CAPITOLO 1

INTRODUZIONE MATEMATICA

Iniziamo ora dalle definizioni e dalle nozioni matematiche principali, che saranno *la base* di questo corso. Ovviamente alcune di esse sono state semplificate, viste le finalità di questo corso, ma resta comunque necessario fissare i concetti e la notazione relativa a questi oggetti, così che tutti possiate seguire tranquillamente le lezioni e concentrarvi poi sulla comprensione dei nuovi argomenti.

KeyWords: campo complesso, spazio vettoriale, prodotto scalare, operatori e matrici

1.1 Campi

Prima di procedere con le definizioni cardine di questo corso, ci servono delle conoscenze preliminari di matematica, in particolar modo ci servirà sapere cos'è un campo e le proprietà principali dei numeri complessi.

Definizione 1.1. Un insieme non vuoto K dotato di due operazioni binarie $+, *$ è un **campo** se $(K, +)$ è un gruppo abeliano con elemento neutro 0, ossia se per ogni $a, b, c \in K$

- $(a + b) + c = a + (b + c)$
- $a + b = b + a$
- $0 + a = \ a + 0 = a$
- $\forall a, \exists \ -a \mid a + (-a) = -a + a = 0$

se $(K^* = K \setminus \{0\}, *)$ è anch'esso un gruppo abeliano con elemento neutro 1, ossia se per ogni $a, b, c \in K^*$

- $(a * b) * c = a * (b * c)$
- $a * b = b * a$
- $1 * a = \ a * 1 = a$
- $\forall a, \exists\, a^{-1} \mid a * a^{-1} = a^{-1} * a = 1$

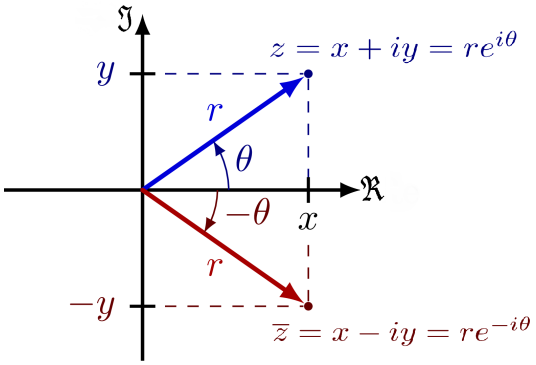
e se la moltiplicazione è distributiva rispetto all'addizione, ossia se per ogni $a, b, c \in K$

$$a * (b + c) = (a * b) + (a * c)$$

I campi che useremo sono, in particolare, il campo dei reali $\mathbb{R}$ e il campo dei complessi $\mathbb{C}$, di cui procediamo ad analizzarne gli elementi.

Definizione 1.2. Un **numero complesso** è un numero della forma $a + ib$, con $a, b \in \mathbb{R}$ e dove i, detta **unità immaginaria**, è soluzione dell'equazione $x^2 = -1$. La scrittura $a + ib$ è detta **forma algebrica** di un numero complesso.

I numeri complessi sono usati in tutti i campi della matematica, in molti campi della fisica (notoriamente in meccanica quantistica, come vedremo), nonché in ingegneria (come ad esempio nelle telecomunicazioni) per la loro utilità nel rappresentare onde elettromagnetiche e correnti elettriche. In matematica, sono generalmente rappresentati come punti di un piano, i cui assi sono la **parte reale** $\Re$ e la **parte immaginaria** $\Im$.



Una scrittura equivalente dei numeri complessi è la cosiddetta **forma esponenziale**, definita come

$$z = re^{i\theta} = \cos(\theta) + i\sin(\theta)$$

dove $r = |z| = \sqrt{a^2 + b^2}$ è detto **modulo** (o **magnitudine**) di z e $\theta \in [0, 2\pi]$ è l'angolo che z forma con il semiasse positivo della parte reale.
Riportiamo ora le operazioni principali che si possono applicare ai numeri complessi.

Definizione 1.3. Dato un numero complesso $z = a + ib$, definiamo il suo **modulo** come

$$|z| := \sqrt{a^2 + b^2}$$

che ha le seguenti proprietà per ogni $z, w \in \mathbb{C}$:

- $|z + w| \leq |z| + |w|$, detta **disuguaglianza triangolare**;
- $|zw| = |z||w|$
- se $w \neq 0$, $\left| \frac{z}{w} \right| = \frac{|z|}{|w|}$

Possiamo dunque definire la distanza tra due punti nel piano complesso come

$$d(z, w) = |z - w|$$

Definizione 1.4. Dato un numero complesso $z = a + ib$, definiamo il suo **coniugato** $\overline{z}$ come

$$\overline{z} := a - ib$$

e osserviamo che, nel piano complesso, equivale a ribaltare il nostro punto rispetto all'asse reale. Il coniugato gode delle seguenti proprietà per ogni $z, w \in \mathbb{C}$

- $\overline{z+w} = \overline{z} + \overline{w}$
- $\overline{zw} = \overline{z}\,\overline{w}$ e $\overline{z/w} = \overline{z}/\overline{w}$
- $\overline{\overline{z}} = z$
- $\overline{z} = z$ se e solo se $z \in \mathbb{R}$
- $|z| = |\overline{z}|$
- $|z|^2 = z\overline{z}$

In notazione esponenziale, il complesso coniugato di z è semplicemente $\overline{z} = re^{-i\theta}$.

Oltre a quello dei complessi, un altro campo importante nella matematica e nell'informatica (e che useremo molto durante questo corso) è $\mathbb{F}_2$, definito su $\{0, 1\}$ con le operazioni di somma e moltiplicazione binarie

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

1.2 Spazi Vettoriali e Spazi di Hilbert

Definizione 1.5 (Spazi e Sottospazi)**.**

1. Uno **spazio vettoriale** V su un campo $\mathbb{F}$ è un insieme di oggetti (chiamati **vettori**) che soddisfano la seguente relazione

   $$\text{se } \mathbf{v}, \mathbf{w} \in V \text{ allora } c\mathbf{v} + d\mathbf{w} \text{ per ogni } c, d \in \mathbb{F}$$

   ossia V è chiuso rispetto all'addizione e alla moltiplicazione scalare. Come notazione, useremo il grassetto per indicare i **vettori**, i.e. gli elementi di V, e le lettere normali per indicare gli **scalari**, i.e. gli elementi di $\mathbb{F}$.

2. Un **sottospazio** W è un sottoinsieme $W \subseteq V$ che "eredita" da V la chiusura rispetto all'addizione e alla moltiplicazione scalare. Si denota con $W \leq V$.

Riportiamo ora alcuni esempi di spazio vettoriale e di sottospazio:

- $V = \mathbb{C}^d$ è lo spazio $d-$dimensionale dei numeri complessi, ossia lo spazio formato da vettori colonna con d entrate complesse. Un possibile sottospazio $W \leq V$ è ad esempio l'insieme dei vettori le cui due prime entrate sono nulle;
- $V = \{0, 1\}^d$ è lo spazio dei vettori binari a d entrate, dotato di somma binaria *entry-wise*, definito sul campo $\mathbb{F}_2$. Un possibile sottospazio $W \leq V$ è l'insieme di vettori le cui prime due entrate sono uguali.

Definizione 1.6 (Indipendenza, Span e Basi)**.**

1. I vettori $v_1, \ldots, v_m \in V$ sono **linearmente indipendenti** se è vero che

   $$\sum_{i=1}^{m} a_i v_i = 0 \Leftrightarrow a_1 = a_2 = \cdots = a_m = 0$$

2. Lo **span** di un insieme di vettori $S = \{v_1, \ldots, v_m\} \subseteq V$ è l'insieme $\text{Span}(S)$ formato dai vettori ottenibili come **combinazioni lineari** dei vettori in S, ossia

$$u \in \text{Span}(S) \text{ sse } u = \sum_{i=1}^{m} a_i v_i, \text{ con } a_i \in \mathbb{F}$$

3. Una **base** di V è un insieme $\mathcal{B}$ di vettori linearmente indipendenti tale che $\text{Span}(S) = V$. É possibile dimostrare che tutte le basi hanno la stessa cardinalità, massimale rispetto all'indipendenza, e tale cardinalità è dunque detta **dimensione** dello spazio vettoriale V. Se fissiamo una base ordinata $\mathcal{B} = \{v_1, \ldots, v_m\}$, allora ogni $w \in V$ può essere scritto come una combinazione lineare dei vettori della base

$$w = \sum_{i=1}^{m} w_i v_i$$

e tale combinazione è unica. Inoltre, otteniamo che che il vettore w può essere scritto come $(w_1, \ldots, w_m)^T$ rispetto alla base $\mathcal{B}$.

Definizione 1.7 (Prodotto scalare)**.**
Siano $\mathbf{a}, \mathbf{b}$ due vettori in $\mathbb{R}^m$. Allora il **prodotto scalare** è definito come

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b} = \sum_{k=1}^{m} a_k b_k$$

Se invece i due vettori sono in $\mathbb{C}^m$ il prodotto interno è definito come

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^* \mathbf{b} = \sum_{k=1}^{m} \overline{a}_k b_k$$

dove con $\mathbf{a}^*$ indichiamo il vettore complesso trasposto con entrate coniugate.

In particolare, il prodotto scalare complesso gode delle seguenti proprietà

$$\langle \psi | \varphi \rangle = \overline{\langle \varphi | \psi \rangle} \tag{1.1}$$

$$\langle \psi | \psi \rangle \geq 0 \tag{1.2}$$

$$\langle \psi | \psi \rangle = 0 \Leftrightarrow \psi = 0 \tag{1.3}$$

$$\langle \psi | a\varphi_1 + b\varphi_2 \rangle = a \langle \psi | \varphi_1 \rangle + b \langle \psi | \varphi_2 \rangle \tag{1.4}$$

$$\langle a\psi | \varphi \rangle = \overline{a} \langle \psi | \varphi \rangle \quad \text{e} \quad \| a\varphi \| = |a| \, \| \varphi \| \tag{1.5}$$

$$\langle \psi | \varphi \rangle = 0 \text{ per ogni } \varphi \in \mathbb{C}^m \Leftrightarrow \psi = 0 \tag{1.6}$$

per ogni $\varphi, \psi, \varphi_1, \varphi_2 \in \mathbb{C}^m$ e per ogni $a, b \in \mathbb{C}$. Osserviamo che dalla Proprietà (1.4), possiamo dedurre che il prodotto scalare sia **lineare** nel primo argomento e **anti-lineare** nel secondo, mentre dalla Proprietà (1.1) otteniamo che $\langle \psi | \psi \rangle$ è un numero reale.

Definizione 1.8 (Spazi di Hilbert)**.**
Uno **spazio di Hilbert** $\mathbb{H}$ è uno spazio vettoriale su cui è definito un **prodotto scalare** $\langle \cdot | \cdot \rangle : \mathbb{H} \times \mathbb{H} \longrightarrow \mathbb{C}$ tale che

$$(\psi, \varphi) \mapsto \langle \psi | \varphi \rangle$$

Tale prodotto induce su $\mathbb{H}$ una norma $\| \cdot \| : \mathbb{H} \longrightarrow \mathbb{R}$

$$\psi \mapsto \sqrt{\langle \psi | \psi \rangle} = \sqrt{\sum_i |\psi_i|^2}$$

che rende tale spazio uno **spazio metrico completo**.

Per quello che concerne questo corso, ci limiteremo a considerare lo spazio di Hilbert come un semplice spazio vettoriale in campo complesso. Osserviamo dunque che uno spazio vettoriale $\mathbb{H}_{sub} \subseteq \mathbb{H}$ mantiene, per ereditarietà, sia il prodotto scalare, sia la norma, e che dunque possiamo parlare di **sottospazio** di $\mathbb{H}$.

Definizione 1.9 (Ortogonalità e Ortonormalità)**.**
Due vettori $\mathbf{v}, \mathbf{w}$ si dicono **ortogonali** se $\langle \mathbf{v}|\mathbf{w}\rangle = 0$. Un insieme di vettori $\{\mathbf{v}_i\}_i$ è **ortogonale** se tutti i vettori sono ortogonali due a due, ossia se

$$\langle \mathbf{v}_i|\mathbf{v}_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

dove δ_{ij} è la **funzione di Kronecker**.
Inoltre, se due vettori ortogonali hanno norma unitaria, essi si dicono **ortonormali**.

Vedremo più avanti come i concetti di ortogonalità e ortonormalità siano indispensabili per definire delle basi ottime per gli spazi vettoriali.

1.3 Operatori e Matrici

Dall'algebra lineare, sappiamo che è possibile rappresentare univocamente ogni operatore lineare tra due spazi vettoriali come una matrice e perciò i due termini saranno spesso intercambiati, com'è possibile anche notare dal fatto che entrambi siano denotati con delle lettere maiuscole e in grassetto.

