Alma Mater Studiorum · Università di Bologna

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

L'ARTE DELLA SEGRETEZZA, tra mondo classico e quantistico.

Relatore:
Prof. Elisa Ercolessi

Presentata da:
Alessandro Bagagli

Anno Accademico 2015/2016

Abstract

La tesi si pone l'obiettivo di descrivere l'evoluzione temporale della crittografia, prima dal punto di vista classico, partendo dalla più antica tecnica a chiave privata fino alla più recente a chiave pubblica, e poi, soprattutto, da quello quantistico, passando per i postulati e l'introduzione del qubit. Questo rappresenta un sistema a due stati che corrisponde all'analogo quantistico del bit classico: se ne illustrano le proprietà e una possibile rappresentazione negli stati di polarizzazione del fotone, andando poi a descrivere il fenomeno dell'Entanglement, un particolare stato legato puramente quantistico in cui le parti del sistema si influenzano a distanza. Infine si focalizza sulla Quantum Key Distribution, illustrando tre protocolli che ne mostrano il funzionamento: questa consiste in una tecnica tutta quantistica per la creazione, praticamente a prova di intercettazione, di una chiave privata tra una coppia di utenti.

Indice

Introduzione

La realizzazione di messaggi segreti è antica forse quanto la comunicazione stessa, e si basa sull'offuscamento di un messaggio che si vuole comunicare da un mittente ad un destinatario senza che terzi non autorizzati possano interpretarlo correttamente.
Nel corso della storia sono state affinate tecniche crittografiche le più sofisticate e indecifrabili possibile, mettendo sempre al centro della disciplina tre entità: un Mittente che vuole inviare un messaggio segreto, un Destinatario che segretamente vuole riceverlo e un Crittoanalista che, a seconda dei casi per scopi corretti o malvagi, vuole origliare.
La storia ha sempre visto crittografi e crittoanalisti ricorrersi l'un l'altro: i primi per creare cifrari che i secondi non potessero risolvere e i secondi per vanificare il lavoro dei primi. Questo studio quindi si pone l'obiettivo di descrivere i sistemi crittografici usati classicamente: dal più antico metodo a Private Key, ovvero con una chiave segreta condivisa tra gli utenti (come il cifrario di Vernam), fino a quello più recente a Public Key, un sistema asimmetrico in cui tutti possono criptare un messaggio facilmente, ma solo uno, il destinatario, può in tempi brevi risalire al messaggio originale.
Questo ultimo metodo è basato sull'utilizzo di alcune funzioni che si suppone siano difficili da invertire per un computer (come ad esempio il problema della fattorizzazione): proprio l'assunzione che esistano queste funzioni cade in seguito allo sviluppo di calcolatori che sfruttano i principi della meccanica quantistica, su questi dispositivi infatti, disponendo di una potenza di calcolo enormemente più elevata di quelli classici, è possibile realizzare algoritmi che risolvano in tempi decenti problemi che classicamente sono ritenuti computazionalmente difficili (ovvero la cui risoluzione, per dati in input molto grandi, richiede tempi fuori dalla portata dell'uomo). Queste sono quindi le tematiche affrontate nella prima parte.
La seconda invece è orientata alla descrizione dell'informatica quantistica, partendo da un excursus di meccanica quantistica, passando per il concetto di qubit, una versione quantomeccanica del bit classico, fino alla descrizione del potente strumento dato dall'entanglement, un tipo di interazione che non possiede controparte classica attraverso cui è possibile realizzare il fenomeno del teletrasporto quantistico. Infine si dà una descrizione qualitativa del funzionamento dei laboriosi algoritmi che sfruttano la trasformata di Fourier quantistica per risolvere il problema della fattorizzazione. L'ultima parte è dedicata completamente alle applicazioni alla crittografia della meccanica quantistica: questa in-

fatti oltre a far cadere i capisaldi della crittografia classica a chiave pubblica, fornisce tecniche naturali per lo scambio di chiavi segrete da implementare in sistemi consolidati a private key. Queste sono particolarmente utili perchè riferendoci alla peculiarità quantistica sull'impossibilità di effettuare una misura su uno stato senza disturbarlo, una eventuale intercettazione non può avvenire senza una contemporanea modifica quantificabile dei qubit della chiave. Questo fatto quindi sembra porre fine all'infinita lotta tra crittografi e crittoanalisti, vedendo questi ultimi definitivamente sconfitti, e fornisce uno strumento solido per rendere sicure le conversazioni del futuro.

È interessante vedere come la National Security Agency, che sostanzialmente detta lo standard mondiale nel campo della sicurezza, ha manifestato la propria preoccupazione riguardo alla possibilità che quelle tecniche di cifratura, che per anni aveva ritenuto impenetrabili, possano essere superate con l'uso di computer quantistici. Il problema di fondo risiede nel fatto che nessuno sa come realizzare un tipo di cifratura che sia rigorosamente e formalmente a prova di computer quantistici: l'unica raccomandazione che può fare la NSA è di utilizzare alcuni algoritmi ritenuti sicuri dall'attacco di un computer quantistico potente, ma nulla di più.[1]

[1]Per maggiori dettagli si veda [6].

Capitolo 1

Crittografia Classica

La Crittologia si suddivide in due parti costituenti: la crittografia che prevede diversi metodi di cifratura e la crittoanalisi che comprende le varie tecniche di interpretazione di un messaggio criptato. Per introdurre un po' di notazione chiamiamo Mitt il mittente del messaggio, Dest il destinatario e X il crittoanalista, inoltre denoteremo come MSG lo spazio dei messaggi e CRITTO lo spazio dei crittogrammi[1].
Con questa notazione la crittologia si basa su una coppia di funzioni C e D tali che:

$$\text{C: MSG} \longmapsto \text{CRITTO}$$
$$\text{D: CRITTO} \longmapsto \text{MSG}$$

in questo modo preso un m∈MSG e un c∈CRITTO, si ha che C(m)=c e D(c)=m, con la condizione che, perché m decifrata da Dest attraverso D non ammetta ambiguità, occorre che C sia iniettiva, per cui se C(m)=C(m') inevitabilmente m e m' devono essere uguali. Il crittoanalista X può intervenire in maniera passiva o attiva: ovvero ascoltando c, e cercando di risalire a m con i mezzi a sua disposizione, o interferendo con c in modo da disturbare la comunicazione e non fare interpretare correttamente a Dest il messaggio m di Mitt.

Nel tredicesimo secolo Ruggero Bacone elencò quelli che erano per l'epoca secondo lui i tre principi che caratterizzano la robustezza di un cifrario:
i) C e D devono essere facili da calcolare;
ii) deve essere impossibile ricavare D se C non è nota;
iii) c=C(m) deve apparire un messaggio innocente.

Ora per quello che riguarda la crittografia ad oggi, visto che i messaggi di cui si parla sono generalmente sequenze di bit, tutti a primo avviso sono innocenti; inoltre quello che era facile da calcolare, a mano, per Bacone, adesso si traduce in funzioni computazionalmente facili da calcolare, e al posto di impossibile ci si accontenta di *computazionalmente*

[1] Il testo di riferimento per questo capitolo è [3]

difficile. Quest'ultimo punto diventerà più avanti centrale nella differenza tra le tecniche passate, considerate impenetrabili e quindi computazionalmente difficili da calcolare, a quelle più attuali che disponendo di maggiore potenza di calcolo, rendono possibile ciò che prima non lo era.
Citando le parole di Ronald Rivest[2], uno dei massimi crittografi contemporanei, finché non conosciamo alcuni fatti fondamentali sulla difficoltà computazionale, non possediamo ancora gli strumenti per provare senza ombra di dubbio che i sistemi di cifratura siano computazionalmente sicuri.
Il massimo che si può fare, da un punto di vista teorico, è di provare che un sistema è sicuro basandosi sull'assunzione che per un computer sia difficile risolvere un certo tipo di ben definiti problemi matematici, sapendo bene che tali assunzioni potrebbero essere false.

