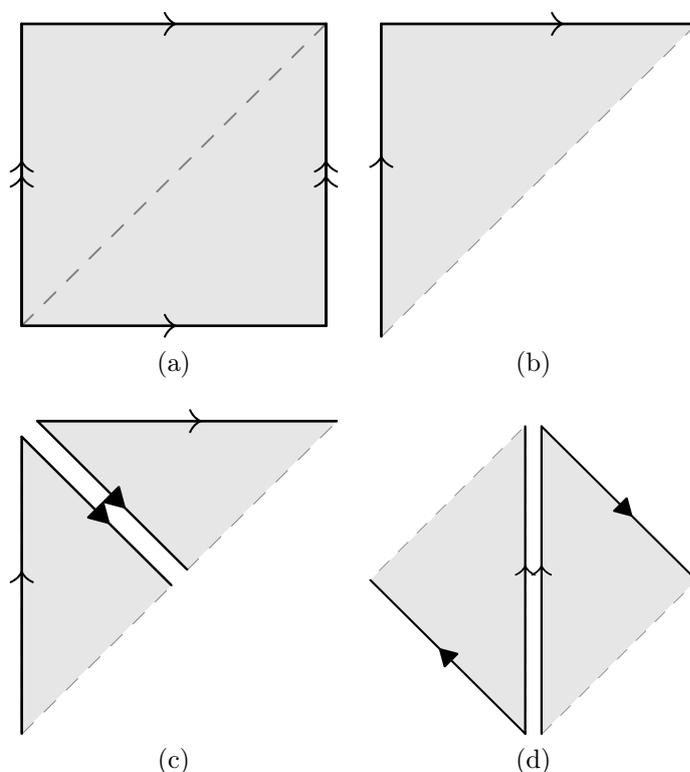


Figura 3.3: I diversi passi del "taglia e cuci" che mostra l'omeomorfismo dello spazio delle configurazioni di due particelle identiche sulla circonferenza con il nastro di Möbius.

Questa striscia ha solo due lati opposti, perché gli altri erano la diagonale che è stata rimossa. I due lati opposti vanno ricuciti, ma non si possono attaccare semplicemente come in un cilindro: per far combaciare il verso delle frecce è necessario girare un'estremità di 180° e solo successivamente incollarla. Lo spazio delle configurazioni ottenuto è un *nastro di Möbius* M ,

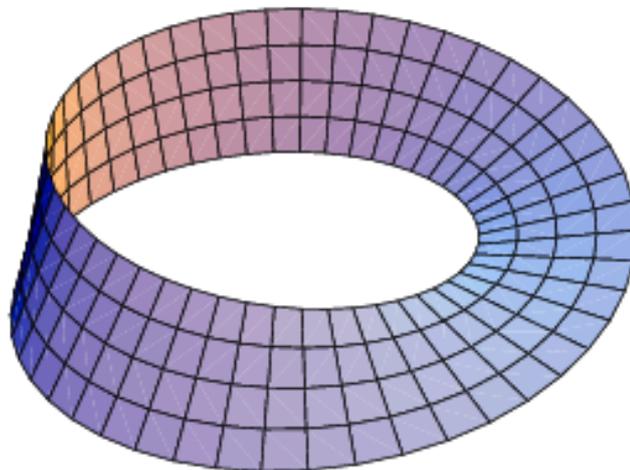


Figura 3.4: Tipica rappresentazione del nastro di Möbius come superficie immersa in \mathbb{R}^3 .

rappresentato nella Figura 3.4, e ad esso vanno considerati eliminati i bordi esterni, siccome corrispondono a Δ_2 .