Assumeremo familiarità con le operazioni di somma e moltiplicazione tra matrici, riporteremo solo le notazioni adottate.
Data una matrice $\mathbf{A}$, indicheremo con $(\mathbf{A})_{ij} = a_{ij}$ l'entrata nella posizione (i, j) e con $\mathbf{A}^T$ la matrice **trasposta** di $\mathbf{A}$, tale che $(\mathbf{A}^T)_{ij} = (\mathbf{A})_{ji}$. Denoteremo con $\mathbb{I}_d$ la **matrice identità** di dimensioni $d \times d$, definita come

$$(\mathbb{I})_{ij} \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Se la dimensione è ovvia dal contesto, ometteremo il pedice d. Inoltre, se $\mathbf{A}$ è una matrice quadrata per cui esiste una matrice $\mathbf{B}$ tale che $\mathbf{AB} = \mathbf{BA} = \mathbb{I}$, diremo che $\mathbf{B}$ è l'**inversa** di $\mathbf{A}$ e la denoteremo con $\mathbf{A}^{-1}$. Osserviamo che $(AB)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
Dato che il quantum computing è studiato in ambito complesso, ci servirà anche il concetto di **matrice autoggiunta**, definita come la matrice $\mathbf{A}^\dagger$ le cui entrate sono trasposte e coniugate rispetto alla matrice $\mathbf{A}$ (per tale motivo si parla anche di matrice **trasposta coniugata**.

1.3.1 Matrici unitarie

Come vedremo più avanti nel corso della giornata, gli operatori quantistici si basano sul concetto di **unitarietà**: una matrice $\mathbf{A}$ è detta **unitaria** se la sua inversa coincide con la sua trasposta coniugata, ossia se $\mathbf{A}^{-1} = \mathbf{A}^\dagger$. Condizioni equivalenti a tale definizione sono

1. $\mathbf{A}$ è unitaria;
2. $\mathbf{A}$ preserva il prodotto interno: $\langle \mathbf{Av}|\mathbf{Aw}\rangle = \langle \mathbf{v}|\mathbf{w}\rangle$ per ogni $\mathbf{v}, \mathbf{w}$;
3. $\mathbf{A}$ preserva la norma: $\| \mathbf{Av} \| = \| \mathbf{v} \|$ per ogni $\mathbf{v}$;

4. $\| \mathbf{Av} \| = 1$ se $\| \mathbf{v} \| = 1$.

Ora, (1) implica (2) perché se $\mathbf{A}$ è unitaria allora $\mathbf{A}^\dagger\mathbf{A} = \mathbb{I}$ e dunque $\langle \mathbf{Av}|\mathbf{Aw}\rangle = (\mathbf{v}^*\mathbf{A}^\dagger)\mathbf{Aw} = \langle \mathbf{v}|\mathbf{w}\rangle$. (2) implica (1) supponendo per assurdo che $\mathbf{A}$ non sia unitaria, ottenendo dunque che esista un vettore $\mathbf{w}$ per cui $\mathbf{A}^\dagger\mathbf{Aw} \neq \mathbf{w}$, da cui segue che $\langle \mathbf{v}|\mathbf{w}\rangle \neq \langle \mathbf{v}|\mathbf{A}^\dagger\mathbf{Aw}\rangle = \langle \mathbf{Av}|\mathbf{Aw}\rangle$, che è una contraddizione. Ovviamente (2) implica (3), mentre per dimostrare che (2) implica (3) ci basta sfruttare la seguente identità:

$$\| \mathbf{v} + \mathbf{w} \|^2 = \| \mathbf{v} \|^2 + \| \mathbf{w} \|^2 + \langle \mathbf{v}|\mathbf{w}\rangle + \langle \mathbf{w}|\mathbf{v}\rangle$$

L'equivalenza tra (3) e (4) è ovvia. Osserviamo che da (4) segue che gli autovalori di una matrice unitaria hanno valore assoluto pari a 1.

1.3.2 Diagonalizzazione e autovalori

Definizione 1.10 (Autovalori e Autovettori)**.**
Un numero complesso λ è un **autovalore** di una matrice quadrata $\mathbf{A}$ se esiste un vettore non nullo $\mathbf{v}$, detto **autovettore**, tale che $\mathbf{Av} = \lambda\mathbf{v}$.

Una matrice **diagonale** è una matrice tale che $(\mathbf{D}_{ij}) = 0$ per ogni $i \neq j$. Sia $\mathbf{S}$ una qualche matrice per cui risulta verficata l'uguaglianza $\mathbf{AS} = \mathbf{SD}$, con $\mathbf{D}$ matrice diagonale. Sia $\mathbf{v}_i$ la $i-$esima colonna di $\mathbf{S}$ e sia λ_i l'$i-$esima entrata della matrice diagonale $\mathbf{D}$; allora

$$\underbrace{\begin{bmatrix} \vdots & & \vdots \\ \mathbf{Av}_1 & \cdots & \mathbf{Av}_d \\ \vdots & & \vdots \end{bmatrix}}_{\mathbf{AS}} = \underbrace{\begin{bmatrix} \vdots & & \vdots \\ \lambda_1\mathbf{v}_1 & \cdots & \lambda_d\mathbf{v}_d \\ \vdots & & \vdots \end{bmatrix}}_{\mathbf{SD}}$$

da cui possiamo dedurre che $\mathbf{v}_i$ è un autovettore di $\mathbf{A}$ associato all'autovalore λ_i. Viceversa, se $\mathbf{v}_1, \ldots, \mathbf{v}_d$ sono autovalori di $\mathbf{A}$ con autovalori associati $\lambda_1, \ldots, \lambda_d$, allora si ha $\mathbf{AS} = \mathbf{SD}$, con $\mathbf{S}$ formata dagli autovettori come colonne e $\mathbf{D}$ è la matrice diagonale con gli autovalori. Dunque, diremo che una matrice quadrata $\mathbf{A}$ è **diagonalizzabile** se esiste una matrice $\mathbf{S}$ tale che $\mathbf{A} = \mathbf{SDS}^{-1}$, da cui otteniamo che $\mathbf{A}$ è diagonalizzabile se e solamente se ha d autovettori indipendenti.
Diremo inoltre che $\mathbf{A}$ è **unitariamente diagonalizzabile** se la matrice $\mathbf{S}$ è unitaria, ossia formata da autovettori ortonormali. Osserviamo che l'insieme di autovettori di una matrice forma una base ortonormale dello spazio vettoriale e che per tanto possiamo utilizzarla come base di riferimento oltre a quella canonica.
L'ultima classe di matrici importanti da definire in questo preambolo sono le matrici Hermitiane:

Definizione 1.11 (Matrici Hermitiane)**.** Una matrice $\mathbf{A}$ si dice **Hermitiana** se $\mathbf{A} = \mathbf{A}^\dagger$.

Dato che se $\mathbf{A} = \mathbf{UDU}^{-1}$ (i.e. se $\mathbf{A}$ è unitariamente diagonalizzabile) abbiamo anche che $\mathbf{A}^\dagger = \mathbf{UD}^\dagger\mathbf{U}^{-1}$, possiamo concludere che gli autovalori di una matrice Hermitiana sono reali.

1.3.3 Caratteristiche delle matrici

Diamo anche la definizione di alcune quantità matriciali molto importanti.

Definizione 1.12 (Rango)**.**
Il **rango** di una matrice **A** è la taglia massimale dell'insieme di vettori colonna (o riga, è equivalente) linearmente indipendenti. É possibile dimostrare che il rango di una matrice è uguale al numero di autovalori non nulli (considerandone anche la molteplicità) e che una matrice è invertibile solo se ha rango massimo rispetto alla dimensione dello spazio.

Definizione 1.13 (Traccia)**.**
La **traccia** di una matrice **A** è la somma degli elementi sulla sua diagonale:

$$\mathrm{Tr}(\mathbf{A}) = \sum_i (\mathbf{A})_{ii}$$

Si può dimostrare che la traccia di una matrice è uguale alla somma dei suoi autovalori.

Definizione 1.14 (Determinante)**.**
Il **determinante** di una matrice **A** è, geometricamente parlando, una quantità che identifica la scala di modifica dei volumi applicando **A**, dove il segno del determinante indica anche se preserva o meno l'orientazione degli assi. Viene definito ricorsivamente in maniera costruttiva:

- se $\dim(\mathbf{A}) = 1$, allora $\det(\mathbf{A}) = a_{11}$;
- se $\dim(\mathbf{A}) = 2$, allora $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$;
- se $\dim(\mathbf{A}) = 3$, allora $\det(\mathbf{A}) = a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$, ottenibile tramite **formula di Sarrus**;
- se $\dim(\mathbf{A}) > 3$, possiamo usare lo **sviluppo di Laplace**[1] per calcolare il determinante: scelta una riga i, si ha che

  $$\det(\mathbf{A}) = \sum_{j=1}^{d} (-1)^{i+1}\, a_{ij}\mathbf{C}_{ij}$$

  dove $\mathbf{C}_{ij}$ è il **complemento algebrico**, ossia è il determinante della matrice di ordine $d-1$ ottenuta dalla matrice **A** rimuovendo la riga $i-$esima e la colonna $j-$esima.

Le proprietà interessanti del determinante sono molteplici, ma noi vogliamo evidenziare che

- il determinante di una matrice sarà nullo se sono presenti delle righe (o delle colonne) linearmente dipendenti tra loro;
- è uguale al prodotto degli autovalori di **A** contati con le rispettive molteplicità;
- vale il **Teorema di Binet**, ossia

  $$\det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{AB})$$

  da cui segue che una matrice è invertibile se e solamente se $\det(\mathbf{A}) \neq 0$ e, in tal caso, vale che $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$;
- il determinante è invariante per trasposizione.

Inoltre, il determinante permette di calcolare esplicitamente gli autovalori come le radici del **polinomio caratteristico** di **A**:

$$p(x) = \det(\mathbf{A} - x\mathbb{I})$$

[1] Tecnicamente, la definizione del determinante è data tramite formula di Leibniz, usando le permutazioni, ma noi salteremo qualche passaggio per dare direttamente una definizione operativa più facilmente applicabile.

CAPITOLO 2

SPAZI DI PROBABILITÀ

Prima di spostarci in ambito quantistico, è importante assicurarci di comprendere i sistemi classici definiti attraverso una distribuzione di probabilità, dato che essi sono molto più intuitivi e facilmente immaginabili rispetto alla loro controparte quantistica.

KeyWords: vettori probabiità, notazione di Dirac, operazioni deterministiche, operazioni probabilistiche, sistemi singoli e multipli

2.1 Sistemi singoli

Supponiamo di avere un **sistema** che contiene delle informazioni: più in particolare, assumiamo che tale sistema possa trovarsi in uno stato qualsiasi tra un numero finito di **stati classici** ad ogni tempo, dove con "stato classico" intendiamo, intuitivamente, una configurazione che sia possibile riconoscere e descrivere senza ambiguità. Chiameremo questo insieme di configurazioni **spazio degli stati** Ω, mentre identificheremo il nostro sistema con X. Riportiamo ora qualche esempio:

- il sistema più semplice a cui possiamo pensare è sicuramente il *bit*, che ha spazio degli stati $\Omega = \{0, 1\}$;
- possiamo definire un sistema a partire dalle nucleobasi del DNA, identificando lo spazio degli stati come $\Omega = \{adenina, citosina, guanina, timina, uracile\}$;
- se giochiamo a DnD, possiamo definire uno spazio di probabilità sugli esiti del lancio del nostro dado da 20 facce, ossia sullo spazio degli stati $\Omega = \{1, \dots, 20\}$

e così via. Oltre ad assumere che Ω sia finito, assumeremo anche che non sia vuoto, dato che non ha senso per un sistema fisico non avere possibili configurazioni.

Ora, se X è in uno stato *noto*, siamo certi della sua configurazione e delle sue proprietà, solo che molte volte, nel processare informazioni, la conoscenza di X è *probabilistica*, ossia non conosciamo per certo in che stato sia, ma solo quali sono le probabilità associate alle possibili configurazioni, rendendo così gli stati classici degli **stati probabilistici**.
Per fissare le idee, supponiamo che X sia un bit e supponiamo che tale sistema si trovi nello stato classico 0 con probabilità $3/4$ e nello stato 1 con probabilità $1/4$: possiamo rappresentare queste informazioni che abbiamo come

$$\mathbb{P}(X = 0) = \frac{3}{4} \qquad \mathbb{P}(X = 1) = \frac{1}{4}$$

o, in maniera più compatta, come

$$\begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix}$$

dove le probabilità sono ordinate nello stesso ordine in cui definiamo lo spazio degli stati, ossia $\Omega = \{0, 1\}$.
In generale, possiamo rappresentare un qualsiasi stato probabilistico di un sistema definito su un qualsiasi spazio degli stati tramite un **vettore di probabilità**, le cui entrate possono essere ordinate in qualsiasi modo ci sia più utile, a patto di essere fatto in accordo con lo spazio degli stati.