A seconda del livello di segretezza che si vuole mantenere per un certo canale, si utilizzano funzioni C e D diverse: queste dovranno essere sicure in ogni aspetto per garantire la riservatezza per comunicazioni diplomatiche e militari, mentre risulta pressoché impossibile che ciò avvenga per un utilizzo esteso alla crittografia di massa.
Questa parte infatti si concentra su un altro tipo di funzioni C e D, pubblicamente note, ma il cui funzionamento dipende strettamente da una chiave k segreta e nota solo alla coppia Mitt e Dest, anche detto metodo a Private Key.

1.1 Private Key

Per questo genere di cifratura ci si rifà a funzioni C e D leggermente diverse: introducendo con KEY lo spazio di tutte le chiavi possibili, C e D vengono definite come

$$\text{C: MSG}\times\text{KEY}\longmapsto\text{CRITTO}$$
$$\text{D: CRITTO}\times\text{KEY}\longmapsto\text{MSG}$$

in modo che, con k$\in$KEY, C(m,k)=c e D(c,k)=m dove k una chiave tendenzialmente più corta dei messaggi da inviare tra Mitt e Dest che può essere scambiata una tantum, attraverso un canale sicuro anche se costoso.
Inoltre la comodità di questo metodo sta anche nel fatto che in seguito ad una eventuale scoperta da parte di X di k, per la coppia Mitt-Dest è sufficiente introdurre una nuova k'$\neq$k per riaprire un canale di comunicazione sicuro. Al contrario per una coppia di funzioni C-D non a private key, una volta manomesso il sistema di cifratura è necessario cambiare completamente queste funzioni.
Visto che la sicurezza del canale dipende solo da k, lo spazio KEY deve essere scelto opportunamente, in modo tale che per X la verifica di tutte le D(c,k) per un dato c

[2]per maggiori dettagli vedi [5]

intercettato sia complicato.
Un aspetto importante da considerare il fatto che all'interno di un gruppo di n utenti che vogliano scambiarsi informazioni private a coppie sono necessarie $\frac{n(n-1)}{2}$ chiavi segrete, per cui cresce quadraticamente nel numero di utenti.
Inoltre questo sistema di cifratura è detto simmetrico perché è totalmente speculare nell'azione di Mitt e Dest.
Di seguito si dà un esempio di sistema a private key semplice ma altamente efficace, il cifrario di Vernam, e si analizza più nel dettaglio questo genere di metodo di cifratura.

1.1.1 Cifrario di Vernam

Il cifrario di Vernam è, nato nel 1917, è l'unico sistema crittografico classico la cui sicurezza è verificabile matematicamente, e perciò si è guadagnato il titolo di cifrario perfetto: questo significa che dato un crittogramma c relativo ad un messaggio m codificato con la chiave k, la probabilità che c corrisponda ad m è la stessa che corrisponda ad un qualsiasi m'$\neq$m di lunghezza al massimo k.
Diamo di seguito un esempio: Paola e Chiara partono ognuna con una password segreta identica composta di n-bit. Paola codifica il suo messaggio di n-bit aggiungendo in senso binario a questo la chiave, Chiara, ricevuto il crittogramma, lo decodifica semplicemente sottraendogli la stessa chiave segreta.
Questo metodo, finchè i bit di k sono tenuti segreti, è certamente sicuro, a patto però che si utilizzi una chiave lunga almeno quanto il messaggio da inviare e che sia diversa ogni volta: per questo viene detto anche cifrario *one time pad* o *blocco monouso*. Se infatti si utilizza la stessa chiave più di una volta il sistema passa dall'essere inattaccabile a facilmente attaccabile attraverso il metodo Kasiski, basato sulla ricerca di sequenze di caratteri identiche poste fra loro ad una certa distanza, questa lunghezza con una certa probabilità potrebbe corrispondere alla dimensione della chiave stessa o a un suo multiplo.
Il problema principale dei sistemi a private key è la distribuzione della chiave: in particolare per il cifrario di Vernam se il messaggio è molto lungo, questo richiede una chiave altrettanto lunga, il che rende poco pratico l'utilizzo. Inoltre questa è indispensabile che sia scambiata prima, custodita al sicuro finchè non usata e successivamente distrutta: altrimenti tale chiave potrebbe essere copiata senza disturbare l'originale, e quindi compromettendo la sicurezza dell'intero protocollo.
Sebbene il protocollo sia laborioso ed economicamente svantaggioso il cifrario di Vernam è tutt'ora utilizzato per la sua provata impenetrabilità.

1.2 Complessità Computazionale

È giunto il momento di specificare che cosa si intende per facile o difficile per un computer.
L'efficienza di un algoritmo si misura in base alle risorse disponibili: per un calcolatore queste sono memoria e soprattutto il tempo T(n) necessario a eseguire un calcolo, dove n il parametro che dà la dimensione dell'ingresso, o per meglio dire il suo ordine di grandezza. Infatti uno stesso input ha una dimensione diversa a seconda della base computazionale in cui è espresso, tuttavia data la proprietà del cambio di base dei logaritmi per cui rappresentazioni diverse di uno stesso dato sono proporzionali tra loro

$$\log_a x = log_a b \times log_b x$$

ciò che interessa è l'ordine di grandezza del suo andamento.
Introduciamo quindi alcune funzioni che aiutano a comprendere il significato di ordine di grandezza: si definisce O(g(n))) l'insieme di tutte le funzioni che, da un certo punto in poi, sono superiormente limitate da cg(n) con c costante positiva, si definisce invece Θ(g(n)), l'insieme delle funzioni f(n) tali che, da un certo punto in poi:

$$c_1 g(n) \leq f(n) \leq g(n) c_2$$

con c_1 e c_2 costanti positive.
Un algoritmo si dice quindi di complessità polinomiale se T(n)$\in\Theta$(p(n)), con p(n) polinomio arbitrario al più di grado n, si dirà invece di complessità esponenziale se non esiste alcun p(n) tale che T(n)$\in$O(p(n)).
Un algoritmo di questo tipo è utilizzabile solo per piccoli valori di n: per dare un esempio di questo andamento si prendano come esempio tre algoritmi A_1, A_2 e A_3 che risolvono lo stesso problema su n dati con rispettivamente T(n) dell'ordine di n, n^2 e 2^n. Per n=50 i calcoli sono portati a termine rispettivamente in $5 \cdot 10^{-8}$ secondi, $2{,}5 \cdot 10^{-6}$ secondi e trenta milioni di anni, a prova del fatto che un algoritmo esponenziale può essere utilizzato solo per piccolissimi valori di n.
Un buon programma quindi deve evitare il più possibile questo genere di andamento, tuttavia non sempre si riesce perché l'esponenzialità è legata a strutture combinatorie di base come le 2^n configurazioni possibili di n bit, o le n! possibili permutazioni di n elementi.
Ad esempio il problema della primalità, molto caro alla crittografia, consiste nel determinare se un dato numero N, rappresentato da n cifre binarie, è primo oppure no. Se N non è primo uno dei suoi divisori deve essere minore o uguale alla sua radice quadrata, per cui si testano tutti i numeri interi tra 2 e $\sqrt{N}$, ovvero rappresentati da sequenze binarie di lunghezza massima $\frac{n}{2}$.
Per generare tutte le possibili configurazioni di $\frac{n}{2}$ bit occorre un tempo $O(2^{\frac{n}{2}})$ composto con D(N) costo polinomiale della divisione. Per cui l'algoritmo è inteso computazionalmente difficile dal momento che il tempo per il suo calcolo è di ordine esponenziale.