Il nastro di Möbius si può vedere come una circonferenza S^1 , in cui ad ogni punto è stato incollato un segmento, omeomorfo all'intervallo reale $[0, 1]$, e questo incollamento è stato fatto "ruotando" sempre più i segmenti, in modo da farli riattaccare girati di un angolo piano quando il cerchio si chiude. Anche se non interessa molto per quanto segue, è interessante sottolineare come questa descrizione mostri chiaramente la struttura di fibrato localmente banale del nastro di Möbius, siccome è localmente omeomorfo a $S^1 \times [0, 1]$, ma non lo è globalmente. Comunque, in base a questo è evidente che sia possibile deformare con continuità un nastro di Möbius nella circonferenza che costituisce il suo spazio base, contraendo ogni intervallo $[0, 1]$ in un unico punto su S^1 . Questo tipo di trasformazione è chiamata *ritrazione* ed è un caso particolare di equivalenza di omotopia (per una descrizione completa si veda [6]). Allora, M e S^1 sono omotopicamente equivalenti e, secondo il Teorema 1.7, hanno lo stesso gruppo fondamentale, per cui

$$\pi_1(M) = \pi_1(S^1) = \mathbb{Z}.$$

Abbiamo ottenuto che il gruppo fondamentale dello spazio delle configurazioni per due particelle vincolate su una circonferenza è

$$\pi_1(Q_2) = \mathbb{Z}. \quad (3.3.1)$$

Il rivestimento universale del nastro di Möbius è identificabile con la striscia infinita $\mathbb{R} \times [0, 1]$, in cui i punti $\tilde{x} = (x, y)$, $\tilde{x}' = (x', y')$ appartenenti alla stessa fibra sono legati dalla relazione di equivalenza

$$\tilde{x}' \sim \tilde{x} \text{ se } (x', y') = f(x, y) = (x + 1, 1 - y).$$

I domini fondamentali diventano i diversi quadrati come $[0, 1] \times [0, 1]$, in cui la relazione di equivalenza introdotta determina le stesse identificazioni tipiche del nastro di Möbius, cioè i lati verticali sono identificati ma invertendone l'orientazione.

La funzione d'onda sarà ben definita su tutto il rivestimento, per cui polidroma sullo spazio delle configurazioni iniziali. Come già detto più volte, restringendosi sui domini fondamentali, si lavorerà con funzioni d'onda univoche (i rami della funzione polidroma) ma discontinue nei bordi che vengono identificate. Le condizioni al contorno che vanno imposte per tenere conto di questa discontinuità si ottengono imponendo che, sotto l'azione di un loop sullo spazio delle configurazione, la funzioni d'onda cambi per una rappresentazione unitaria di esso. Il gruppo fondamentale dello spazio delle configurazioni è ancora \mathbb{Z} , per cui le sue rappresentazioni su $U(1)$ sono già state trovate e sono

$$a : [\gamma] \in \pi_1(Q_2) = \mathbb{Z} \longmapsto a([\gamma]) = e^{in\theta} \in U(1), \quad \theta \in [0, 2\pi[,$$

per cui si dovrà imporre la condizione

$$\psi(\tilde{q} \star [\gamma]_n) = e^{in\theta} \psi(\tilde{q}), \quad (3.3.2)$$

dove $[\gamma]_n$, $n \in \mathbb{Z}$ sono gli elementi del gruppo fondamentale isomorfo a \mathbb{Z} . Come nel caso di una particella singola su S^1 , si hanno infiniti settori di superselezione etichettati dall'angolo θ che varia nell'intervallo $[0, 2\pi[$.

Proiettando un ramo della funzione d'onda, esso è definito sullo spazio delle configurazioni $Q_2 = M$. Per lavorarci, è comodo individuare un generico punto sul nastro di Möbius con un punto sul triangolo iniziale, ottenuto dalle opportune identificazioni già descritte sul quadrato $[0, 1] \times [0, 1]$. Reinserendo una nozione di lunghezza, che era stata "dimenticata" perché ininfluenza sulla topologia, possiamo affermare che partire dal quadrato $[0, 1] \times [0, 1]$ sia equivalente a considerare una circonferenza di lunghezza unitaria. Nel problema della particella singola si era assunto unitario il raggio, e dunque la lunghezza era 2π . Per lavorare più in generale, chiamiamo L la lunghezza della circonferenza, per cui un dominio fondamentale ($\sim M$) sarà il triangolo ottenuto dal quadrato $[0, L] \times [0, L]$ ripiegato, come già descritto, e identificheremo i punti di esso con le coppie (x_1, x_2) . I bordi di questo triangolo