Definizione 2.1. Un vettore di probabilità è un vettore colonna che descrive lo stato probabilistico di un sistema ed è tale che

1. tutte le entrate sono numeri reali non negativi;
2. la somma delle entrate è uguale a 1.

Viceversa, un qualsiasi vettore colonna che soddisfi queste due proprietà può essere considerato una rappresentazione di uno stato probabilistico.

Ma cosa succede quando misuriamo un sistema che si trova in uno stato probabilistico? Dato che *misurare* un sistema vuol dire interagire con esso e riconoscere inequivocabilmente lo stato in cui si trova, ma, dato che non è possibile vedere uno stato probabilistico, la nostra misurazione darà come output un solo stato classico tra quelli possibili, cambiando così la distribuzione di probabilità degli stati. Più in dettaglio, supponiamo di trovare X nello stato $a \in \Omega$: allora, dopo la misurazione, il vettore probabilità associato al sistema sarà un vettore colonna con tutte entrate nulle ad eccezione di quella relativa ad a, che invece sarà 1. Un vettore probabilità con questa forma è un vettore che esprime *certezza assoluta* riguardo allo stato del sistema e, pertanto, lo identificheremo come il *ket* dello stato corrispondente, ossia $|a\rangle$. Vettori di questo tipo sono anche chiamati **vettori della base standard**.

Definizione 2.2 (**Notazione di Dirac**)**.** Identificheremo con $|\varphi\rangle$ il vettore colonna

$$\begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix}$$

In spazi a valori reali, identificheremo con $\langle\varphi|$ il vettore riga

$$(\varphi_1, \ldots, \varphi_n)$$

ossia è il trasposto di $|\varphi\rangle$. Vedremo più avanti una definizione più dettagliata e una spiegazione di come questa notazione sia ottimale per discutere di meccanica quantistica.

Prendiamo nuovamente il nostro bit: la nostra base standard sarà formata dunque da

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

e osserviamo che qualsiasi vettore colonna di dimensione 2 può essere espresso come combinazione lineare di questi due vettori:

$$\begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix} = \frac{3}{4}|0\rangle + \frac{1}{4}|1\rangle$$

Questo fatto si può facilmente generalizzare a qualsiasi spazio degli stati classico, ossia è sempre possibile scrivere un vettore colonna come una combinazione lineare dei vettori della base standard, i cui vettori sono associati agli stati classici.

Esercizio 2.1. Supponiamo di avere una moneta bilanciata, di lanciarla in aria, riprenderla e coprirla con la mano prima di poterla guardare. Qual è il vettore di probabilità prima di guardare la moneta? Supponendo che sia uscita testa, come sarà il vettore delle probabilità dopo averla guardata?
Supponiamo ora di ricoprire la moneta con la mano e di chiedere a una persona appena entrata informazioni sulla moneta: dal suo punto di vista, quale sarà il vettore probabilità?

Passiamo ora alle possibili operazioni che possiamo applicare a degli stati probabilistici iniziando dalle **operazioni deterministiche** $f : \Omega \to \Omega$, per cui, dato uno stato classico $a \in \Omega$, otteniamo $f(a)$ senza ambiguità. Ad esempio, se $\Omega = \{0, 1\}$, possiamo definire quattro funzioni deterministiche, il cui comportamento è riportato nelle tabelle seguenti:

a	$f_0(a)$
0	0
1	0

a	$id(a)$
0	0
1	1

a	**NOT**(a)
0	1
1	0

a	$f_1(a)$
0	1
1	1

La prima funzione e l'ultima sono le funzioni costanti pari a 0 e 1 rispettivamente, mentre le funzioni nel mezzo sono *bilanciate*, dato che i possibili output appaiono con uguale frequenza. Nello specifico, la seconda funzione è la funzione identità, mentre la terza è la funzione negazione, la cui azione equivale al cosiddetto *bit flip*.

Tutte le operazioni deterministiche possono essere espresse sotto forma matriciale, così da poter trattare la loro applicazione come una moltiplicazione matrice-vettore. In particolare, fissata una certa funzione $f : \Omega \to \Omega$, con $n = \dim(\Omega)$, una matrice $\mathbf{M}$ di dimensioni $n \times n$ esprime f se verifica

$$\mathbf{M}\,|a\rangle = |f(a)\rangle$$

per ogni $a \in \Omega$. Come abbiamo discusso nell'introduzione matematica, per ogni funzione f, una matrice di questo tipo esiste sempre ed è unica: la denoteremo pertanto come $\mathbf{M}_f$.
Per esempio, riprendendo le funzioni definite poco sopra, possiamo scrivere

$$\mathbf{M}_{id} \equiv \mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{M}_0 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Un altro modo per esprimere tale operazioni è tramite il **prodotto esterno**: per una scelta arbitraria di due stati classici in Ω, moltiplicare una colonna $|b\rangle$ per una riga $\langle a|$ dà origine a una matrice quadrata le cui entrate sono tutte nulle ad eccezione dell'entrata (b, a), che invece vale 1. Ad esempio:

$$|0\rangle\langle 1| = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Con questa notazione, otteniamo che, per una qualsiasi funzione $f : \Omega \to \Omega$, possiamo esprimere la matrice $\mathbf{M}$ come

$$\mathbf{M} = \sum_{a \in \Omega} |f(a)\rangle\langle a|$$

Infatti, possiamo osservare che per ogni $b \in \Omega$ si ha

$$\mathbf{M}\,|b\rangle = \left(\sum_{a \in \Omega} |f(a)\rangle\langle a| \right) |b\rangle = \sum_{a \in \Omega} |f(a)\rangle\langle a|b\rangle = |f(b)\rangle$$

dato che $\langle a|b\rangle$ corrisponde al **prodotto scalare** (o **prodotto interno**) tra i due vettori ed è tale che

$$\langle a|b\rangle = \begin{cases} 1 & \text{se } a = b \\ 0 & \text{se } a \neq b \end{cases}$$

Prima di passare a un altro tipo di operazioni, è interessante osservare perché la notazione di Dirac faccia riferimento a *bra* e *ket*: tali nomi, infatti, se letti di fila nel prodotto interno, danno luogo alla parola "bracket", ossia parentesi.

Esercizio 2.2. Scrivere le matrici corrispondenti alla funzione **NOT** e alla funzione costante a 1.

Oltre alle operazioni deterministiche, abbiamo anche le **operazioni probabilistiche**.
Ad esempio, supponiamo di definire la seguente operazione su un bit: se il bit è nello stato classico 0, allora non facciamo nulla; se invece è nello stato classico 1, allora invertiamolo con probabilità $1/2$. Possiamo rappresentare questa matrice come

$$\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$$

Per verificare che fa esattamente ciò che ci aspettiamo ci basta moltiplicarla con i vettori della base standard.

Per uno spazio degli stati arbitrario, possiamo descrivere l'insieme di tutte le operazioni probabilistiche come quelle rappresentabili da **matrici stocastiche**.

Definizione 2.3. Una matrice è detta stocastica se tuttte le sue entrate sono numeri reali non negativi e se la somma delle entrate di una colonna è uguale a 1, per tutte le colonne. Equivalentemente, una matrice è stocastica se le sue colonne sono vettori di probabilità.

Le matrici stocastiche sono indispensabili per poter analizzare e formulare situazioni in cui ci sono degli elementi randomici che interferiscono con l'evoluzione del sistema e in cui ci serve mantenere il sistema in uno stato probabilisitico: infatti, le matrici stocastiche sono le uniche che mandano i vettori di probabilità in altri vettori di probabilità e, viceversa, se una matrice manda vettori di probabilità in altri vettori di probabilità allora è una matrice stocastica.
Possiamo interpretare le operazioni probabilistiche come scelte aleatorie di operazioni deterministiche. Ad esempio, riprendendo la matrice appena definita, possiamo vedere come essa equivalga ad applicare una funzione costante a 0 con probabilità $1/2$ o una funzione identità sempre con probabilità $1/2$:

$$\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Possiamo chiederci ora come potremmo fare se volessimo applicare più operazioni probabilistiche in sequenza, dato che esse preservano la forma di stato probabilistico. Supponiamo allora di voler applicare $\mathbf{M}_1, \ldots, \mathbf{M}_m$ al nostro sistema X con spazio degli stati Ω. Supponiamo inoltre che, inizialmente, il nostro sistema sia descritto da un qualche stato probabilistico $|\varphi\rangle$. Se la prima operazione che effettuiamo è $\mathbf{M}_1$, otteniamo come risultato un vettore probabilistico $\mathbf{M}_1 |\varphi\rangle$, a cui poi possiamo applicare la nostra seconda operazione $\mathbf{M}_2$:

$$\mathbf{M}_2(\mathbf{M}_1 |\varphi\rangle) = (\mathbf{M}_2\mathbf{M}_1) |\varphi\rangle$$

dove l'uguaglianza segue dalla proprietà associativa del prodotto matriciale. Osserviamo che, per come è definita la matrice stocastica, il prodotto di due matrici stocastiche è a sua volta una matrice stocastica e dunque il prodotto $\mathbf{M}_2\mathbf{M}_1$ equivale a comporre le due operazioni probabilistiche nell'ordine appena descritto. Generalizzando, se applichiamo in ordine $\mathbf{M}_1, \ldots, \mathbf{M}_m$, la matrice totale di tutte le operazioni che applichiamo è data da

$$\mathbf{M}_m \cdots \mathbf{M}_1$$

Osserviamo che l'ordine delle operazioni nel prodotto è "al contrario", dato che il suo risultato viene applicato al vettore di partenza tramite una moltiplicazione a destra, e che solitamente il prodotto matriciale non è commutativo. Ad esempio, se

$$\mathbf{M}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad \text{e} \qquad \mathbf{M}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

allora

$$\mathbf{M}_2\mathbf{M}_1 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \mathbf{M}_1\mathbf{M}_2$$

Dunque è importante stare attenti all'ordine in cui applichiamo le operazioni probabilistiche al nostro sistema.

2.2 Sistemi multipli

Finora abbiamo analizzato il caso in cui abbiamo un solo sistema, ma, nella realtà, l'elaborazione di dati e di informazioni richiede molto più frequentemente di operare su *più sistemi* allo stesso tempo. Solitamente questi sistemi, detti **sistemi multipli**, sono costruiti a partire da sistemi più piccoli e possono essere studiati in due modi differenti:

(a) se consideriamo il sistema nel suo *insieme*, possiamo trattarlo come un sistema singolo, così da poter applicare quanto visto fin qui;

(b) altrimenti, possiamo decidere di interagire in maniera diversa con i vari sotto-sistemi che lo compongono, così da poter sfruttare delle proprietà diverse che li caratterizzano.

Per semplicità, iniziamo discutendo di sistemi multipli formati da due sistemi: supponiamo dunque di avere un sistema X con spazio degli stati classico Ω e Y sistema con spazio degli stati classico Λ, dove, come in precedenza, stiamo assumendo che Ω e Λ siano entrambi finiti e non vuoti. Osserviamo che i due spazi degli stati potrebbero anche essere uguali, ma ciò non inficia in alcun modo l'esposizione dei concetti che seguono. Se immaginiamo ora di prendere entrambi i sistemi e di volerli analizzare insieme, possiamo scrivere il loro sistema multiplo come (X, Y), il cui spazio degli stati classico è formato dal prodotto cartesiano di Ω e Λ, ossia

$$\Omega \times \Lambda = \{(a, b) : a \in \Omega \text{ e } b \in \Lambda\}$$

In termini semplici, il prodotto cartesiano è la nozione matematica che rappresenta l'idea di vedere elementi di diversi insiemi come un insieme solo. Nel nostro caso, dire che (X, Y) è nello stato (a, b) equivale a dire che il sistema X si trova nello stato $a \in \Omega$, mentre Y si trova nello stato $b \in \Lambda$; e viceversa, se lo stato di X è $a \in \Omega$ e lo stato di Y è $b \in \Lambda$, allora lo stato del sistema multiplo (X, Y) è (a, b).