1.3 Public Key

Introdotto nel 1976 da Diffie ed Hellman, e indipendentemente da Merkle, il metodo di cifratura a Public Key si contrappone a quello a chiave segreta: questo infatti si basa su un sistema in cui chiunque può cifrare un messaggio ma solo un lo può decifrare.
Infatti questo metodo prevede la presenza di due chiavi, di cui una nota pubblicamente k_{pub} e una privata k_{priv} nota solo al destinatario, per questo motivo è anche detto sistema di cifratura asimmetrico. Si noti che a differenza del metodo a Private Key, questo sistema prevede, per un insieme di n utenti, la necessità di creare solo n coppie di k_{priv}-k_{pub}, fatto da non sottovalutare nell'ottica di una globalizzazione delle comunicazioni che rende praticamente impossibile gestire lo scambio segreto di un numero quadratico di chiavi.
Per realizzare questa configurazione è risulta che C deve essere di tipo one-way trap-door, ossia che c=C(m,k_{pub}) sia semplice da calcolare mentre m=D(c,k_{priv}) sia difficile (da cui one-way), a meno che non si conosca un meccanismo segreto (trap-door) che la semplifica.
Si basano infatti su funzioni facilmente calcolabili ma la cui inversione legata alla risoluzione di problemi ritenuti difficilissimi, fino a prova contraria, a meno che non si conosca il ruolo di k_{priv}.
Nella vita di tutti i giorni abbiamo a che fare con l'esistenza di questo genere di funzioni, basti pensare ad esempio un semplice sistema di protezione domestica della posta: prendendo delle lettere come messaggi da cifrare, identificando con k_{pub} come l'indirizzo di casa e con k_{priv} la chiave della buchetta delle lettere, è immediato il fatto che reso noto l'indirizzo è facile recapitare un messaggio, tuttavia è difficile, per chiunque non possieda la chiave scoprire quali messaggi siano stati recapitati.
L'esistenza in natura di questo genere di funzioni fa supporre che queste esistano anche in matematica, fatto non scontato a priori, anche se è indispensabile accontentarsi di un problema noto per cui non si conosce un algoritmo polinomiale o comunque di funzioni che siano computazionalmente difficili da invertire, fino al momento improbabile in cui si trovi una soluzione efficiente a questi problemi.
Il metodo a chiave pubblica tuttavia mostra alcuni svantaggi immediati:
i) Il fatto che chiunque possa cifrare un messaggio espone il sistema ad attacchi del tipo chosen plain-text: infatti nulla vieta al crittoanalista X di cifrare n messaggi utilizzando la funzione pubblica C e la corrispondente chiave pubblica k_{pub}, ottenendo così n crittogrammi da confrontare con i messaggi in viaggio verso Dest. In questo modo egli può ricavare informazione sia se un crittogramma intercettato c_i combacia con uno degli n in suo possesso, sia se questo non succede sapendo che il messaggio relativo a c_i non corrisponde a nessuno degli n messaggi da lui cifrati. Questo attacco è particolarmente pericoloso se X sospetta che Dest debba ricevere un messaggio particolare, oppure se c_i rappresenta un messaggio breve e di struttura prevedibile come un indirizzo Internet o una password.

ii) Inoltre questi sistemi sono molto pi lenti di quelli basati su cifrari simmetrici, il rapporto delle velocit è stimato nell'ordine del millesimo nelle realizzazioni hardware e un centesimo in quelle software.
Per sfruttare i pregi di entrambi i sistemi e limitarne i difetti, spesso si usa un approccio ibrido che combina gli aspetti positivi ed esclude quelli negativi: infatti spesso si usa il metodo pubblico per scambiare una chiave da utilizzare nel protocollo privato, senza incontri fisici tra gli utenti. Inoltre l'attacco chosen plain-text diventa inefficace se la password da scambiare attraverso il canale pubblico è scelta in modo da essere imprevedibile al crittoanalista.
In questo studio prenderemo in esame due problemi in particolare: quello del calcolo del logaritmo discreto e della fattorizzazione; focalizzandoci poi su quest'ultimo, essendo alla base del cifrario RSA che descriveremo.

Calcolo del logaritmo discreto

Questo problema richiede, noti x, y e s interi, di trovare il valore di z tale che $y=x^z \bmod s$.
Nell'algebra non modulare il problema è semplice e corrisponde a calcolare il logaritmo di y in base x, ma con il modulo il problema diventa computazionalmente difficile e non ammette sempre soluzione.
Per un teorema dell'algebra modulare, se s è primo esiste una soluzione per ogni y, se e solo se x è un generatore dell'insieme dei numeri primi con s, quindi gli algoritmi ad oggi noti hanno la stessa complessità della fattorizzazione anche se non è stato dimostrato che i due problemi siano computazionalmente equivalenti.
La dimostrazione di come introdurre una trap-door nel logaritmo discreto è piuttosto complessa, tuttavia basti sapere che ciò è possibile.

Fattorizzazione

Dati due fattori interi p e q, calcolare n è un problema polinomialmente facile perché richiede tempo quadratico nella lunghezza della loro rappresentazione. Tuttavia invertire la funzione significa ricostruire p e q, sapendo solo n: questo possibile farlo univocamente solo se p e q sono primi.
Per tale inversione è improbabile l'esistenza di un algoritmo polinomiale, infatti è legato ad un problema difficile, esponenziale, ma diventa banale se si conosce k_{priv}, ovvero uno dei due fattori.

1.4 Il cifrario RSA

Nato nel 1978, prende il nome dalle iniziali degli autori Rivest, Shamir e Adleman. La sua fortuna è stata enorme perché combina una grande semplicità strutturale a una

sostanziale inviolabilità, legata ad alcuni risultati di algebra modulare. Se ne illustra di seguito l'organizzazione generale.

1.4.1 Creazione della chiave

Ogni utente, come possibile Dest, esegue le seguenti operazioni:
i) sceglie due numeri primi p, q molto grandi;
ii) ne calcola il prodotto n e la funzione di Eulero $\Phi(n)$[3];
iii) sceglie un intero x minore di $\Phi(n)$ e primo con esso;
iv) calcola l'intero d inverso di x modulo $\Phi(n)$, che è unico perché x e $\Phi(n)$ sono primi fra loro;
v) rende pubblica la chiave k_{pub}= (x,n) e mantiene segreta la chiave k_{priv} = (d).

1.4.2 Messaggio

Il messaggio è codificato come sequenza binaria, che viene trattata come un numero intero m. Per poter utilizzare il cifrario occorre che m sia strettamente minore di n, ciò è possibile dividendo il messaggio in blocchi di al più $\log_2 n$ bit. La dimensione massima del blocco dipende dalla chiave pubblica del destinatario, che è fissata per tutti i destinatari (cifrario a blocchi). Infatti se m fosse maggiore di n allora i messaggi m e m mod n genererebbero lo stesso crittogramma. Poiché però il valore di n è relativo ad un singolo Dest, occorre uno standard per la decomposizione in blocchi nell'intero sistema, ovvero stabilire un limite inferiore per la dimensione dei blocchi uguale per tutti.

1.4.3 Cifratura e decifrazione

Un utente, per inviare a Dest un messaggio m, calcola la funzione

$$c=C(m,k_{pub})=m^x \bmod n$$

dove x e n sono contenuti nella k_{pub}. Dest, una volta ricevuto c, deve solamente calcolare

$$m=D(c,k_{priv})=c^d \bmod n$$

dove d contenuto nella sua k_{priv}.

La correttezza del cifrario, ovvero la sua invertibilità e la iniettività di C, si dimostra attraverso dei teoremi dell'algebra modulare[4], che garantiscono il fatto che dati due crittogrammi c_1 e c_2 diversi, essi corrispondono a due messaggi m_1 e m_2 diversi.
È opportuno però fare alcune considerazioni: data la potenza di calcolo di un calcolatore

[3]Per n strettamente maggiore di 1 si definisce la funzione di Eulero $\Phi(n)$ come il numero di interi minori di n e relativi primi con esso.

[4]per approfondimenti si veda [4]

classico, p e q si intendono grandi se contengono almeno un migliaio di cifre binarie, per cui la fattorizzazione di n è praticamente inattuabile e la probabilità che due utenti scelgano la stessa k_{priv} è trascurabile.
Le chiavi possono essere create in maniera efficiente, dal momento che tutte le operazioni che coinvolgono la loro generazione, ovvero scelta di p e q primi, moltiplicazione e l'inverso di x mod Φ(n), sono di ordine polinomiale.

1.4.4 Attacchi all'RSA

Se il crittoanalista X non è in grado di prevedere il contenuto dei messaggi che vengono scambiati, attacchi del tipo chosen plain-text sono sostanzialmente inoffensivi.
Inoltre la sicurezza del cifrario è legata alla difficoltà di fattorizzare un numero intero molto grande, infatti noto n dalla chiave pubblica, determinare p e q divisori di n, equivale a trovare la chiave privata relativa a Dest.
Tuttavia non è ancora dimostrato che valga l'implicazione inversa, ovvero che per forzare il cifrario RSA si debba passare per forza per la fattorizzazione, questo infatti è stato solo congetturato da Rivest, Shamir e Adleman, anche se risulta essere molto plausibile.
Alcune varianti dell'RSA sono dimostrate essere equivalenti al problema della fattorizzazione, ma tali sono fondamentalmente più complesse e non hanno riscosso lo stesso successo.