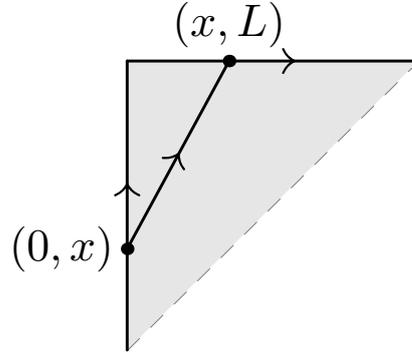


Figura 3.5: Il cammino che manda $(0, x)$ in (x, L) è un loop in Q_2 che genera gli scambi delle due particelle sulla circonferenza

sono i due lati adiacenti che, per la discussione precedente, sono identificati in modo che $(0, x) \sim (x, L)$: su questi punti che va vincolata la polidromia della funzione d'onda con la condizione al contorno precedente. Il cammino (rappresentato nella Figura 3.5) che manda $(0, x)$ in (x, L) è un loop che lega domini fondamentali adiacenti, è un generatore del gruppo fondamentale ed è associato al numero intero $n = 1$. Così, il vincolo (3.3.2) si traduce nella condizione al contorno

$$\psi(x, L) = e^{i\theta} \psi(0, x). \quad (3.3.3)$$

Lo spazio di Hilbert \mathcal{H} per questo problema consiste sempre nelle funzioni complesse a quadrato sommabile definite sullo spazio delle configurazioni, che è identificato in questo triangolo \mathcal{T} :

$$\mathcal{H} = \{ \psi : \mathcal{T} \rightarrow \mathbb{C} \mid \langle \psi, \psi \rangle < +\infty \},$$

dove il prodotto interno è appunto l'integrale definito sul triangolo

$$\langle \psi, \phi \rangle = \int_0^L dx_2 \int_0^{x_2} dx_1 (\psi^* \phi)(x_1, x_2).$$

Vogliamo trovare le autofunzioni dell'energia per le due particelle libere, per cui l'operatore hamiltoniano è

$$\hat{H} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right)$$

Esattamente come nel caso di una particella singola, la condizione (3.3.3) non è sufficiente a determinare il dominio di autoaggiunzione della hamiltoniana, ma serve anche una condizione sulle derivate analoga alla (2.4.3):

$$\frac{\partial \psi}{\partial x_2}(x, L) = e^{i\theta} \frac{\partial \psi}{\partial x_1}(0, x). \quad (3.3.4)$$

Le condizioni (3.3.3) e (3.3.4) non sono ancora sufficienti per rendere \hat{H} autoaggiunto: lo spazio delle configurazioni è identificato con il triangolo \mathcal{T} della Figura 3.5 e la diagonale è un'altra componente della frontiera di \mathcal{T} , su cui vanno imposte opportune condizioni al contorno. La diagonale è costituita nei punti coincidenti, per cui $x_1 = x_2$, e su essi si impone la condizione

$$\left. \frac{(\partial_2 - \partial_1)\psi(x_1, x_2)}{\psi(x_1, x_2)} \right|_{x_1 \rightarrow x_2} = \lambda, \quad (3.3.5)$$

dove $\lambda \in [0, +\infty]$ (infinito incluso!) e si è abbreviato $\partial_i \equiv \partial/\partial x_i$. Assieme, le relazioni (3.3.3), (3.3.4), (3.3.5) determinano il dominio dell'operatore hamiltoniano \hat{H} . Si potrebbe passare nelle coordinate del centro di massa, definendo come coordinata relativa $r = x_2 - x_1$ e la componente dell'energia associata al moto relativo sarebbe descritta dall'operatore ∂_r^2 . Questo operatore chiaramente può agire su funzioni in r definite nel dominio $0 < r < +\infty$ e, integrando per parti, si può vedere che la condizione (3.3.5) (che diventa $\partial_r \psi(r \rightarrow 0) = \lambda \psi(r \rightarrow 0)$) è quella che ne determina l'autoaggiunzione, mentre le (3.3.3) e (3.3.4) rendono autoaggiunta la parte del centro di massa.