Generalizzando, se abbiamo $n > 0$ sistemi $X_1, \dots, X_n$, rispettivamente con spazi degli stati $\Omega_1, \dots, \Omega_n$, otteniamo il sistema multiplo $(X_1, \dots, X_n)$ con spazio degli stati

$$\Omega_1 \times \cdots \times \Omega_n = \{(a_1, \dots, a_n) : a_1 \in \Omega_1, \dots, a_n \in \Omega_n\}$$

Osserviamo che, molto spesso, è utile rappresentare uno stato classico $(a_1, \dots, a_n)$ come una **stringa** $a_1 \cdots a_n$, specie se gli spazi degli stati $\Omega_1, \dots, \Omega_n$ sono associati a insiemi di *simboli* o *caratteri*. Di solito, ci si riferisce ai simboli usati per costruire le stringhe come elementi di un **alfabeto**, che però ha la stessa definizione matematica di uno spazio degli stati, ossia si tratta di un insieme finito e non vuoto.
Ad esempio, supponiamo che $X_1, \dots, X_{10}$ siano bit e che dunque gli spazi degli stati siano tutti uguali:

$$\Omega_1 = \Omega_2 = \cdots = \Omega_{10} = \{0, 1\}$$

Ci riferiremo a tale spazio $\{0, 1\}$ come **alfabeto binario**. All'interno dello spazio del sistema multiplo $(X_1, \dots, X_{10})$ ci sono quindi $2^{10} = 1024$ stati classici appartenenti all'insieme

$$\Omega_1 \times \Omega_2 \times \cdots \times \Omega_{10} = \{0, 1\}^{10}$$

In formato di stringa, gli stati classici appaiono come segue

$$\begin{array}{c} 0000000000 \\ 0000000001 \\ 0000000010 \\ 0000000011 \\ 0000000100 \\ \vdots \\ 1111111111 \end{array}$$

Come nel caso di un sistema singolo, anche in un sistema multiplo possiamo parlare di **stati probabilistici**, in cui ad ogni stato classico del sistema multiplo viene associata una probabilità di realizzazione. Dunque, nuovamente, è importante definire un ordine per gli elementi dello spazio degli stati multiplo, così da ordinare analogamente le entrate del vettore delle probabilità. L'ordinamento più utilizzato è l'**ordine lessicografico**: le entrate di ogni $n-$upla (o, equivalentemente, gli elementi di ogni stringa) sono ordinati in base alla loro importanza, la quale diminuisce da sinistra verso destra.

Ad esempio, supponiamo che X e Y siano bit, così che i loro spazi degli stati siano $\Omega = \Lambda = \{0, 1\}$. Allora, un possibile stato probabilistico del sistema (X, Y) è

$$\begin{array}{l} \mathbb{P}((X, Y) = (0, 0)) = \frac{1}{2} \\ \mathbb{P}((X, Y) = (0, 1)) = 0 \\ \mathbb{P}((X, Y) = (1, 0)) = 0 \\ \mathbb{P}((X, Y) = (1, 1)) = \frac{1}{2} \end{array}$$

dove stiamo supponendo che gli unici stati verificabili siano quelli in cui i sistemi si trovano nello stato stato singolo. In casi come questo, diremo che i due sistemi sono **correlati** tra loro. Per quanto riguarda il relativo vettore delle probabilità si ha

$$\begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \\ \frac{1}{2} \end{pmatrix} \begin{array}{ll} \leftarrow & \text{probabilità associata allo stato 00} \\ \leftarrow & \text{probabilità associata allo stato 01} \\ \leftarrow & \text{probabilità associata allo stato 10} \\ \leftarrow & \text{probabilità associata allo stato 11} \end{array} \tag{2.1}$$

Un tipo speciale di stato probabilistico è quello che riguarda due sistemi *indipendenti*. Intuitivamente, due sistemi sono indipendenti se scoprire lo stato classico di uno dei due sistemi non influenza in alcun modo la nostra conoscenza (i.e. le probabilità) dello stato classico dell'altro sistema. Matematicamente, invece, diremo che due sistemi X, Y, con rispettivi spazi degli stati Ω, Λ, sono **indipendenti** se

$$\mathbb{P}((X,Y)=(a,b)) \equiv \mathbb{P}(X=a, Y=b) = \mathbb{P}(X=a)\mathbb{P}(Y=b) \tag{2.2}$$

per ogni scelta di $a \in \Omega, b \in \Lambda$.
Per esprimere tale condizione in termini di vettori di probabilità, assumiamo di fissare uno stato probabilistico di (X, Y) e che esso sia descritto da un qualche vettore di probabilità, scritto in notazione di Dirac come

$$\sum_{(a,b)\in\Omega\times\Lambda} p_{ab} \,|ab\rangle$$

La condizione (2.2) per l'indipendenza equivale allora all'esistenza di due vettori

$$|\varphi\rangle = \sum_{a\in\Omega} q_a \,|a\rangle \qquad \text{e} \qquad |\psi\rangle = \sum_{b\in\Lambda} r_b \,|b\rangle \tag{2.3}$$

che rappresentino le probabilità associate agli stati classici di X e Y e che tali che

$$p_{ab} = q_a r_b \tag{2.4}$$

per ogni $a \in \Omega, b \in \Lambda$.
Ad esempio, lo stato probabilistico di un paio di bit (X, Y) rappresentato dal vettore

$$\frac{1}{6}\,|00\rangle + \frac{1}{12}\,|01\rangle + \frac{1}{2}\,|10\rangle + \frac{1}{4}\,|11\rangle$$

è uno stato in cui X e Y sono indipendenti. In particolare, la condizione di indipendenza è verificata per i seguenti vettori probabilità:

$$|\varphi\rangle = \frac{1}{4}\,|0\rangle + \frac{3}{4}\,|1\rangle \qquad \text{e} \qquad |\psi\rangle = \frac{2}{3}\,|0\rangle + \frac{1}{3}\,|1\rangle$$

come possiamo verificare tramite conti diretti.
D'altra parte, lo stato probabilistico definito in (2.1), che possiamo scrivere come

$$\frac{1}{2}\,|00\rangle + \frac{1}{2}\,|11\rangle$$

non identifica due sistemi tra loro indipendenti, che dunque risultano **correlati**.

Esercizio 2.3. Verificare che i sistemi X, Y su cui è definito lo stato (2.1) non sono indipendenti.

La condizione di indipendenza appena descritta può essere espressa più sinteticamente attraverso la nozione di **prodotto tensoriale**: dati due vettori

$$|\varphi\rangle = \sum_{a\in\Omega} \alpha_a \,|a\rangle \qquad \text{e} \qquad |\psi\rangle = \sum_{b\in\Lambda} \beta_b \,|b\rangle$$

il loro prodotto tensoriale $|\varphi\rangle \otimes |\psi\rangle$ è un nuovo vettore nello spazio congiunto $\Omega \times \Lambda$ definito come

$$|\varphi\rangle \otimes |\psi\rangle := \sum_{(a,b)\in\Omega\times\Lambda} \alpha_a \beta_b \,|ab\rangle$$

che in coordinate cartesiane equivale a

$$\underbrace{\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_m\end{pmatrix}}_{\dim=m} \otimes \underbrace{\begin{pmatrix}\beta_1\\ \vdots\\ \beta_k\end{pmatrix}}_{\dim=k} = \begin{pmatrix}\alpha_1\begin{pmatrix}\beta_1\\ \vdots\\ \beta_k\end{pmatrix}\\ \vdots\\ \alpha_m\begin{pmatrix}\beta_1\\ \vdots\\ \beta_k\end{pmatrix}\end{pmatrix} = \underbrace{\begin{pmatrix}\alpha_1\beta_1\\ \vdots\\ \alpha_1\beta_k\\ \alpha_2\beta_1\\ \vdots\\ \alpha_2\beta_k\\ \vdots\\ \alpha_m\beta_1\\ \vdots\\ \alpha_m\beta_k\end{pmatrix}}_{\dim=mk}$$

Spesso si omette il simbolo operatoriale e dunque denoteremo il prodotto tensoriale semplicemente come $|\varphi\rangle\,|\psi\rangle$. Dunque, possiamo riformulare la condizione di indipendenza come segue: due sistemi sono indipendenti se il vettore probabilità $|\pi\rangle$ del sistema multiplo (X, Y) può essere scritto come un prodotto tensoriale di un vettore probabilità $|\varphi\rangle$ in X e di un vettore probabilità $|\psi\rangle$ in Y, ossia

$$|\pi\rangle = |\varphi\rangle \otimes |\psi\rangle \equiv |\varphi\rangle\,|\psi\rangle$$

Se tale relazione è verificata, si dice che $|\pi\rangle$ è uno **stato prodotto**.
Osserviamo che per vettori della base standard vale che

$$|a\rangle\,|b\rangle = |ab\rangle \equiv |a, b\rangle$$

Un'importante proprietà del prodotto tensoriale è la **bilinearità**:

- assumendo fissato il secondo argomento, c'è linearità rispetto al primo

$$(|\varphi_1\rangle + |\varphi_2\rangle) \otimes |\psi\rangle = |\varphi_1\rangle \otimes |\psi\rangle + |\varphi_2\rangle \otimes |\psi\rangle$$
$$(\alpha\,|\varphi\rangle) \otimes |\psi\rangle = \alpha(|\varphi\rangle \otimes |\psi\rangle)$$

- assumendo fissato il primo argomento, c'è linearità rispetto al secondo

$$|\varphi\rangle \otimes (|\psi_1\rangle + |\psi_2\rangle) = |\varphi\rangle \otimes |\psi_1\rangle + |\varphi\rangle \otimes |\psi_2\rangle$$
$$|\varphi\rangle \otimes (\alpha\,|\psi\rangle) = \alpha(|\varphi\rangle \otimes |\psi\rangle)$$

da cui possiamo dunque osservare che gli scalari si "muovono" liberamente all'interno dei prodotti tensoriali.

Le nozioni di indipendenza e prodotto tensoriale si generalizzano facilmente nel caso in cui si abbiano tre o più sistemi. Siano $X_1, \ldots, X_n$ sistemi definiti, rispettivamente, su spazi degli stati $\Omega_1, \ldots, \Omega_n$. Allora il vettore probabilità del sistema multiplo $(X_1, \ldots, X_n)$ è uno stato prodotto se il vettore probabilità associato è tale che

$$|\pi\rangle = |\varphi_1\rangle \otimes \cdots \otimes |\varphi_n\rangle$$

dove $|\varphi_1\rangle, \ldots, |\varphi_n\rangle$ sono i vettori probabilità che descrivono $X_1, \ldots, X_n$. In questo caso, il prodotto tensoriale è definito ricorsivamente in termini di prodotto tensoriale tra due vettori:

$$|\varphi_1\rangle \otimes \cdots \otimes |\varphi_n\rangle \coloneqq (|\varphi_1\rangle \otimes \cdots \otimes |\varphi_{n-1}\rangle) \otimes |\varphi_n\rangle$$

dove stiamo assumendo che $n \geq 3$. Similmente al caso con due vettori, il prodotto tensoriale generalizzato è **multilineare**, ossia è lineare in qualsiasi dei suoi argomenti, a patto di fissare tutti gli altri. Un'altra importante generalizzazione è quella che riguarda i vettori della base standard: infatti

$$|a_1\rangle \otimes \cdots \otimes |a_n\rangle = |a_1 \cdots a_n\rangle \equiv |a_1, \ldots, a_n\rangle$$

Per quanto riguarda l'indipendenza, non ci addentrereremo molto nel concetto generalizzato al momento: ci basta sapere che se un sistema multiplo è in uno stato prodotto allora i sistemi $X_1, \ldots, X_n$ sono indipendenti.

Passando alle misurazioni dei sistemi multipli, dobbiamo innanzitutto scegliere come vogliamo interpretarlo: infatti, se lo consideriamo come un singolo sistema, possiamo semplicemente generalizzare quanto visto prima, a patto di supporre che *tutti i sistemi siano misurati*. Ad esempio, se lo stato probabilistico del sistema di due bit (X, Y) è descritto dal vettore probabilità

$$\frac{1}{2}|00\rangle + \frac{1}{2}|11\rangle$$

allora misurarlo darà come output 00 o 11 con uguale probabilità, ossia $1/2$. In ogni caso, dopo la misurazione modificheremo il vettore probabilità in accordo con lo stato ottenuto, rendendolo $|00\rangle$ oppure $|11\rangle$.

Supponiamo ora di voler invece misurare solo qualcuno dei sotto-sistemi del nostro sistema composito. Senza perdita di generalità, possiamo supporre di avere solo due sotto-sistemi e di misurare uno di essi: infatti, il caso più generale si riduce al caso di due sistemi se consideriamo i sotto-sistemi misurati come un primo sistema e quelli non misurati come un secondo sistema.