Nella seconda parte della trattazione vedremo come sostanzialmente cada l'infallibilità di questo metodo con l'applicazione della meccanica quantistici all'informatica.
Questa nuova branca della computazione infatti rivoluziona in particolare il tema della complessità e provvede a fornire automaticamente nuove tecniche crittografiche che emergono naturalmente.

Capitolo 2

Meccanica quantistica

Per lo studio della meccanica quantistica è necessaria una conoscenza di base dell'algebra lineare: infatti il concetto di sovrapposizione di stati porta con sé la necessità di reinterpretare lo stato di un sistema come elemento di un insieme che sia chiuso rispetto all'operazione di somma, in questo modo assumerà un significato, magari non chiaro, ma di sicuro coerente una loro combinazione lineare.[1]
Un riassunto dei risultati principali di Algebra Lineare che torneranno utili in seguito è riportato in Appendice. Invece in questo capitolo elencheremo i quattro postulati della meccanica quantistica, traducendo in linguaggio fisico quelle che sono le proprietà che emergono da un'analisi puramente algebrica. Introdurremo poi il concetto di operatore densità, di importanza vitale per la descrizione di sistemi interagenti, e quello di qubit, con l'esempio pratico dato dalla polarizzazione dei fotoni. Mostreremo infine le proprietà dell'entanglement, la realizzazione del teletrasporto quantistico, e per finire qualche cenno sugli algoritmi quantistici utili alla crittografia.

2.1 Postulati

2.1.1 Stato

Si definisce stato un vettore $|\psi\rangle$ in uno spazio degli stati complesso dotato di prodotto interno (uno spazio di Hilbert), che descrive completamente un sistema fisico: in particolare considereremo equivalenti due stati che possiedono le stesse informazioni dinamiche, ovvero che differiscano al più per quantità non osservabili.
Pur riassumendo tutte le proprietà fisicamente interessanti di un sistema, il vettore d'onda $|\psi\rangle$ non rappresenta nessuna quantità direttamente misurabile, il ruolo che essa gioca all'interno della descrizione fisica di un fenomeno non è infatti immediatamente ricono-

[1] Per la parte di Meccanica Quantistica si è preso come testo di riferimento [2], mentre per la sezione informatica e crittografica è stata seguita la trattazione di [4].

scibile, tuttavia una chiara interpretazione la possiede la sua norma al quadrato $\langle\psi|\psi\rangle$: essa infatti è un numero positivo e, una volta normalizzata $|\psi\rangle$, rappresenta una probabilità. Inoltre la proprietà di sovrapposizione degli stati trova la sua forma vettoriale nell'operazione di somma tra gli elementi dello spazio di Hilbert, a meno di fattori di normalizzazione: infatti dato un set, assumiamo finito, di n autostati di base normalizzati e ortogonali $|\psi_i\rangle$. Un qualsiasi vettore $|\psi\rangle$ normalizzato si può espandere in questa base nel modo seguente

$$|\psi\rangle = \Sigma_i a_i |\psi_i\rangle$$

considerandone la norma quadra si ottiene

$$\Sigma_{i,j} a_j * a_i \langle\psi_j|\psi_i\rangle = \Sigma_i |a_i|^2$$

che per la normalizzazione di $|\psi\rangle$ fa 1. Ora questo significa che gli $|a_i|^2$ rappresentano la probabilità che il sistema stia nello stato $|\psi_i\rangle$.

2.1.2 Osservabile

È naturale chiedersi ora a quali quantità algebriche corrispondano le osservabili e quindi che ruolo giochino nella nuova interpretazione quantomeccanica: classicamente una grandezza osservabile (come posizione, momento, energia, ecc...) è una variabile sperimentalmente misurabile (direttamente o indirettamente) e quindi numericamente quantificabile.
Nella trattazione hamiltoniana della meccanica classica, fissate posizione e momento nell'istante iniziale, lo stato del sistema è fissato ad ogni tempo, per cui una qualsiasi osservabile a=f(q,p) è univocamente determinata in ogni momento. Quantisticamente lo stato di un sistema è descritto completamente da un vettore nello spazio di Hilbert e le osservabili devono quindi essere collegate a quantità opportune che agendo su uno spazio di vettori restituiscano valori numerici: tali quantità sono operatori A:H$\longmapsto$H lineari autoaggiunti.
La richiesta di autoaggiunzione porta con sé attraverso il teorema spettrale alcune caratteristiche fondamentali delle osservabili: infatti come precedentemente detto un operatore autoaggiunto possiede uno spettro (ovvero l'insieme dei suoi autovalori) reale, questo è ciò che comunemente ci aspettiamo da una misura sperimentale, per cui la scelta di operare in uno spazio complesso diventa sensata perché comunque i risultati numerici che otteniamo sono sempre reali, inoltre gli autovalori a_i di A sono tutte e sole le misure che si possono ottenere per A.
Quindi un'osservabile, in quanto associato ad un operatore autoaggiunto, ammette sempre una base ortonormale dello spazio di Hilbert: questo significa che dato un qualsiasi stato $|\psi\rangle$ si può espandere rispetto agli autostati $|a_i\rangle$ associati alla misura a di A, con coordinate $\langle a|\psi\rangle$. Da quello che abbiamo precedentemente detto, il quadrato di queste

coordinate dà la probabilità che lo stato generico $|\psi\rangle$ dia il risultato a in seguito ad una misura dell'osservabile A.
Questo è tutto e solo ciò che è possibile sapere a priori di una misura sperimentale: ovvero mentre dalla meccanica classica il valore di un'osservabile per uno stato (q,p) è determinato univocamente dal valore di p e q all'istante desiderato, quantisticamente ciò è vero solo per gli autostati di A: per uno stato qualsiasi è nota a priori solo la probabilità che una misura dia il valore a_i. Inoltre classicamente dati due osservabili A e B si possono determinare ad ogni istante i valori a e b dalla dinamica dei q(t) e p(t), quantisticamente, avendo convertito in operatori le osservabili, queste non è detto che commutino, per cui in generale la misura di A e poi di B dà risultati diversi dell'osservazione di B e poi di A.
Questo algebricamente significa che non esiste una base ortonormale in cui entrambi gli operatori siano simultaneamente diagonali, per cui, se ciò accade, significa che le osservabili non sono compatibili.

2.1.3 Evoluzione temporale

Mentre il risultato di una misura per uno stato generico non può essere predetto, il determinismo è completamente conservato per la funzione d'onda $|\psi\rangle$: infatti questa obbedisce ad un'equazione differenziale del primo ordine nel tempo, l'eqazione di Schroedinger, la cui evoluzione temporale è univocamente determinata fissato il suo valore all'istante iniziale.

$$i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle$$

Dove H è l'operatore hermitiano associato all'osservabile energia e $\hbar$ è la costante di plank ridotta. Se associamo $|\psi\rangle$ ad una particella richiediamo che questa non possa essere creata o distrutta, per cui la probabilità deve essere conservata ad ogni istante: di conseguenza se ammettiamo che

$$|\psi(t)\rangle = U|\psi(t=0)\rangle$$

risulta che U deve essere un operatore unitario.
Estendendo le proprietà di analiticità dell'esponenziale di una variabile reale, all'esponenziale di un operatore nello spazio di Hilbert, si può formalmente esprimere l'operatore U(t)=exp($\frac{i}{\hbar}$Ht) in modo che $U(t)|\psi(t=0)\rangle$ sia soluzione dell'eqazione di Schroedinger.
È facile verificare che U è unitario dal momento che è esponenziale di un operatore autoaggiunto.