È interessante notare che questa condizione generalizza degli analoghi unidimensionali per le statistiche bosoniche e fermioniche, interpolandole: se $\lambda = +\infty$ si riduce a $\psi(x_1, x_2)|_{x_1 \rightarrow x_2} = 0$, come per i fermioni (caratterizzati da una ψ antisimmetrica rispetto allo scambio), e se $\lambda = 0$ si riduce a $(\partial_2 - \partial_1)\psi(x_1, x_2)|_{x_1 \rightarrow x_2} = 0$, come per i bosoni (caratterizzati da una ψ simmetrica).

Vogliamo risolvere il problema gli autovalori per le due particelle libere $\hat{H}\psi = E\psi$, per cui, nella rappresentazione delle posizioni, l'equazione di Schrödinger indipendente dal tempo è

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \psi(x_1, x_2) = E \psi(x_1, x_2). \quad (3.3.6)$$

Pur non essendo a rigore completamente dimostrato, si pensa che tutte le soluzioni di questa equazione siano date dal *Bethe ansatz* [2]

$$\psi_{k_1 k_2}(x_1, x_2) = A e^{i(k_1 x_1 + k_2 x_2)} + B e^{i(k_2 x_1 + k_1 x_2)}.$$

L'idea è che, essendo le due particelle indistinguibili, non sarà possibile assegnare alla prima in posizione x_1 il numero d'onda k_1 e alla seconda in x_2 il numero d'onda k_2 , per cui si considera in sovrapposizione anche il caso con i numeri d'onda scambiati. Si sono usate in sovrapposizione onde piane perché sono particelle libere. Raccogliendo A , possiamo scrivere l'ansatz come

$$\psi_{k_1 k_2}(x_1, x_2) = e^{i(k_1 x_1 + k_2 x_2)} + C e^{i(k_2 x_1 + k_1 x_2)}, \quad (3.3.7)$$

lasciando a parte il fattore di normalizzazione. Sostituendo questo ansatz nell'equazione (3.3.6) è immediato trovare che effettivamente sono autofunzioni dell'energia e che l'autovalore corrispondente a $\psi_{k_1 k_2}$ è

$$E_{k_1 k_2} = \frac{\hbar^2}{2m} (k_1^2 + k_2^2). \quad (3.3.8)$$

Imponendo la condizione (3.3.5) si trova

$$C = \frac{k_1 - k_2 - i\lambda}{k_1 + k_2 + i\lambda},$$

mentre da (3.3.3) e (3.3.5) risultano le relazioni

$$\begin{aligned} e^{ik_1 L} &= C^{-1} e^{i\theta}, \\ e^{ik_2 L} &= C e^{i\theta}. \end{aligned}$$

Questi vincoli determinano le condizioni di quantizzazione e dunque lo spettro del sistema, se $k_1 \neq k_2$. Ad esempio, per due bosoni $\lambda = 0 \Rightarrow C = 1$ (se $k_1 \neq k_2$) e dalle relazioni precedenti si ha che

$$k_1 L - \theta = 2\pi \cdot n \quad \text{e} \quad k_2 L - \theta = 2\pi \cdot n',$$

con $n, n' \in \mathbb{Z}$, per cui la quantizzazione degli impulsi è data da

$$\begin{aligned} k_1 &= \frac{2\pi}{L} n + \frac{\theta}{L} \quad n \in \mathbb{Z}, \\ k_2 &= \frac{2\pi}{L} n' + \frac{\theta}{L} \quad n' \in \mathbb{Z}. \end{aligned}$$

Possiamo calcolare allora lo spettro energetico per $\lambda = 0$, che è

$$E_{n, n'} = \frac{\hbar^2}{2m} \left[\left(\frac{2\pi}{L} n + \frac{\theta}{L} \right)^2 + \left(\frac{2\pi}{L} n' + \frac{\theta}{L} \right)^2 \right];$$

come nel caso di una particella singola, l'angolo θ che identifica il settore di superselezione è misurabile, siccome influenza lo spettro. Notiamo anche che questo spettro è coerente con quella in cui si ha una particella dato dalla (2.4.5), in cui si era assunto che la circonferenza fosse unitaria e dunque $L = 2\pi$. Nel caso dei fermioni $\lambda \rightarrow +\infty$ e $C \rightarrow -1$, per cui, analogamente si ottengono