Dunque, più formalmente, siano X, Y sistemi con spazi degli stati classici Ω, Λ rispettivamente e assumiamo che si trovino in un qualche stato probabilistico come sistema composito. Decidiamo di misurare X e lasciare Y invariato (il caso a parte invertite è analogo per simmetria). Osserviamo innanzitutto che la probabilità di osservare uno stato classico $a \in \Omega$ misurando solo X deve essere consistente con le probabiltà che otterrremmo misurando anche Y, ossia deve valere che

$$\mathbb{P}(X = a) = \sum_{b \in \Lambda} \mathbb{P}((X, Y) = (a, b))$$

Tale formula definisce quella che è chiamata **densità marginale** di X. Dato che stiamo misurando solo X, in generale potrebbe esserci ancora dell'incertezza riguardo allo stato classico di Y, incertezza che non ci permette di aggiornare il vettore probabilità di (X, Y) in maniera arbitraria. Per mantenere tale insicurezza sullo stato di Y necessitiamo della **probabilità condizionata**, definita come

$$\mathbb{P}(Y = b | X = a) := \frac{\mathbb{P}(X = a, Y = b)}{\mathbb{P}(X = a)}$$

dove $\mathbb{P}(Y = b|X = a)$ indica la probabilità che si verifichi $Y = b$ *sapendo* che $X = a$. Osserviamo che affinché tale definizione sia ben posta, si deve avere $\mathbb{P}(X = a) \neq 0$, ma ciò non è una limitazione: infatti, se anche tale probabilità fosse zero non avremmo problemi, dato che semplicemente vuol dire che non possiamo mai osservare a e dunque non ha senso considerarlo a priori come condizionamento.
Per esprimere queste formule in termini di vettori probabilità possiamo considerare un vettore $|\psi\rangle$ che descrive lo stato di (X, Y)

$$|\psi\rangle = \sum_{(a,b) \in \Omega \times \Lambda} p_{ab} |ab\rangle$$

da cui otteniamo che ogni possibile risultato del misurare X si ottiene con probabilità

$$\mathbb{P}(X = a) = \sum_{b \in \Lambda} p_{ab}$$

Dunque, il vettore che rappresenta lo stato probabilistico del solo X è

$$\sum_{a \in \Omega} \left(\sum_{c \in \Lambda} p_{ac} \right) |a\rangle$$

Fissato il risultato $a \in \Omega$ della misurazione di X, lo stato probabilistico di Y viene aggiornato in accordo con la definizione di probabilità condizionata, ossia viene rappresentato dal vettore

$$|\pi_a\rangle = \frac{\sum_{b \in \Lambda} p_{ab} \, |b\rangle}{\sum_{c \in \Lambda} p_{ac}}$$

da cui otteniamo che il nuovo stato probabilistico del nostro sistema composito (X, Y) è $|a\rangle \otimes |\pi_a\rangle$.

Esercizio 2.4. Supponiamo di avere X definito su $\Omega = \{0, 1\}$, Y definito su $\{1, 2, 3\}$ e sistema multiplo (X, Y) descritto dallo stato probabilistico

$$|\psi\rangle = \frac{1}{2} |0, 1\rangle + \frac{1}{12} |0, 3\rangle + \frac{1}{12} |1, 1\rangle + \frac{1}{6} |1, 2\rangle + \frac{1}{6} |1, 3\rangle$$

Determinare le probabilità dei due possibili output nel misurare X e i possibili stati probabilistici di Y dopo tale misurazione.

Per concludere questa discussione dei sistemi classici, consideriamo ora le operazioni che possiamo applicare a un sistema multiplo che si trovi in uno stato probabilistico. Supponendo sempre di avere due sistemi X, Y, consideriamo il loro sistema composito (X, Y): da quanto visto nella sezione precedente, le operazioni classiche applicabili sono rappresentate da una matrice stocastica le cui righe e colonne sono indicizzate come gli elementi del prodotto cartesiano $\Omega \times \Lambda$.
Ad esempio, supponiamo che X e Y siano bit e consideriamo l'operazione qui definita:

Se $X = 1$ allora applichiamo un **NOT** *su Y; altrimenti non facciamo nulla.*

Tale operazione è nota come **controlled-NOT**, dove X è il **bit di controllo** e Y è il **bit target**. La matrice associata a tale operazione è

$$\mathbf{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

ed agisce sulla base standard nel modo seguente:

$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto |01\rangle \\ |10\rangle &\mapsto |11\rangle \\ |11\rangle &\mapsto |10\rangle \end{aligned}$$

Se dovessimo scambiare i ruoli di X e Y otterremo la seguente matrice

$$\mathbf{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

che agisce sulla base standard nel modo seguente:

$$\begin{aligned} |00\rangle &\mapsto |00\rangle \\ |01\rangle &\mapsto |11\rangle \\ |10\rangle &\mapsto |10\rangle \\ |11\rangle &\mapsto |01\rangle \end{aligned}$$

Lo stesso procedimento può essere applicato ad un numero qualsiasi di sistemi: immaginiamo ad esempio di avere tre bit e di volerli incrementare modulo 8, ossia identificheremo i numeri con le rispettive stringhe in binario, le modificheremo aggiugendo 1 e poi prenderemo il resto della divisione per 8. Possiamo scrivere tale operazione come

$$\sum_{k=0}^{7} |(k+1) \mod 8\rangle\langle k| \overset{in\ bin.}{=} |001\rangle\langle 000| + |010\rangle\langle 001| + \cdots + |111\rangle\langle 110| + |000\rangle\langle 111|$$

da cui ricaviamo la matrice

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Osserviamo che potevamo definire tale matrice direttamente per entrate secondo la seguente regola:

$$\mathbf{M}_{ij} = \begin{cases} 1 & \text{se } |j\rangle = \mathbf{M}\,|i\rangle \\ 0 & \text{altrimenti} \end{cases}$$

Esercizio 2.5. Data l'operazione così definita

Con uguale probabilità, pari a 1/2, porre Y uguale a X o X uguale a Y

definire la matrice associata e determinarne l'azione sui vettori della base standard.

Supponiamo ora di avere dei sistemi multipli e di applicare ai suoi sottosistemi delle operazioni *indipendenti*. Ad esempio, prendiamo il nostro solito sistema (X, Y) formato da due bit e assumiamo di voler applicare una matrice stocastica $\mathbf{M}$ su X e un'altra matrice stocastica $\mathbf{N}$ su Y. Sorge ora spontaneo chiedersi dunque come possiamo rappresentare questo tipo di operazione dal punto di vista del sistema composito: a tal fine, definiamo il **prodotto tensoriale per matrici**.

Definizione 2.4. Date due matrici $\mathbf{M}, \mathbf{N}$, rispettivamente di dimensioni $m \times m$ e $n \times n$, tali che

$$\mathbf{M} = \sum_{a,b \in \Omega} \alpha_{ab} \left|a\right\rangle \left\langle b\right| \qquad \mathbf{N} = \sum_{c,d \in \Lambda} \beta_{cd} \left|c\right\rangle \left\langle d\right|$$

definiamo il loro prodotto tensoriale $\mathbf{M} \otimes \mathbf{N}$ come la matrice di dimensioni $mn \times mn$ tale che

$$\mathbf{M} \otimes \mathbf{N} = \sum_{a,b \in \Omega} \sum_{c,d \in \Lambda} \alpha_{ab} \beta_{cd} \left|ac\right\rangle \left\langle bd\right|$$

Per un numero maggiore di matrici, il prodotto tensoriale può essere invece definito ricorsivamente.

In coordinate cartesiane, ciò equivale a

$$\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1m} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mm} \end{bmatrix} \otimes \begin{bmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{kk} \end{bmatrix} = \begin{bmatrix} \alpha_{11} \begin{bmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{kk} \end{bmatrix} & \cdots & \alpha_{1m} \begin{bmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{kk} \end{bmatrix} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} \begin{bmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{kk} \end{bmatrix} & \cdots & \alpha_{mm} \begin{bmatrix} \beta_{11} & \cdots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \beta_{k1} & \cdots & \beta_{kk} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{11}\beta_{11} & \cdots & \alpha_{11}\beta_{1k} & & \alpha_{1m}\beta_{11} & \cdots & \alpha_{1m}\beta_{1k} \\ \vdots & \ddots & \vdots & \vdots & & & \\ \alpha_{11}\beta_{k1} & \cdots & \alpha_{11}\beta_{kk} & & \alpha_{1m}\beta_{k1} & \cdots & \alpha_{1m}\beta_{kk} \\ & \vdots & & \ddots & & \vdots & \\ \alpha_{m1}\beta_{11} & \cdots & \alpha_{m1}\beta_{1k} & & \alpha_{mm}\beta_{11} & \cdots & \alpha_{mm}\beta_{1k} \\ \vdots & \ddots & \vdots & \vdots & & & \\ \alpha_{m1}\beta_{k1} & \cdots & \alpha_{m1}\beta_{kk} & & \alpha_{mm}\beta_{k1} & \cdots & \alpha_{mm}\beta_{kk} \end{bmatrix}$$

Il prodotto tensoriale di due matrici agisce su uno stato prodotto nel seguente modo:

$$(\mathbf{M} \otimes \mathbf{N})(\left|\varphi\right\rangle \otimes \left|\psi\right\rangle) = (\mathbf{M} \left|\varphi\right\rangle) \otimes (\mathbf{N} \left|\psi\right\rangle)$$

e dunque anche in questo caso il prodotto tensoriale rappresenta indipendenza. Inoltre, il prodotto tensoriale è **moltiplicativo**, dato che per ogni $\mathbf{M}_1, \ldots, \mathbf{M}_n, \mathbf{N}_1, \ldots, \mathbf{N}_n$ si ha

$$(\mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_n)(\mathbf{N}_1 \otimes \cdots \otimes \mathbf{N}_n) = (\mathbf{M}_1\mathbf{N}_1) \otimes \cdots \otimes (\mathbf{M}_n\mathbf{N}_n)$$

a patto ovviamente che i prodotti $\mathbf{M}_i\mathbf{N}_i$ siano ben definiti, ed è tale che il prodotto tensoriale di matrici stocastiche sia a sua volta una matrice stocastica.

Una situazione che incontreremo spesso è quella in cui a uno dei due sottosistemi non viene applicata alcuna operazione, mentre all'altro sì: la non-interazione sul primo sistema è dunque rappresentata dalla matrice identità $\mathbb{I}$, mentre la matrice totale applicata al sistema sarà il prodotto tensoriale tra l'identità e l'operazione applicata (ovviamente tenendo conto dell'ordine dei bit nell'ordinare le matrici). Ad esempio, supponiamo di resettare il bit X a 0 e di non agire sul bit Y: allora, l'operazione probabilistica (in realtà deterministica) su (X, Y) è rappresentata dalla matrice

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

CAPITOLO 3

SISTEMI QUANTISTICI

Siamo ora pronti per spostarci in ambito quantistico, introducendo così il contesto in cui lavoreremo durante questo corso. Come nel capitolo (sezione forse a livello concettuale ndr) precendente, partiremo dal considerare un sistema quantistico semplice formato da un solo elemento, per poi procedere alla sua generalizzazione.

Anche in questo caso, assumeremo che il nostro sistema sia associato a uno spazio degli stati classico finito e non vuoto.

KeyWords: qubit, sistemi singoli e multipli, misurazioni, entanglement, operatori unitari

3.1 Sistemi quantistici a un qubit

Iniziamo definendo uno **stato quantistico** relativo a un sistema, il quale è rappresentato da un vettore colonna a entrate complesse tale che la somma dei quadrati dei suoi moduli sia pari a 1. Formalmente, si tratta di un vettore $|\psi\rangle$ tale che

$$\sqrt{\sum_{k=1}^{n} |\psi_k|^2} = 1$$

e osserviamo che stiamo dunque richiedendo che i vettori siano dotati di *norma unitaria*.

Il termine **qubit** (ossia *quantum bit*) fa riferimento all'unità di informazione quantistica e, convenzionalmente, è associato a un sistema quantistico con spazio degli stati classici $\{0, 1\}$: si tratta dunque di un analogo al bit classico che però viene rappresentato come stato quantistico. Alcuni esempi di qubit sono la **base standard**

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{e} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

dove associeremo agli stati $|0\rangle, |1\rangle$, rispettivamente, il fatto che il sistema sia nello stato classico 0 o 1, oppure

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle \tag{3.1}$$

$$\frac{1+2i}{3} |0\rangle - \frac{2}{3} |1\rangle \begin{pmatrix} \frac{1+2i}{3} \\ -\frac{2}{3} \end{pmatrix} = \frac{1+2i}{3} |0\rangle - \frac{2}{3} |1\rangle \tag{3.2}$$

Questi due stati sono un esempio di **sovrapposizione** degli stati della base standard, che formalmente equivale a una combinazione lineare dei vettori della base.