2.1.4 Misura

Come anticipato precedentemente, la meccanica quantistica offre una ricetta per l'operazione di misura: infatti i risultati osservabili su uno stato qualsiasi sono solo valori noti a

priori dipendenti solamente dall'osservabile A, ciò che dipende dallo stato soggetto alla misura è solo la probabilità con cui un particolare risultato esca.
Quando un determinato autovalore dell'osservabile viene estratto come misura dell'osservazione la funzione d'onda si dice che collassa nell'autostato corrispondente: tuttavia riguardo al modo in cui ci avvenga non si può sapere nulla, quello che succede in questo fenomeno è fisicamente inesplorabile.
Tutto ciò che si può conoscere è lo stato in cui il sistema è preparato e quello in cui si trova a seguito dell'osservazione, infatti lo stato a dopo la misura appartiene all'autospazio corrispondente al dato numerico ottenuto dal dispositivo sperimentale, quindi se l'autovalore è degenere, ovvero se più autostati danno lo stesso valore della misura, il sistema, a seguito dell'osservazione, sarà una qualsiasi combinazione lineare di questi autostati.
Nel caso particolare di autovalori non degeneri si avrà esattamente l'autostato corrispondente al valore numerico trovato, e qualsiasi misura successiva si faccia della stessa osservabile su questo stato darà lo stesso valore numerico.
Classicamente una misura di una quantità può avvenire con una discrezione tale da non modificare lo stato del sistema: si suppone infatti che l'interazione osservatore-osservando possa avvenire in maniera arbitrariamente piccola, dal punto di vista quantistico invece l'oggetto in esame è talmente piccolo che per misurarne le proprietà è necessaria una interazione che non pu più essere discreta e che quindi inevitabilmente ne modifica radicalmente lo stato.
In particolare una osservabile A è associata ad un insieme di operatori M_m di proiezione che agiscono sullo stato che deve essere misurato, dove il pedice m rappresenta il risultato numerico della misura che potrebbe uscire nell'esperimento. La probabilità che il risultato della misura sia m è pari a

$$p(m) = \langle\psi|M_m|\psi\rangle,$$

con il sistema lasciato nel corrispondente stato

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m|\psi\rangle}}.$$

Come sappiamo dall'algebra gli operatori M_m agendo su un vettore lo proiettano sul sottospazio corrispondente al valore m della misura, sono quindi idempotenti ($\mathrm{M}_m^2 = \mathrm{M}_m$), infatti l'operatore di proiezione su un sottospazio non fa nulla se agisce su un vettore di quel sottospazio. Essi soddisfano la relazione di completezza

$$\Sigma_m M_m = I,$$

e per il teorema spettrale

$$\Sigma_m m M_m = A.$$

Dove le somme sono estese a tutti i possibili risultati m della misura di A.
Inoltre logicamente $M_m M_{m'}=M_{m'}M_m=0$ se m≠m' e gli autostati corrispondenti sono ortonormali. Ad esempio dati due autostati $|+\frac{1}{2}\rangle$ e $|-\frac{1}{2}\rangle$ corrispondenti ad autovalori $+\frac{1}{2}$ e -$\frac{1}{2}$ della componente z dello spin. Gli operatori di proiezione sono dati da

$$M_{+\frac{1}{2}} = |+\tfrac{1}{2}\rangle\langle+\tfrac{1}{2}|$$
$$M_{-\frac{1}{2}} = |-\tfrac{1}{2}\rangle\langle-\tfrac{1}{2}|.$$

Essi soddisfano tutte le proprietà precedentemente elencate e danno una rappresentazione dell'operatore di spin lungo l'asse z

$$S_z=+\tfrac{1}{2}|+\tfrac{1}{2}\rangle\langle+\tfrac{1}{2}|-\tfrac{1}{2}|-\tfrac{1}{2}\rangle\langle-\tfrac{1}{2}|.$$

2.2 Operatore densità

Un sistema quantistico può essere descritto allo stesso modo sia attraverso il concetto di stato, che abbiamo descritto, sia con l'uso dell'operatore (o matrice) densità. Questo è particolarmente conveniente per descrivere sistemi quantistici il cui stato non sia completamente conosciuto. Più precisamente supponiamo che un sistema sia in uno tra un certo numero di stati $|i\rangle$ con probabilità p_i, dove chiamiamo le coppie $p_i, |i\rangle$ un ensamble di stati puri. Per questo sistema di definisce l'operatore densità come

$$\rho \equiv \Sigma_i p_i |i\rangle\langle i|.$$

Risulta che i postulati della meccanica quantistica possono essere riformulati nel nuovo linguaggio della matrice densità per cui i due approcci saranno equivalenti. Vederemo in questa parte come si può realizzare ciò e in quali casi è utile.
Supponiamo che la descrizione dell'evoluzione di un sistema quantistico chiuso sia descritta da un operatore unitario U. Se inizialmente il sistema è nello stato $|i\rangle$ con probabilità p_i, allora in seguito all'evoluzione il sistema sarà nello stato $U|i\rangle$ con probabilità p_i. Il corrispondente operatore densità è descritto dall'equazione

$$\rho = \Sigma_i p_i |i\rangle\langle i| \longmapsto \rho' = \Sigma_i p_i U|i\rangle\langle i|U^+ = U\rho U^+.$$

Analogamente per le misure, se gli operatori di proiezione M_m di un'osservabile A agiscono su uno stato $|\psi\rangle = \Sigma_i \sqrt{p_i}|i\rangle$ portandolo nello stato

$$M_m|\psi\rangle = \Sigma_i \sqrt{p_i} M_m |i\rangle$$

Pertanto, usando la propietà della traccia per cui

$$tr(A|\psi\rangle\langle\psi|) = \langle\psi|A|\psi\rangle,$$

se la probabilità che una misura di A sullo stato $|i\rangle$ dia come risultato m è

$$p(m|i) = \langle i|M_m^+ M_m|i\rangle = tr(M_m|i\rangle\langle i|),$$

Allora la probabilità totale che il risultato della misura di $|\psi\rangle$ sia m è

$$p(m) = \Sigma_i p_i p(m|i) = \Sigma_i p_i tr(M_m|i\rangle\langle i|M_m^+) = tr(M_m \rho M_m^+).$$

La matrice densità relativa ai nuovi stati misurati $|i^m\rangle = \frac{M_m|i\rangle}{\langle i|M_m^+ M_m|i\rangle}$

$$\rho_m = \Sigma_i p(i|m)|i^m\rangle\langle i^m| = \Sigma_i p(i|m)\frac{M_m|i\rangle\langle i|M_m^+}{\langle i|M_m^+ M_m|i\rangle}$$

e sapendo dalla teoria della probabilità che $p(i|m) = p(m|i)\frac{p_i}{p(m)}$ l'operatore densità assume la forma

$$\rho_m = \frac{M_m \rho M_m^+}{tr(M_m \rho M_m^+)}.$$

Inoltre come lo spazio degli stati di un sistema fisico composto si esprime come prodotto tensoriale degli n stati componenti ($|\psi\rangle = |\phi\rangle_1 \otimes |\phi\rangle_2 \otimes ... \otimes |\phi\rangle_n$), così in termini della matrice densità

$$\rho = \rho_1 \otimes \rho_2 \otimes ... \otimes \rho_n.$$

Un sistema il cui stato $|\psi\rangle$ sia conosciuto esattamente si dice essere *puro*, e la sua matrice densità è semplicemente $\rho = |\psi\rangle\langle\psi|$, invece se il sistema è in una miscela statistica di stati $|\psi_i\rangle$ con probabilità classiche p_i allora si dirà stato *misto*, con matrice densità

$$\rho = \Sigma_i p_i \rho_i,$$

ovvero una somma pesata con le probabilità degli operatori di proiezione sugli stati $|\psi_i\rangle$. È facile vedere che nel caso di uno stato puro $tr(\rho^2) = 1$, mentre in uno stato misto $tr(\rho^2) < 1$. Dati due sistemi A e B il cui stato è descritto dall'operatore densità ρ^{AB}, l'operatore densità ridotto si definisce come

$$\rho^A = tr_B(\rho^{AB}),$$

dove tr_B è una mappa che agisce dallo spazio A⊗B ad A chiamata traccia parziale sopra il sistema B:

$$tr_B(|a_1\rangle\langle a_2 \otimes |b_1\rangle\langle b_2|) \equiv |a_1\rangle\langle a_2 tr(|b_1\rangle\langle b_2|),$$

con $|a_1\rangle$ e $a_2\rangle$ vettori qualsiasi di A e $|b_1\rangle$ e $b_2\rangle$ elementi di B. ρ^A è quindi una descrizione dello stato del sistema A, nel senso che fornisce le corrette distribuzioni statistiche per le misure fatte sul sistema A.