$$\begin{aligned} k_1 &= \frac{2\pi}{L} n + \frac{\pi + \theta}{L} \quad n \in \mathbb{Z}, \\ k_2 &= \frac{2\pi}{L} n' + \frac{\pi + \theta}{L} \quad n' \in \mathbb{Z} \end{aligned}$$

e sostituendole nella (3.3.8) si trova lo spettro

$$E_{n,n'} = \frac{\hbar^2}{2m} \left[\left(\frac{2\pi}{L}n + \frac{\pi + \theta}{L} \right)^2 + \left(\frac{2\pi}{L}n' + \frac{\pi + \theta}{L} \right)^2 \right].$$

Nei casi intermedi con $0 < \lambda < +\infty$ si devono risolvere equazioni trascendenti, ma la situazione è analoga. Vediamo comparire due angoli che descrivono la quantizzazione del sistema: uno è l'angolo θ del settore di superselezione e l'altro è un angolo ϕ , legato a λ , che nel primo caso vale $\phi = 0$ e nel secondo $\phi = \pi$. Studiando anche il caso $k_1 = k_2$, si trova [2] che è ammesso solo per $\lambda = 0$ e che conduce alla condizione di quantizzazione già descritta.

Conclusioni

Solitamente, la meccanica quantistica introduttiva viene presentata su spazi semplici, come \mathbb{R}^n . Il percorso compiuto ci ha portato a vedere che quando lo spazio delle configurazioni ha topologia non banale (come invece è la topologia di \mathbb{R}^n) emergono complicazioni che rendono evidente l'importanza della fase di una funzione d'onda. I problemi sono dovuti alla non semplice connessione dello spazio delle configurazioni: si ottengono in generale funzioni d'onda polidrome, le quali variano con continuità lungo percorsi continui, ma che lo fanno in modo diverso percorso per percorso, dando risultati differenti per ogni classe di omotopia. In questi casi, la meccanica quantistica va costruita sul rivestimento universale dello spazio delle configurazioni, che è uno spazio semplicemente connesso, e poi va riproiettata sullo spazio fisico di base in modo coerente con la struttura del rivestimento. Per proiettare ragionevolmente le funzioni d'onda, si sono imposte delle condizioni che vincolavano la polidromia a quella della fase e questo è l'origine degli effetti interessanti che si hanno quando lo spazio delle configurazioni ha topologia non banale. Nonostante solitamente la fase di una funzione d'onda non dia contributo ai valori attesi delle osservabili, è importante determinarla perché, in questi casi, diverse fasi portano a differenze misurabili di comportamento e a diverse quantizzazioni dello stesso sistema classico. In particolare, abbiamo visto con l'effetto Aharonov - Bohm che l'evoluzione non banale della fase della funzione d'onda che descrive il sistema porti a modifiche ad effetti di interferenza misurabili; queste modifiche sono stati effettivamente osservati sperimentalmente, ad esempio da Osakabe ed altri nel 1986 [14]. Studiando ramo per ramo le funzioni d'onda a più valori, i vincoli sulla proiettabilità diventano condizioni al contorno di discontinuità e attraverso gli esempi di una e di due particelle su una circonferenza, abbiamo visto come queste condizioni al contorno determinassero lo spettro energetico del sistema, cioè i valori ammessi per l'energia. La fase della funzione d'onda non potrà contribuire alla forma funzionale della densità di probabilità $|\psi|^2$, ma resta di fondamentale importanza perché esistono casi in cui determina gli autovalori ammessi delle osservabili e le corrispondenti autofunzioni: è evidente che variazioni di