Lo stato definito in (3.1) è anche chiamato $|+\rangle$ e fa parte di un'altra base del nostro spazio di Hilbert insieme allo stato $|-\rangle$, definito come

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{2}} |1\rangle$$

Osserviamo dunque che possiamo chiamare uno stato quantistico come ci risulta più consono al contesto; in ogni caso, la notazione più comune è quella di utilizzare[1] un generico ψ per indicare uno stato $|\psi\rangle$ arbitrario, che non sia necessariamente parte di una base, oppure usare la scrittura binaria dello stato classico corrispondente (a patto ovviamente che sia un intero non negativo o che lo spazio degli stati classici sia in biezione con i naturali). Inoltre, è utile osservare che, dato un generico stato $|\psi\rangle$, si possono far coincidere i suoi indici con gli elementi di un qualche spazio degli stati classici Ω, ottenenendo che il prodotto scalare

$$\langle a|\psi\rangle \quad a \in \Omega$$

equivale al coefficiente di $|\psi\rangle$ nell'entrata a. Ad esempio, per $\Omega = \{0, 1\}$ e per lo stato definito in (3.2), si ha

$$\langle 0|\psi\rangle = \frac{1+2i}{3} \quad \text{e} \quad \langle 1|\psi\rangle = -\frac{2}{3}$$

Consideriamo ora cosa succede se proviamo a misurare il nostro sistema quantistico, supponendo di misurare rispetto alla base standard (esistono anche altri riferimenti per la misurazione, ma concentriamoci al momento su quello più comune). Proprio come accade nel caso probabilistico, misurando uno stato quantistico si ottiene come output uno stato classico: possiamo dunque interpretare la misurazione come il mezzo di traduzione dall'informazione quantistica a quella classica, che ci permette di interagire a tutti gli effetti con le informazioni di un sistema quantistico.
La regola è semplice: dato uno stato $|\psi\rangle$, sovrapposizione banale o meno dei vettori della base standard, la probabilità di ottenere uno degli stati classici è pari al *quadrato del modulo* del coefficiente relativo allo stato quantistico associato allo stato classico. Per rendere più chiara questa regola, supponendo di avere uno stato

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$$

la probabilità di misurare lo stato classico 0 è $|\alpha_0|^2$ mentre la probabilità di misurare lo stato classico 1 è $|\alpha_1|^2$. Osserviamo che tali probabilità sono ben definite, dato che avevamo supposto che gli stati quantistici fossero rappresentati da vettori di norma unitaria. La caratteristica delle misurazioni di un sistema quantistico è però il cosiddetto **collasso**: infatti, una volta misurato il nostro sistema, esso assume lo stato classico misurato, perdendo tutte le informazioni dello stato quantistico in cui si trovava precedentemente alla misurazione.
Riprendiamo per un momento l'esempio della moneta che abbiamo visto per il caso classico e supponiamo di avere a disposizione una *moneta quantistica*: lanciata in aria la moneta e poi coperta, tutte le informazioni che abbiamo su di essa sono rappresentabili da uno stato quantistico e i possibili output della misurazione sono tutti probabilistici. Supponiamo ora di osservare la moneta, di interagire con essa, e di scoprire dunque che è uscita testa. Se ora arrivasse qualcun altro nella stanza e noi ci chiedessimo, coprendo nuovamente la moneta, in che stato sia il nostro sistema, la risposta sarebbe comunque testa, anche per un nuovo

[1] Che effettivamente è la notazione che abbiamo adottato anche a inizio sezione.

osservatore, a causa del collasso.
Data la natura probabilistica della misurazione e la presenza di questo effetto, è dunque necessario ripetere più volte la stessa computazione, ri-inizializzando di volta in volta il sistema, e tenere traccia dei risultati così da avere un output finale di maggioranza.
Un primo esempio di misurazione riguarda gli stati quantistici della base stessa: se infatti misuriamo $|0\rangle$ otteniamo lo stato classico 0 con probabilità pari a 1 e analogamente per $|1\rangle$, il che effettivamente è in accordo con la nostra identificazione di *ket* e omonimo stato classico.
Un secondo esempio consiste nel prendere lo stato quantistico $|+\rangle$ definito poco fa e calcolare le probabilità dei due possibili output:

$$\mathbb{P}(\text{l'output è } 0) = |\langle 0|+\rangle|^2 = \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}$$
$$\mathbb{P}(\text{l'output è } 1) = |\langle 1|+\rangle|^2 = \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}$$

É interessante osservare che se misuriamo lo stato $|-\rangle$ le probabilità sono esattamente le stesse:

$$\mathbb{P}(\text{l'output è } 0) = |\langle 0|-\rangle|^2 = \left| \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}$$
$$\mathbb{P}(\text{l'output è } 1) = |\langle 1|-\rangle|^2 = \left| -\frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2}$$

Da ciò possiamo dedurre che effettivamente, almeno per quanto riguarda le misurazioni nella base standard, gli stati $|+\rangle$ e $|-\rangle$ sono indistinguibili: vedremo tra poco che in realtà questi due stati sono perfettamente distinguibili, a patto di applicare prima un corretto operatore.

Ma che tipo di operatori possiamo davvero applicare affinché l'evoluzione del nostro sistema sia valida e ben definita? Innanzitutto, dato che si tratta di un'operazione tra stati, vogliamo che il nostro operatore sia **lineare** e, inoltre, che esso **conservi la norma**, così che uno stato quantistico di norma unitaria venga mappato in un altro stato quantistico di norma unitaria. C'è un ultima richiesta che vogliamo sia soddisfatta da questo operatore: vogliamo che **conservi l'ortogonalità** tra gli stati, ossia

$$\langle \psi|\varphi\rangle = 0 \Leftrightarrow \langle \psi|\mathbf{U}^\dagger\mathbf{U}|\varphi\rangle = 0$$

dove con $\mathbf{U}$ indichiamo tale operatore. Ciò che stiamo richiedendo è dunque che $\mathbf{U}^\dagger\mathbf{U} = \mathbb{I}$, ossia che $\mathbf{U}$ sia **unitario**. Osserviamo che tale condizione soddisfa anche le prime due condizioni richieste per tale operatore. Da questa osservazione, siamo in grado di capire perché utilizzare un operatore unitario, ma perché vogliamo che l'ortogonalità venga conservata? La risposta si può trovare nei **postulati della meccanica quantistica**, che ora procediamo a enunciare:

1. Le quantità osservabili della meccanica quantistica sono rappresentate da operatori hermitiani;

2. I possibili risultati di una misurazione sono gli autovalori dell'operatore associato. Inoltre, misureremo l'autovalore λ_i se e solo se il sistema è nello stato $|\lambda_i\rangle$;

3. Stati che possono essere misurati senza ambiguità sono rappresentati da vettori ortogonali;

4. Se $|\psi\rangle$ è lo stato-vettore di un sistema e misuriamo l'osservabile $\mathbf{L}$, la probabilità di ottenere il valore λ_i è
$$\mathbb{P}(\lambda_i) = \langle\psi|\lambda_i\rangle\langle\lambda_i|\psi\rangle = |\langle\psi|\lambda_i\rangle|^2$$
Usando questo principio risulta ovvio poter calcolare $\langle\mathbf{L}\rangle = \sum_i \lambda_i P(\lambda_i)$.

5. L'evoluzione temporale degli stati-vettori è unitaria.

Gli operatori unitari sono alla base dei *gate quantistici*, i quali possono essere tradotti materialmente in circuiti quantistici, ottenendo una struttura simile ai gate logici usati con i bit classici. Riportiamo dunque alcuni degli operatori più utlizzati che agiscono su un singolo qubit:

- Operatori di Pauli
Le quattro matrici di Pauli sono definite come segue
$$\mathbb{I} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \quad \mathbf{X} \equiv \sigma_x \coloneqq \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \quad \mathbf{Y} \equiv \sigma_y \coloneqq \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} \quad \mathbf{Z} \equiv \sigma_z \coloneqq \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$
Abbiamo già incontrato la matrice σ_x, che effettua il *bit flip* sui nostri qubit
$$\mathbf{X}|0\rangle = |1\rangle \qquad \mathbf{X}|1\rangle = |0\rangle$$
mentre l'operatore $\mathbf{Z}$ è conosciuto anche come *phase flip*, dato che
$$\mathbf{Z}|0\rangle = |0\rangle \qquad \mathbf{Z}|1\rangle = -|1\rangle$$

- Hadamard Gate
L'operatore di Hadamard è definito come
$$\mathbf{H} \coloneqq \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$
e ciò che fa è generare una sovrapposizione equiprobabile degli stati della base. Ad esempio:
$$\mathbf{H}|0\rangle = |+\rangle \qquad \mathbf{H}|1\rangle = |-\rangle$$
$$\mathbf{H}|+\rangle = |0\rangle \qquad \mathbf{H}|-\rangle = |1\rangle$$
Tornando alla domanda del paragrafo precedente, usando un Hadamard gate siamo in grado di distinguere perfettamente gli stati $|+\rangle, |-\rangle$: supponiamo che il nostro sistema sia preparato in uno di questi due stati, preparazione di cui però noi non abbiamo informazioni. Se misurassimo subito non otterremo alcun tipo di informazione sullo stato del nostro sistema, dato che entrambi gli esiti 0, 1 sono equiprobabili, ma se prima di misurare applichiamo un Hadamard otterremo con assoluta certezza un output di misurazione pari a 0 se lo stato preparato era $|+\rangle$ oppure 1 se lo stato preparato era $|-\rangle$, permettendo dunque una discriminazione perfetta e definendo la necessità di utilizzare dei vettori diversi dagli stati probabilistici.

- Operatori di Fase
Un operatore che modifica la fase di uno stato è del tipo
$$\mathbf{P}_\theta \coloneqq \begin{bmatrix}1 & 0\\ 0 & e^{i\theta}\end{bmatrix}$$
per qualsiasi scelta di $\theta \in \mathbb{R}$. Due dei più usati di questo tipo di operatori sono
$$\mathbf{S} \coloneqq \mathbf{P}_{\pi/2} = \begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix} \qquad \mathbf{T} \coloneqq \mathbf{P}_{\pi/4} = \begin{bmatrix}1 & 0\\ 0 & \frac{1+i}{\sqrt{2}}\end{bmatrix}$$

Dato che si tratta di semplici operatori, si possono facilmente comporre tramite moltiplicazione matriciale. Ad esempio, possiamo definire la seguente matrice

$$\mathbf{R} = \mathbf{HSH}$$

la cui particolarità è che coincide con la *radice dell'operatore* **X**. Tale affermazione si può dimostrare tramite conti diretti, ma noi preferiamo sfruttare questa occasione per presentarvi il Quantum Composer di Qiskit, uno strumento molto utile e intuitivo per costruire circuiti e visualizzarne l'effetto sui qubit. Osserviamo che il poter ottenere un **NOT** da una matrice come **R** è dovuto al fatto che siamo in ambito quantistico: infatti, sarebbe impossibile effettuare la radice di un elemento negativo se fossimo nel caso classico, ma data la definizione in campo complesso tale operazione è ben definita.

E non è solo in questo tipo di operazioni che la potenza della meccanica quantistica si presenta: infatti, dato uno stato quantistico, questo può evolvere in molti modi diversi, inclusi quelli che comportano *l'annullamento di alcune ampiezze*, rendendo nulla la probabilità di misurare il valore associato. Questo tipo di avvenimento è chiamato **interferenza quantistica** e può essere **distruttiva** o **costruttiva** a seconda che vada ad amplificare o a diminuire le ampiezze che determinano lo stato. Vedremo nel corso delle giornate come questa proprietà sia la chiave di volta di numerosi algoritmi quantistici, specie quelli basati sulla diffusione come Grover.

Concludiamo la presentazione dei sistemi quantistici a un qubit parlando della **fase**, in particolare defininendo una fase globale e una fase relativa.
Supponiamo di avere due stati $|\psi\rangle, |\varphi\rangle$ e che esista un numero $\alpha \in \mathbb{C}$ tale che $|\alpha| = 1$ e per cui

$$|\varphi\rangle = \alpha\,|\psi\rangle$$

Diremo allora che i due stati differiscono di una **fase globale** α. Analizziamo cosa succede se applichiamo una misurazione rispetto alla base standard di questi due vettori:

$$\mathbb{P}(|\psi\rangle \text{ è nello stato } a) = |\langle a|\psi\rangle|^2$$

$$\mathbb{P}(|\varphi\rangle \text{ è nello stato } a) = |\langle a|\varphi\rangle|^2 = |\alpha|^2|\langle a|\psi\rangle|^2 \overset{|\alpha|=1}{=} |\langle a|\psi\rangle|^2$$

Dunque, a livello di misurazione, i due stati danno uguali probabilità, proprio come accadeva nel caso degli stati $|+\rangle\,|-\rangle$. Viene quindi spontaneo chiedersi se anche in questo caso i due stati siano distinguibili applicando un qualche operatore unitario, però stavolta la risposta è negativa: infatti, se supponiamo di applicare l'operatore **U** otteniamo

$$\mathbf{U}\,|\psi\rangle \quad \text{e} \quad \mathbf{U}\,|\varphi\rangle = \alpha\mathbf{U}\,|\psi\rangle$$

e quindi anche le successive evoluzioni dei nostri stati continueranno ad avere un differenza di fase globale pari ad α. Dunque, gli stati che differiscono di una fase globale sono *perfettamente indistinguibili*: a prescindere dall'operazione, o dalla sequenza di operazioni, che applichiamo, essi continueranno a mantenere la loro relazione di fase, rendendo uguali in probabilità gli output della misurazione. Per tale motivo, due stati che differiscono di una fase globale sono considerati a tutti gli effetti *equivalenti* e vengono trattati come se fossero lo stesso stato. Ad esempio, gli stati

$$|-\rangle = \frac{1}{\sqrt{2}}\,|0\rangle - \frac{1}{\sqrt{2}}\,|1\rangle \quad \text{e} \quad -|-\rangle = -\frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle$$

differiscono per una fase globale pari a -1 e possono essere dunque considerati come un unico stato.