Si definisce ora una generica operazione quantistica E, una mappa che manda operatori densità in operatori densità con tre proprietà:

i) tr[E(ρ)] è la probabilità che il processo rappresentato da E avvenga quando ρ è lo stato iniziale. Quindi, $0 \leq tr[E(\rho)] \leq 1$ per ogni stato ρ.

ii) E è una mappa lineare e convessa, per cui

$$E(\Sigma_i p_i \rho_i) = \Sigma_i p_i E(\rho_i).$$

iii) E è una mappa completamente positiva, ovvero E(A) deve essere postivo per ogni A operatore densità positivo.

2.3 Qubit

Un bit è una variabile a due stati che può assumere valori 0 o 1, il suo ruolo è centrale nell'informatica classica e rappresenta l'unità di definizione di uno stato logico. Il salto concettuale da classico a quantistico risiede, in buona parte, nel passaggio da stati distinti, ovvero di configurazioni diverse in cui un sistema può realizzarsi, a sovrapposizione di stati, cioè configurazioni intermedie tra stati distinti.
Abbattuto il primo scoglio concettuale correlato all'umana impossibilità di concepire la natura in questo modo, vediamo come i principi della meccanica quantistica si applicano all'informatica e come la crittografia si evolve di conseguenza.

Il qubit, anche detto quantum bit, è un sistema a due stati: $|0\rangle$ e $|1\rangle$. Tuttavia, vista la natura non classica specificata dal suo nome, per un qubit sono possibili anche stati intermedi di sovrapposizione:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

dove α e β sono numeri complessi. In altri termini un qubit è un vettore in uno spazio vettoriale complesso bidimensionale (C^2), di cui $|0\rangle$ e $|1\rangle$ sono stati ortonormali che chiameremo base computazionale.
Dalla precedente digressione sulla meccanica quantistica risulta quindi che α e β rappresentano attraverso il loro modulo quadro la probabilità che in seguito ad una misura di $|\psi\rangle$ si abbia rispettivamente lo stato $|0\rangle$ o $|1\rangle$. Ne segue che $|\alpha|^2 + |\beta|^2 = 1$, visto che la probabilità deve sommarsi a 1. Inoltre dall'algebra lineare si sa che dato uno spazio vettoriale qualsiasi vettore può essere espanso in una generica sua base, quindi oltre alla base ortonormale computazionale (che è evidentemente la più comoda) si può rappresentare un qubit in qualsiasi coppia di sovrapposizioni linearmente indipendenti degli stati $|0\rangle$ e $|1\rangle$, e quindi di cambiare a piacere base di riferimento a seconda delle necessità.
Un qubit poi essendo un vettore unitario in uno spazio di Hilbert bidimensionale complesso, può stare in un continuo di stati tra $|0\rangle$ e $|1\rangle$ finché non viene osservato. In quel caso il qubit $|\psi\rangle$ collassa con una certa probabilità nello stato $|0\rangle$ e $|1\rangle$.

È bene sottolineare il fatto che i qubit in natura esistono, essi non sono pura speculazione matematica: ad esempio in un modello atomico l'elettrone può essere misurato solo o nello stato fondamentale o in quello eccitato. Illuminando l'atomo con un'appropriata energia e per un determinato lasso di tempo è possibile far transitare l'elettrone dallo stato fondamentale a quello eccitato e viceversa. Tuttavia riducendo il tempo di esposizione un'elettrone inizialmente nello stato fondamentale può spostarsi a metà strada tra $|0\rangle$ e $|1\rangle$, ovvero nello stato

$$\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$$

che chiameremo $|+\rangle$. Denoteremo invece con $|-\rangle$ lo stato $\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$. Una misura di $|+\rangle$ dà come risultato un elettrone, il 50% delle volte nello stato fondamentale e il 50% delle volte nello stato eccitato.
A causa di ciò un qubit può teoricamente contenere una quantità infinita di informazione, visto che infiniti sono i modi possibili in cui un qubit può trovarsi. Tuttavia questa informazione esiste solo finché non viene misurata: in quel momento il qubit viene in un certo senso *resettato* con una certa probabilità o nello stato $|0\rangle$ o nello stato $|1\rangle$, che quindi sono la base ortonormale completa di autostati dell'energia.
Per quanto detto nella sezione introduttiva di meccanica quantistica, quindi questi due stati sono gli unici per cui sia determinato a priori il risultato della misura, e rappresentano un bit classico.
Tuttavia finché tale misurazione non viene fatta, questa infinità di informazione viene conservata, informazione che inoltre cresce esponenzialmente con il numero di qubit.

Misura in base non computazionale

Fino ad ora abbiamo descritto una misura del qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ come un'operazione che dà come risultato 0 o 1 e lascia il qubit nel corrispondente stato $|0\rangle$ o $|1\rangle$ con probabilità rispettivamente di $|\alpha|^2|$ e $|\beta|^2$. Tuttavia lo spettro di tutte le possibili misurazioni a nostra disposizione è molto ampio: infatti $|0\rangle$ e $|1\rangle$ è solo una delle infinite basi ortonormali che possiamo prendere per generare tutto C^2. Un'altra possibile scelta è data dai vettori $|+\rangle$ e $|-\rangle$, per cui l'arbitrario stato $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ si può riesprimere nella nuova base come

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \alpha\frac{|+\rangle+|-\rangle}{\sqrt{2}} + \beta\frac{|+\rangle-|-\rangle}{\sqrt{2}} = \frac{\alpha+\beta}{\sqrt{2}}|+\rangle + \frac{\alpha-\beta}{\sqrt{2}}|-\rangle.$$

È quindi possibile trattare gli stati $|+\rangle$ e $|-\rangle$ come nuova base rispetto a cui misurare i qubits. Naturalmente in seguito ad un misura la probabilità che il sistema sia nello stato $|+\rangle$ è data da

$$p(+) = \langle\psi|M_+|\psi\rangle = |\langle+|\psi\rangle|^2 = \frac{|\alpha+\beta|^2}{2},$$

e analogamente per lo stato $|-\rangle$

$$p(-) = \langle\psi|M_-|\psi\rangle = |\langle-|\psi\rangle|^2 = \frac{|\alpha-\beta|^2}{2}.$$

Fotoni

L'esempio più semplice di qubit che si può trovare in natura è dato dal fotone. In particolare lo stato di polarizzazione di un fascio di luce è motivato microscopicamente dalla polarizzazione dei fotoni costituenti il fascio. Classicamente facendo incidere un'onda

elettromagnetica polarizzata linearmente di intensità I su un polarizzatore, una porzione $Icos^2(\theta)$ viene trasmessa, e una parte $Isin^2(\theta)$ viene assorbita. Quindi θ rappresenta l'angolo che forma la direzione di polarizzazione dell'onda con la normale all'asse del polarizzatore. Quantisticamente visto che la luce è composta da fotoni con energia $\hbar\omega$ significa che una porzione $N(\theta) = \frac{Icos^2(\theta)}{\hbar\omega}$ viene trasmessa mentre l'altra viene assorbita. Si potrebbe pensare che visto che l'effetto è quello di ridurre l'energia del fascio trasmesso, ciò sia dovuto ad una diminuzione media dell'energia di fotoni: tuttavia questo sarebbe correlato ad una diversa colorazione del fascio trasmesso, fatto che non è osservato sperimentalmente.
Semplicemente, i risultati possibili della misura sono due: o il fotone passa, o viene assorbito. In particolare il fotone avrà probabilità 1 di passare se è polarizzato perpendicolarmente all'asse ottico, mentre ha probabilità 0 se è polarizzato parallelamente. Questi sono gli unici due stati in cui una misura della polarizzazione dà risultato conosciuto a priori, negli altri casi invece si può conoscere prima di compiere la misura solamente la probabilità con cui il fotone passi o venga assorbito. Quindi uno fascio polarizzato obliquamente rispetto all'asse ottico e alla sua normale, riducendo l'intensità della luce in modo da generare un fotone alla volta, farà passare mediamente un fotone su due e con polarizzazione perpendicolare all'asse ottico. Questo fenomeno fisico contiene tutto ciò che serve per capire il funzionamento di una misura quantistica.
Per chiudere il cerchio quindi identifichiamo come base di misura di uno stato di polarizzazione i diversi modi in cui possiamo orientare il polarizzatore nello spazio, in particolare si può considerare la base $|0\rangle$ e $|1\rangle$ come gli autostati relativi alla misura di un polarizzatore con asse ottico verticale, mentre la base $|+\rangle, |-\rangle$ quella derivante dalla disposizione del polarizzatore con l'asse ottico a 45 gradi con a verticale.