essa diano effetti misurabili. Più in generale, si è mostrato che queste condizioni al contorno sulla fase possono non essere determinate univocamente dal sistema classico che si sta cercando di quantizzare: esistono più sistemi quantistici che in un limite semiclassico si riducono allo stesso sistema. Essendo in questi casi la fase delle funzioni d'onda percepibile, si avranno più quantizzazioni inequivalenti dello stesso sistema, ciascuna caratterizzata da diversi autovalori delle osservabili, ma non solo: possono differire anche per i loro comportamenti statistici. Infatti, studiando i sistemi di particelle identiche in \mathbb{R}^n , abbiamo visto che i familiari concetti di bosone e fermione non siano altro che un esempio di quantizzazioni inequivalenti dello stesso sistema classico di particelle identiche: la separazione delle particelle in bosoni e fermioni emerge direttamente da ambiguità non risolubili fino in fondo nella procedura di quantizzazione. Questa procedura di quantizzazione è particolarmente interessante perché ci ha condotto anche all'esistenza degli anioni: la suddivisione delle particelle nelle sole classi di bosoni e fermioni è valida solo in uno spazio tridimensionale (o più), mentre in due dimensioni spaziali sono possibili infinite classi di statistiche che interpolano quelle bosoniche e fermioniche e vengono chiamate, appunto, statistiche anioniche.

Bibliografia

- [1] Y. Aharonov, D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory*, Phys. Rev., **115**(3), 485 (1959).
- [2] C. Aneziris, A. P. Balachandran e D. Sen, *Statistics in One Dimension*, Int. J. Mod. Phys. **A6**, 4721 (1991); Erratum **A7**, 1851 (1992).
- [3] A. P. Balachandran, G. Marmo, B. S. Skagerstam e A. Stern, *Classical Topology and Quantum States* (World Scientific, 1991).
- [4] A. P. Balachandran e E. Ercolessi, *Statistics on Networks*, Int. J. Mod. Phys. **A19**, 4633 (1992).
- [5] D. J. Griffiths, *Introduction to Quantum Mechanics*, Prentice Hall (2005).
- [6] A. Hatcher, *Algebraic Topology*, Cambridge University Press (2002).
- [7] P. A. Horvathy, G. Morandi e E. C. G. Sudarshan, *Inequivalent Quantizations in Multiply Connected Spaces*, Nuovo Cim. **B11**, 201 (1989).
- [8] K. Huang, *Statistical Mechanics*, John Wiley & Sons Inc (1987).
- [9] J. M. Leinaas e J. Myrheim, *On the theory of identical particles*, Nuovo Cim. **B37**, 1 (1977).
- [10] C. R. F. Maunder, *Algebraic Topology*, Dover (1992).
- [11] G. Morandi, *The Role of Topology in Classical and Quantum Physics*, Springer (1992).
- [12] M. Nakahara, *Geometry, Topology and physics*, Institute of Physics Publishing (2003).
- [13] R. D. Richtmeyer, *Principles of Advanced Mathematical Physics*, Springer (Vol. I 1978, Vol. II 1981)

-
- [14] N. Osakabe, T. Matsuda, T. Kawasaki, J. Endo, A. Tonomura, S. Yano, e H. Yamada, *Experimental confirmation of Aharonov-Bohm effect using a toroidal magnetic field confined by a superconductor*, Phys. Rev. **A34**(2), 815 (1986)
- [15] W. Pauli, *Über ein Kriterium für Ein- oder Zweiwertigkeit der Eigenfunktionen in der Wellenmechanik (On a criterion for one- or two-valuedness of eigenfunctions in wave mechanics)*, Helv. Phys. Acta, **12**(2), 147 (1939).
- [16] J. J. Sakurai, J. Napolitano, *Modern Quantum Mechanics*; Cambridge University Press (2017);
- [17] I. M. Singer e J. A. Thorpe, *Lecture Notes on Elementary Topology and Geometry*, Scott, Foresman & Co. (1967)
- [18] E. H. Spanier, *Algebraic Topology*, Springer (1994).
- [19] N. Steendrod, *The Topology of Fibre Bundles*, Princeton University Press (1999).
- [20] A. Stern, *Anyons and the quantum Hall effect - A pedagogical review*, Ann. Phys., **323**, 204 (2008).
- [21] F. Zaccaria, E. C. G. Sudarshan, J. S. Nilsson, N. Mukunda, G. Marmo, e A. P. Balachandran, *Universal unfolding of Hamiltonian systems: From symplectic structure to fiber bundles*, Phys. Rev. **D27**, 2327 (1983)