Al contrario, gli stati

$$|+\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \quad \text{e} \quad |-\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

non differiscono di una fase globale, ma di una **fase relativa**, dato che tale differenza non influisce su tutte le ampiezze ma solo su un sottoinsieme proprio di esse (nel nostro caso, solo sull'entrata relativa a $|1\rangle$).
Concludendo, possiamo dunque ignorare una qualsiasi differenza di fase tra due stati a patto che essa sia globale; se invece è relativa, gli stati sono distinti e vanno trattati come tali.

3.2 Sistemi quantistici a più qubit

Siamo ora pronti a passare a sistemi a più qubit e, proprio come nel caso a un qubit, la formulazione di questi sistemi è molto simile a quella già vista nel caso probabilistico.

Un sistema quantistico multiplo può essere considerato, nuovamente, come un singolo sistema composito, ed essere rappresentato quindi da un vettore colonna di norma unitaria, a patto ovviamente che esso abbia dimensione maggiore di 2. Per l'esattezza, il nuovo spazio avrà dimensione pari a $N = 2^n$, dove con n indicheremo il numero di sistemi singoli che compongono il sistema composito, con spazio degli stati classici formato dal prodotto cartesiano degli spazi degli stati classici dei singoli sottosistemi. Ad esempio, se X, Y sono due qubit, il sistema composito (X, Y) sarà associato a uno spazio degli stati classici pari a $\{0,1\} \times \{0,1\}$, che possiamo rappresentare tramite l'insieme di stringhe binarie $\{00, 01, 10, 11\}$. Allora i seguenti vettori sono esempi di stati quantistici nello spazio composito

$$\frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{6}}|01\rangle + \frac{i}{\sqrt{6}}|10\rangle + \frac{1}{\sqrt{6}}|11\rangle \qquad \frac{3}{5}|00\rangle - \frac{4}{5}|11\rangle \qquad |01\rangle$$

Riportiamo ora delle notazioni equivalenti, che saranno utilizzate a seconda del contesto per rendere più comprensibile il ragionamento:

$$|00\rangle \equiv |0\rangle|0\rangle \equiv |0\rangle \otimes |0\rangle \equiv |0\rangle_X |0\rangle_Y$$

oltre ovviamente alla scrittura vettoriale esplicita.

Analogamente al caso classico dunque, utilizzeremo il prodotto tensoriale per espandere i nostri sistemi e, in particolare, tale prodotto genera un nuovo stato quantistico valido e rappresenta **indipendenza** tra i sottosistemi. Più formalmente, supponiamo di avere due sistemi quantistici e che $|\varphi\rangle$ sia uno stato di X e che $|\psi\rangle$ sia uno stato di Y. Definiremo il prodotto tensoriale $|\varphi\rangle|\psi\rangle$ come **stato prodotto** e, se il sistema composito si trova in tale stato, diremo dunque che i due stati $|\varphi\rangle, |\psi\rangle$ non si influenzano l'un l'altro. Possiamo dimostrare che, effettivamente, uno stato prodotto è ancora uno stato quantistico come segue:

$$\begin{aligned}
\| |\varphi\rangle|\psi\rangle \| &= \sqrt{\sum_{(a,b)\in\Sigma\times\Lambda} |\langle ab|\varphi\otimes\psi\rangle|^2} = \\
&= \sqrt{\sum_{a\in\Sigma}\sum_{b\in\Lambda} |\langle a|\varphi\rangle|^2 |\langle b|\varphi\rangle|^2} = \\
&= \sqrt{\left(\sum_{a\in\Sigma}|\langle a|\varphi\rangle|^2\right)\left(\sum_{b\in\Lambda}|\langle b|\varphi\rangle|^2\right)} = \\
&= \| |\varphi\rangle \| \, \| |\psi\rangle \| = 1
\end{aligned}$$

Abbiamo dunque dimostrato che *la norma euclidea è moltiplicativa rispetto al prodotto tensoriale*, e dunque possiamo facilmente generalizzare a $X_1, \dots, X_n$ sistemi, ottenenendo uno stato prodotto della forma $|\psi_1\rangle \cdots |\psi_n\rangle$ nel sistema composito $(X_1, \dots, X_n)$.

Non tutti gli stati quantistici di un sistema composito sono però degli stati prodotto. Ad esempio,

$$\frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |11\rangle$$

non è uno stato prodotto. Infatti, se lo fosse, esisterebbero due stati $|\varphi\rangle, |\psi\rangle$ tali che

$$|\varphi\rangle \otimes |\psi\rangle = \frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |11\rangle$$

da cui otterremmo che

$$\langle 0|\varphi\rangle \langle 1|\psi\rangle = \langle 01|\varphi \otimes \psi\rangle = 0$$

ossia $\langle 0|\varphi\rangle = 0$ oppure $\langle 1|\psi\rangle = 0$. Questa conclusione è però un assurdo, dato che

$$\langle 0|\varphi\rangle \langle 0|\psi\rangle = \langle 00|\varphi \otimes \psi\rangle = \frac{1}{\sqrt{2}} \quad \text{e} \quad \langle 1|\varphi\rangle \langle 1|\psi\rangle = \langle 11|\varphi \otimes \psi\rangle = \frac{1}{\sqrt{2}}$$

Definiremo gli stati che non sono stati prodotto come **stati entangled** (o in **entanglement**) e sono quegli stati caratterizzati da una correlazione tra i sottosistemi su cui è definito.

In uno spazio composito con $n = 2$, gli stati in entanglement più famosi sono i cosiddetti **stati di Bell**, chiamati così in onore di John Bell, e sono definiti come

$$\begin{aligned}
|\Phi^+\rangle &:= \frac{1}{\sqrt{2}} |00\rangle + \frac{1}{\sqrt{2}} |11\rangle \\
|\Phi^-\rangle &:= \frac{1}{\sqrt{2}} |00\rangle - \frac{1}{\sqrt{2}} |11\rangle \\
|\Psi^+\rangle &:= \frac{1}{\sqrt{2}} |01\rangle + \frac{1}{\sqrt{2}} |10\rangle \\
|\Psi^-\rangle &:= \frac{1}{\sqrt{2}} |01\rangle - \frac{1}{\sqrt{2}} |10\rangle
\end{aligned}$$

La collezione di questi stati $\mathcal{B}_{Bell} = \{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\}$ forma una base, conosciuta per l'appunto come **base di Bell**, e può essere quindi utilizzata per esprimere qualsiasi altro stato a due qubit come combinazione lineare dei suoi elementi; ad esempio

$$|00\rangle = \frac{1}{\sqrt{2}} |\Phi^+\rangle + \frac{1}{\sqrt{2}} |\Phi^-\rangle$$

Se invece ci spostiamo in uno spazio con 3 qubit, possiamo considerare come stati in entanglement gli stati **GHZ**, chiamato così in onore di Daniel Greenberger, Michael Horne, and Anton Zeilinger, e il cosiddetto stato **W**:

$$\begin{aligned}
|GHZ\rangle &:= \frac{1}{\sqrt{2}} |000\rangle + \frac{1}{\sqrt{2}} |111\rangle \\
|W\rangle &:= \frac{1}{\sqrt{3}} |001\rangle + \frac{1}{\sqrt{3}} |010\rangle + \frac{1}{\sqrt{3}} |100\rangle
\end{aligned}$$

Vedremo più avanti alcune caratterische importanti di tali stati e, in particolare, come essi interagiscano con le misurazioni.

Ma come si misura un sistema multiplo? Analogamente al caso probabilistico, abbiamo due strategie di misurazione diverse: misurare tutto il sistema o solo una parte di esso.

Se decidiamo di effettuare la misurazione su *tutto* il sistema, otterremo una formulazione analoga al caso con un sistema solo: dati n sistemi $X_1, \dots, X_n$ con spazi degli stati classici $\Sigma_1, \dots, \Sigma_n$, possiamo vedere il sistema composito $(X_1, \dots, X_n)$ come un sistema singolo con spazio degli stati classici $\Sigma_1 \times \cdots \times \Sigma_n$. Dunque, se il sistema è in uno stato quantistico $|\psi\rangle$, ogni possibile risultato della misurazione $(a_1 \cdots a_n) \in \Sigma_1 \times \cdots \times \Sigma_n$ avrà probabilità di presentarsi pari a $|\langle a_1 \cdots a_n|\psi\rangle|^2$. Ad esempio, supponiamo di avere due sistemi X, Y con spazio degli stati classici $\Sigma = \{0, 1\}$ e $\Lambda = \{\clubsuit, \blacklozenge, \spadesuit, \heartsuit\}$ rispettivamente e che il sistema composito si trovi nello stato

$$\frac{3}{5}|0\rangle|\heartsuit\rangle - \frac{4i}{5}|1\rangle|\spadesuit\rangle$$

Allora, misurare tale sistema nella sua totalità darà come risultato $(0, \heartsuit)$ con probabilità $9/25$ e $(1, \spadesuit)$ con probabilità $16/25$.

Supponiamo ora di voler misurare un sottoinsieme proprio di sottosistemi. Per fissare le idee, partiremo da un sistema composto da due sottosistemi per poi generalizzare. Siano dunque X, Y due sistemi con spazio degli stati Σ, Λ rispettivamente. In generale, uno stato quantistico di (X, Y) è della forma

$$|\psi\rangle = \sum_{(a,b)\in\Sigma\times\Lambda} \alpha_{ab}|ab\rangle$$

dove $\{\alpha_{ab} : (a, b) \in \Sigma \times \Lambda\}$ è una collezione di numeri complessi tale che

$$\sum_{(a,b)\in\Sigma\times\Lambda} |\alpha_{ab}|^2 = 1$$

così che $|\psi\rangle$ sia effettivamente un vettore unitario. Dal caso precedente, sappiamo che se misurassimo entrambi i sistemi otterremo $(a, b) \in \Sigma \times \Lambda$ con probabilità pari a

$$|\langle ab|\psi\rangle|^2 = |\alpha_{ab}|^2$$

Supponiamo ora di voler misurare solo X. Allora, la probabilità che tale misurazione dia come risultato $a \in \Sigma$ è pari a

$$\sum_{b\in\Lambda} |\langle ab|\psi\rangle|^2 = \sum_{b\in\Lambda} |\alpha_{ab}|^2$$

dato che, ovviamente, non è possibile che la probabilità di ottenere a dipenda in qualche modo dallo stato di Y. Assumendo di aver ottenuto lo stato classico $a \in \Sigma$ dalla nostra misurazione, ci aspettiamo che X "collassi" su a, così da essere perennemente nello stato $|a\rangle$, ma come cambia il sistema Y? Per esprimere l'effetto della misurazione di X su Y, riscriviamo $|\psi\rangle$ come

$$|\psi\rangle = \sum_{a\in\Sigma} |a\rangle \otimes |\varphi_a\rangle$$

dove

$$|\varphi_a\rangle = \sum_{b\in\Lambda} \alpha_{ab}|b\rangle$$

per ogni $a \in \Sigma$. Osserviamo che la probabilità di ottenere a misurando può essere dunque riscritta come

$$\sum_{b\in\Lambda} |\alpha_{ab}|^2 = \| \, |\varphi_a\rangle \, \|^2$$

Ora, avendo ottenuto come risultato della misurazione su X l'output a, il sistema composito si troverà nel nuovo stato

$$|a\rangle \otimes \frac{|\varphi_a\rangle}{\| |\varphi_a\rangle \|}$$

Concettualmente, ciò che stiamo dicendo è che del vettore $|\varphi\rangle$ rimane solo la componente relativa allo stato $|a\rangle$, ossia $|a\rangle \otimes |\varphi_a\rangle$, che però deve essere poi normalizzata tramite la sua norma euclidea per ottenere un vettore valido.
Vediamo un esempio: siano X, Y due qubit con spazio degli stati classici $\Sigma = \Lambda = \{0, 1\}$ e supponiamo che (X, Y) sia nello stato

$$|\psi\rangle = \frac{1}{\sqrt{2}} |00\rangle - \frac{1}{\sqrt{6}} |01\rangle + \frac{i}{\sqrt{6}} |10\rangle + \frac{1}{\sqrt{6}} |11\rangle$$

e riscriviamolo nella formulazione più comoda per analizzare le misurazioni parziali, ossia

$$|\psi\rangle = |0\rangle \otimes \left(\frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{6}} |1\rangle \right) + |1\rangle \otimes \left(\frac{i}{\sqrt{6}} |0\rangle + \frac{1}{\sqrt{6}} |1\rangle \right)$$