Qubit Multipli

Il discorso si estende automaticamente al caso in cui i qubits siano due: infatti mentre classicamente questo corrisponde alle quattro possibili configurazioni 00, 01, 10, 11, quantisticamente hanno controparte in quattro stati sovrapponibili $|00\rangle$, $|01\rangle$, $|10\rangle$ e $|11\rangle$ che rappresentano la base computazionale di $\mathbb{C}^2$, che è solo una delle infinite basi ortonormali equivalenti dello spazio vettoriale. Per cui due qubits generici in questa base avranno espressione:

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$

con la solita condizione di normalizzazione degli α_{ij}. In questa situazione però è possibile compiere misure su sottospazi dei qubit, ossia ad esempio vedere lo stato del primo qubit: questo ad sarà 0 con probabilità $|\alpha_{00}|^2 + |\alpha_{01}|^2$ lasciando il qubit nello stato

$$|\psi'\rangle = \frac{\alpha_{00}|00\rangle + \alpha_{01}|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}},$$

mentre sarà 1 con probabilità $|\alpha_{10}|^2 + |\alpha_{11}|^2$ e il corrispondente stato

$$|\psi'\rangle = \frac{\alpha_{10}|10\rangle + \alpha_{11}|11\rangle}{\sqrt{|\alpha_{10}|^2 + |\alpha_{11}|^2}}.$$

Per determinare quindi univocamente lo stato di due qubits è necessario compiere una misura su entrambi i sottospazi C di C^2. In questo caso, prendendo sempre per comodità la base computazionale, avremo che lo stato $|\psi\rangle$ collasserà nello stato $|ij\rangle$ con probabilità $|\alpha_{ij}|^2$.

2.3.1 Computazione Quantistica

Come un computer classico è basato su cavi e porte logiche, un computer quantistico necessita di appropriati circuiti e porte, rispettivamente, per trasferire e manipolare informazioni quantistiche. Per ogni porta logica classica che dati uno o più bit in ingresso restituisce il risultato di un'operazione (completamente specificata dalla tavola di verità), così i quantum gates agiscono su qubit in ingresso modificandoli in qualche modo. Tuttavia la conoscenza della tavola di verità di una porta comunica come essa agisce sulla basa ortonormale computazionale, e non ci dice come agisce su un generico stato sovrapposizione, a meno che il quantum gate non agisca linearmente, in tal caso infatti nota l'azione di una porta sullo stato $|0\rangle$ e $|1\rangle$ è nota la sua azione sul generico stato

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle.$$

Inoltre visto che l'azione su un generico stato normalizzato deve dare un corrispondente stato normalizzato, risulta che gli operatori associati alle porte logiche devono essere unitari, e viceversa qualsiasi operatore unitario rappresenta una potenziale porta logica per un circuito quantistico.
Risulta che mentre nel caso classico l'unica porta non triviale a un bit è il NOT logico, quantisticamente vi sono un'infinità di possibili gates non banali per un singolo qubit.
Per dare qualche esempio, in notazione vettoriale il generico vettore $|\psi\rangle$ si può esprimere come un vettore bidimensionale cui la prima componente è riferita allo stato $|0\rangle$, e la seconda allo stato $|1\rangle$. Per cui qualsiasi matrice quadrata bidimensionale unitaria rappresenta una porta possibile, ad esempio la porta NOT si rappresenta come:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

e la sua azione su $|\psi\rangle$ dà come risultato

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}$$

come ci aspettiamo.

Introduciamo ora alcune porte che torneranno utili in seguito: lo Z gate

$$Z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

che lascia lo stato $|0\rangle$ invariato e cambia il segno di $|1\rangle$, e l'Hadamard gate

$$H \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Questo gate porta lo stato $|0\rangle$ nello stato $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, e lo stato $|1\rangle$ nello stato $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Inoltre applicando H due volte a qualsiasi stato lo lascia immutato, infatti

$$H^2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Il discorso ora si generalizza automaticamente a porte per n qubit multipli: questi si possono rappresentare come vettori colonna 2n dimensionali e le porte associate come matrici unitarie di dimensione 2n×2n.
Introduciamo ora una porta che ci servirà nella prossima sezione: il controlled-NOT o CNOT gate, ha due qubits di ingresso, di cui il primo funge da controllo mentre il secondo è quello su cui si opera. Infatti se il primo è nello stato $|0\rangle$ il secondo viene lasciato immutato, se invece è nello stato $|1\rangle$ il secondo viene cambiato.
In forma matriciale questo gate è rappresentato da una matrice unitaria 4×4 che chiamiamo U_{CN}, ed è definita come:

$$\mathrm{U}_{CN} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Questa porta è particolarmente interessante perché qualsiasi porta logica per qubits multipli può essere decomposta in una CNOT e una porta per qubit singolo, e rappresenta l'analogo quantistico dell'universialità del NAND gate per bit classici.

Stati di Bell

Introduciamo ora quattro configurazioni a due qubit molto importanti detti stati di Bell o coppie EPR, da Einstein Podolsky e Rosen che per primi ne studiarono le proprietà.
Per ottenerli si parte dalla base computazionale $|00\rangle$, $|01\rangle$, $|10\rangle$ e $|11\rangle$ su cui si applica l'Hadamard gate seguito da un CNOT.

$$\begin{aligned}
|00\rangle &\longmapsto |\beta_{00}\rangle = \tfrac{1}{\sqrt{2}}|00\rangle + \tfrac{1}{\sqrt{2}}|11\rangle, \\
|01\rangle &\longmapsto |\beta_{01}\rangle = \tfrac{1}{\sqrt{2}}|01\rangle + \tfrac{1}{\sqrt{2}}|10\rangle, \\
|10\rangle &\longmapsto |\beta_{10}\rangle = \tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle, \\
|11\rangle &\longmapsto |\beta_{11}\rangle = \tfrac{1}{\sqrt{2}}|01\rangle + \tfrac{1}{\sqrt{2}}|10\rangle.
\end{aligned}$$

In particolare lo stato $|\beta_{00}\rangle$ e $|\beta_{11}\rangle$ sono particolarmente interessanti essendo alla base del così detto teletrasporto quantistico. La loro peculiarità sta nel fatto che eseguendo una misura sul primo quibit, l'altro è automaticamente determinato, nonostante né il primo né il secondo abbiano valore determinato: questo significa che i due qubit sono strettamente correlati, correlazione che persiste anche in seguito all'applicazione di un'operazione su uno dei due qubits.

Consideriamo per lo stato $|\beta_{00}\rangle$ la matrice densità

$$\rho = (\frac{|00\rangle+|11\rangle}{\sqrt{2}})(\frac{\langle 00|+\langle 11|}{\sqrt{2}}) = \frac{|00\rangle\langle 00|+|11\rangle\langle 00|+|00\rangle\langle 11|+|11\rangle\langle 11|}{2}.$$

Calcolando la traccia sopra il secondo qubit troviamo l'operatore denistà relativo allo stato 1:

$$\rho^1 = tr_2(\rho) = \frac{|0\rangle\langle 0|\langle 0|0\rangle+|1\rangle\langle 0|\langle 0|1\rangle+|0\rangle\langle 1|\langle 1|0\rangle+|1\rangle\langle 1|\langle 1|1\rangle}{2} = \frac{|0\rangle\langle 0|+|1\rangle\langle 1|}{2} = \frac{I}{2}.$$

Si noti che questo è uno stato misto, infatti $tr((I/2)^2) = 1/4 < 1$. Questo fatto è particolare perché lo stato del sistema congiunto è puro ($\rho = |+\rangle\langle +|$), tuttavia il primo qubit è in uno stato misto, ovvero uno stato di cui a priori sappiamo solamente la probabilità con cui possa collassare negli stati puri $|0\rangle$ e $|1\rangle$.