Dunque, la probabilità di ottenere 0 misurando X è pari a

$$\left\| \frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{6}} |1\rangle \right\|^2 = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}$$

e lo stato del sistema composito diventa

$$|0\rangle \otimes \frac{\frac{1}{\sqrt{2}} |0\rangle - \frac{1}{\sqrt{6}} |1\rangle}{\sqrt{\frac{2}{3}}} = |0\rangle \otimes \left(\sqrt{\frac{3}{4}} |0\rangle - \frac{1}{2} |1\rangle \right)$$

Analogamente, la probabilità di ottenere 1 è pari a

$$\left\| \frac{i}{\sqrt{6}} |0\rangle + \frac{1}{\sqrt{6}} |1\rangle \right\|^2 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

e lo stato del sistema composito diventa

$$|1\rangle \otimes \frac{\frac{i}{\sqrt{6}} |0\rangle + \frac{1}{\sqrt{6}} |1\rangle}{\sqrt{\frac{1}{3}}} = |1\rangle \otimes \left(\frac{i}{\sqrt{2}} |0\rangle + \frac{1}{\sqrt{2}} |1\rangle \right)$$

Si può ragionare in maniera del tutto simmetrica a quanto visto fin qui se il sistema misurato è Y.

Per generalizzare a più sottosistemi, ci basta suddividere tali sottosistemi in due insiemi, come misurati e non misurati, e ricadere così nel caso di due soli sottosistemi. La caratteristica importante da tenere a mente è che il prodotto tensoriale, proprio come il prodotto cartesiano, **non** è commutativo: dunque, gli stati $|\pi\rangle \otimes |\psi\rangle, |\psi\rangle \otimes |\pi\rangle$ sono diversi tra loro e dunque l'ordine in cui definiamo i sistemi e i nostri stati è di fondamentale importanza.
Riprendiamo gli stati $|GHZ\rangle, |W\rangle$ e vediamo cosa succede nelle misurazioni parziali: iniziamo supponendo di voler misurare solo il primo sistema dei tre che formano il sistema composito che si trova nello stato $|GHZ\rangle$, che ricordiamo essere

$$|GHZ\rangle = \frac{1}{\sqrt{2}} |000\rangle + \frac{1}{\sqrt{2}} |111\rangle$$

Dunque, la probabilità che la misurazione del primo sistema dia 0 è pari a $1/2$, facendo collassare parzialmente il sistema composito nello stato $|000\rangle$, e ugualmente a $1/2$ per l'output 1, facendo collassare parzialmente il sistema composito nello stato $|111\rangle$.
Consideriamo ora invece lo stato $|W\rangle$, che può essere riscritto come

$$\begin{aligned}|W\rangle &= \frac{1}{\sqrt{3}}|001\rangle + \frac{1}{\sqrt{3}}|010\rangle + \frac{1}{\sqrt{3}}|100\rangle \\ &= |0\rangle \otimes \left(\frac{1}{\sqrt{3}}|01\rangle + \frac{1}{\sqrt{3}}|10\rangle\right) + |1\rangle \otimes \left(\frac{1}{\sqrt{3}}|00\rangle\right)\end{aligned}$$

Dunque, la probabilità di ottenere 0 misurando il primo sistema è pari a $2/3$ e il sistema composito collasserebbe nello stato

$$|0\rangle \otimes \frac{\frac{1}{\sqrt{3}}|01\rangle + \frac{1}{\sqrt{3}}|10\rangle}{\sqrt{\frac{2}{3}}} = |0\rangle \otimes \left(\frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle\right) = |0\rangle\,|\Psi^+\rangle$$

mentre la probabilità di ottenere 1 è pari a $1/3$ e il nuovo stato del sistema composito diventerebbe $|100\rangle$.

Ci rimane dunque da capire attraverso quali operazioni è possibile far evolvere tali sistemi multipli. In linea generale, si tratta comunque di matrici unitarie quadrate (ovviamente con dimensioni maggiori) che hanno un numero di righe e colonne pari al numero di elementi in $\Sigma_1 \times \cdots \times \Sigma_n$, che nel nostro caso sarà pari a $N = 2^n$, dato che i sottosistemi saranno dei singoli qubit.
La particolarità delle operazioni effettuate su sistemi multipli è che, se le operazioni agiscono solo su un sottoinsieme di sistemi singoli indipendentemente dagli altri, allora l'operatore totale può essere scomposto nel prodotto tensoriale dei singoli operatori indipendenti. Più precisamente, se supponiamo di avere n sottosistemi $X_1, \ldots, X_n$ e $\mathbf{U}_1, \ldots, \mathbf{U}_n$ operatori unitari su questi sottosistemi, allora l'operatore $\mathbf{U}$ che agisce sul sistema composito $(X_1, \ldots, X_n)$ sarà

$$\mathbf{U} = \mathbf{U}_1 \otimes \ldots \otimes \mathbf{U}_n$$

ottenendo una situazione analoga alla controparte classica. Verifichiamo che il prodotto tensoriale di matrici unitarie sia unitario:

$$\begin{aligned}\mathbf{U}^\dagger\mathbf{U} &= (\mathbf{U}_1 \otimes \ldots \otimes \mathbf{U}_n)^\dagger(\mathbf{U}_1 \otimes \ldots \otimes \mathbf{U}_n) \overset{(1)}{=} \\ &= (\mathbf{U}_1^\dagger \otimes \ldots \otimes \mathbf{U}_n^\dagger)(\mathbf{U}_1 \otimes \ldots \otimes \mathbf{U}_n) \overset{(2)}{=} \\ &= (\mathbf{U}_1^\dagger\mathbf{U}_1) \otimes \cdots \otimes (\mathbf{U}_n^\dagger\mathbf{U}_n) = \\ &= \mathbb{I} \otimes \cdots \otimes \mathbb{I} = \mathbb{I}^{\otimes n}\end{aligned}$$

dove (1) segue dal fatto che

$$(\mathbf{U}_1 \otimes \ldots \otimes \mathbf{U}_n)^\dagger = \mathbf{U}_1^\dagger \otimes \ldots \otimes \mathbf{U}_n^\dagger$$

proprietà facilmente verificabile tramite definizioni di prodotto tensoriale e matrice autoaggiunta, mentre (2) segue dalla moltiplicatività del prodotto tensoriale. Sfruttiamo inoltre l'occasione per introdurre la notazione $\mathbb{I}^{\otimes n}$ per indicare il prodotto tensoriale di n matrici identità a due dimensioni, identificando dunque la matrice identità $2^n \times 2^n$. Dunque $\mathbf{U}$ è unitaria e rappresenta un'operatore valido da applicare sul sistema composito.

Una delle casistiche più comuni è quella di voler applicare un qualche operatore $\tilde{\mathbf{U}}$ solo su parte dei sottosistemi: espliciteremo questo "non far nulla" sugli altri sistemi tramite matrice identità (della relativa dimensione), usandola come fattore nel prodotto tensoriale con $\tilde{\mathbf{U}}$. Prendiamo ad esempio due qubit X, Y e supponiamo di voler applicare un Hadamard ad X e di non voler far agire alcun operatore su Y: allora, l'operatore sul sistema composito (X, Y) sarà

$$\mathbf{H} \otimes \mathbb{I} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & -1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Se invece volessimo lasciare intatto X e applicare Hadamard a Y dovremmo invertire l'ordine:

$$\mathbb{I} \otimes \mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \mathbf{H} & 0 \\ 0 & \mathbf{H} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

Ovviamente non tutti gli operatori unitari di dimensione $N \times N$ sono scomponibili in un prodotto tensoriale di operatori unitari indipendenti di dimensione minore, com'è possibile osservare analizzando il gate **SWAP** e i **controlled gate**.

Iniziando dallo **SWAP**, definiamo tale operatore come agente su due sistemi X, Y il cui effetto è, per l'appunto, quello di scambiare gli stati della coppia (X, Y). Matematicamente, per ogni scelta degli stati $a, b \in \Sigma = \Lambda$, si vuole che

$$\mathbf{SWAP} \, |a\rangle \, |b\rangle = |b\rangle \, |a\rangle$$

A tal fine, è possibile scrivere la matrice associata all'operatore di scambio come segue usando la notazione di Dirac:

$$\mathbf{SWAP} \coloneqq \sum_{c,d \in \Sigma} |c\rangle \langle d| \otimes |d\rangle \langle c|$$

o, tradotto in forma matriciale, come

$$\mathbf{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Osserviamo che effettivamente tale matrice inverte i due qubit di uno stato, lasciando "invariati" gli stati in cui i due qubit hanno lo stesso stato:

$$|01\rangle \to |10\rangle \qquad |10\rangle \to |01\rangle \qquad |00\rangle \to |00\rangle \qquad |11\rangle \to |11\rangle$$

Passiamo ora ai controlled gate, supponiamo di avere due sistemi: un qubit Q e un qualche sistema arbitrario R. Data un'operazione unitaria $\mathbf{U}$ agente su R, definiamo come **controlled U** l'operatore unitario sul sistema (Q, R) definito come

$$\mathbf{CU} \coloneqq |0\rangle \langle 0| \otimes \mathbb{I}_R + |1\rangle \langle 1| \otimes \mathbf{U} = \begin{bmatrix} \mathbb{I}_R & 0 \\ 0 & \mathbf{U} \end{bmatrix}$$

dove con $\mathbb{I}_R$ indichiamo l'identità applicata al sistema R. Dunque Q sarà il nostro **qubit di controllo** e R sarà il nostro **target**. Ad esempio, se R è anch'esso formato da un solo qubit e supponiamo di voler applicare un *bit flip*, avremmo un **CNOT**:

$$\mathbf{CNOT} = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes \mathbf{NOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

mentre se invece vogliamo applicare un *phase flip* avremmo un **CZ**:

$$\mathbf{CZ} = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes \mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Se invece R è un sistema formato da due qubit su cui vogliamo applicare uno **SWAP**, otterremo

$$\mathbf{CSWAP} := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

il cui effetto può essere riassunto in questo schema:

$$|0bc\rangle \overset{\mathbf{CSWAP}}{\longrightarrow} |0bc\rangle \qquad |1bc\rangle \overset{\mathbf{CSWAP}}{\longrightarrow} |1cb\rangle$$

per ogni $c, b \in \Sigma = \{0, 1\}$.
Infine, l'ultimo esempio di questa sezione è il **controlled-controlled-NOT**, denotato come **CCNOT** e anche conosciuto come **Toffoli gate**. Tale operatore agisce su sistemi a tre qubit, usandone due come controllo e uno come target, è definito come

$$\mathbf{CCNOT} \equiv \mathbf{TOF} := (|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|) \otimes \mathbb{I} + |11\rangle\langle 11| \otimes \mathbf{NOT}$$

$$\equiv \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Concludiamo questa lezione introducendo un ultimo operatore unitario molto importante nel quantum computing e che nel corso delle prossime lezioni utilizzeremo molto: la

Quantum Fourier Transform, o più semplicemente **QFT**, la cui forma matriciale è

$$\mathbf{QFT} \coloneqq \frac{1}{\sqrt{N}} \sum_{j,k=0}^{2^n-1} \omega^{jk} \left|k\right\rangle \left\langle j\right| \equiv \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^8 & \dots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \dots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{bmatrix}$$

dove $N = 2^n$ e ω è la radice $N-$esima dell'unità, ossia $\omega = e^{2i\pi/N}$. L'utilizzo principale di questo operatore è quello di creare una sovrapposizione equiprobabile di tutti gli stati quantistici nel sistema multiplo, analogamente a quanto l'Hadamard gate faceva per sistemi a un qubit. Per oggi non andremo oltre la semplice definizione, ma nelle prossime lezioni daremo maggiori spiegazioni sul suo utilizzo e sulla sua validità, specie nel contesto dell'Hidden Subgroup Problem, della Quantum Phase Estimation e dell'Algoritmo di Shor.