2.4 Entanglement

Introduciamo ora uno degli argomenti più importanti della computazione quantistica, ovvero l'entanglement. Il nome deriva da un articolo di Schroedinger sulle coppie EPR, e letteralmente significa *aggrovigliamento* (o intrecciamento): questo infatti rappresenta un tipo di interazione che non possiede analogo classico e che permette a due sistemi di interagire a distanza istantaneamente.
Questo riguarda sistemi accoppiati in cui una misura su una parte del sistema fornisce automaticamente lo stato della parte non soggetta a osservazione: nello stato $|+\rangle$ descritto precedentemente, ad esempio avevamo detto come una misura del primo qubit fornisca automaticamente lo stato del secondo.
Questo legame, inoltre, persiste anche in seguito all'esecuzione di operazioni sulle due parti, causando su uno gli effetti di una misura dell'altro.
Nonostante la descrizione superficiale questi sistemi risultano evidentemente comodissimi come canali di comunicazione e come apparati crittografici: infatti lo stretto legame tra le parti limita quasi totalmente la possibilità di intercettazione, ponendo forse fine una volta per tutte alla lotta tra crittografi e crittoanalisti.

Per dare un chiarimento ed un esempio di questo genere di interazione consideriamo uno dei risultati più interessanti che si possono realizzare con questi sistemi, ovvero il teletrasporto quantistico.

2.4.1 Quantum Teleportation

Paola e Chiara sono amiche di infanzia che purtroppo nel corso degli anni si sono perse di vista. Quando ancora insieme un giorno crearono per gioco una coppia EPR

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

prendendo poi ciascuna un qubit dello stato legato in segno della loro amicizia.
Col passare degli anni le comunicazioni tra le due sono sempre meno frequenti fino ad arrivare alla quasi totale indifferenza, tuttavia entrambe hanno in qualche modo continuato a tenere addosso la propria metà della coppia EPR.
Alcuni anni dopo Paola è costretta a nascondersi e Chiara ha il compito di inviale un qubit $|\psi\rangle$, potendo comunicare solamente dell'informazione classica.
La situazione sarebbe disperata per diversi motivi: Chiara non conosce lo stato $|\psi\rangle$ da mandare a Paola, inoltre la meccanica quantistica, attraverso il teorema di no-cloning che vedremo in seguito, vieta a Chiara di determinare lo stato del qubit avendone a disposizione solo una copia. Inoltre se anche sapesse lo stato del qubit le servirebbe una quantità infinita di informazione classica per descrivere precisamente lo stato di $|\psi\rangle$, potendo questo assumere valori in un range continuo.
Tuttavia, ormai quasi rassegnata, Chiara si ricorda di avere con sé la sua metà della coppia EPR e, memore dei suoi studi di meccanica quantistica, di come inviare un qubit di informazione avendo a disposizione uno stato entangled e un canale di comunicazione classico.
Chiara quindi fa interagire il qubit da inviare

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

dove le ampiezze α e β sono le ampiezze sconosciute, con la sua metà della coppia EPR. Lo stato totale del sistema è allora:

$$|\psi_0\rangle = |\psi\rangle|\beta_{00}\rangle = \frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle)],$$

dove abbiamo usato la convenzione che i primi due qubit a sinistra appartengano a Chiara e il terzo a Paola.
A questo punto Chiara manda i due qubit in suo possesso attraverso un CNOT gate, ottenendo

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}[\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)].$$

Quindi manda il primo qubit attraverso un Hadamard gate, ottenendo

$$|\psi_2\rangle = \frac{1}{2}[\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)].$$

Riordinando i termini, lo stato può essere riscritto come:

$$|\psi_2\rangle = \frac{1}{2}[|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)],$$

dove i primi due qubit rappresentano sempre lo stato dei due qubits in possesso di Chiara, mentre il terzo è quello di Paola. Ora a seconda dello stato in cui Chiara trova i suoi qubits in seguito ad una misura, lo stato del qubit in possesso di Paola sarà una qualche variante dello stato $|\psi\rangle$ di partenza.
A Chiara quindi basterà inviare a Paola una coppia di bit classici, guarda caso proprio l'unico strumento che aveva a disposizione per comunicare, per far risalire Paola allo stato $|\psi\rangle$ di partenza, attraverso l'applicazione di un quantum gate opportuno.

Il fatto di avere bisogno di un messaggio classico con scritto il risultato della misura ottenuta da Chiara per trasferire dell'informazione a Paola, evita l'imbarazzo di dover rivedere i principi della relatività ristretta. Infatti, senza la comunicazione classica, che può essere inviata al massimo alla velocità della luce, del risultato ottenuto da Chiara, lo stato del qubit in possesso di Paola continua ad essere indeterminato, anche se esso cambia istantaneamente in seguito all'applicazione del protocollo di teletrasporto.

Come si vede dal protocollo di teletrasporto quantistico, il messaggio $|\psi\rangle$ viene totalmente perso dalla disponibilità di Chiara nel momento in cui va a misurare la coppia dei qubits in suo possesso, rimanendone traccia solo nel qubit di Paola.
Questo fatto riguarda la meccanica quantistica in generale: infatti esiste un teorema che dimostra l'impossibilità di realizzare una copia perfetta di uno stato $|\psi\rangle$ sconosciuto usando evoluzioni unitarie, senza contemporaneamente distruggere l'originale.

No-cloning Theorem

Supponiamo di avere due sistemi quantistici A e B con spazio di Hilbert comune H=H_A=H_B, con l'obiettivo di realizzare una procedura che permetta di copiare lo stato $|\phi\rangle_A$ del sistema A in B inizialmente in un generico $|e\rangle_B$ normalizzato. Per fare ciò occorre combinare i due sistemi nel prodotto tensoriale dei due stati

$$|\phi\rangle_A \otimes |e\rangle_B.$$

In questa situazione di combinazione è possibile compiere due tipi di operazioni: un'osservazione, che fa collassare il sistema in qualche autostato di un osservabile eliminando l'informazione contenuta nei qubits (che ovviamente non vogliamo), oppure attraverso operatori unitari dello spazio $H \otimes H$ che modifica senza misurare lo stato composto. Tuttavia nessuno di questi operatori può clonare tutti gli stati, cioè non esiste nessun U su H⊗H tale che per tutti gli stati normalizzati $|\phi\rangle_A e|e\rangle_B$ in H

$$U(|\phi\rangle_A|e\rangle_B) = |\phi\rangle_A|\phi\rangle_B,$$

a meno di fattori di fase globali che non producono effetti osservabili.
Per dimostrare ciò prendiamo un'arbitraria coppia di stati $|\phi\rangle_A$ e $|\psi\rangle_A$ in H, e costruiamo la corrispondente coppia in H⊗H come $|\phi\rangle_A|e\rangle_B$ e $|\psi\rangle_A|e\rangle_B$. Ora visto che U è unitario

$$\langle\phi|\psi\rangle\langle e|e\rangle \equiv \langle\phi|_A\langle e|_B|\psi\rangle_A|e\rangle_B = \langle\phi|_A\langle e|_B U^+U|\psi\rangle_A|e\rangle_B = \langle\phi|_A\langle\phi|_B|\psi\rangle_A|\psi\rangle_B = \langle\phi|\psi\rangle^2$$

Per cui avendo supposto $|e\rangle$ normalizzato abbiamo

$$|\langle\phi|\psi\rangle|^2 = |\langle\phi|\psi\rangle|.$$

L'equazione $x^2 = x$ ammette due soluzioni: x=0 e x=1, quindi o i due stati arbitrari $|\phi\rangle_A$ e $|\psi\rangle_A$ sono lo stesso stato a meno di fattori di fase ($|\langle\phi|\psi\rangle| = 1$), oppure sono ortogonali ($|\langle\phi|\psi\rangle| = 0$). In entrambi i casi non sono stati arbitrari, quindi un singolo operatore unitario universale U non può clonare un generico stato quantistico provando il teorema di no-cloning.
Tuttavia quello che possiamo fare è considerare copie imperfette, questo si può realizzare accoppiando un sistema ausiliario più grande a quello da copiare e applicare una trasformazione unitaria a questo nuovo stato. Scegliendo opportunamente l'operatore, diverse componenti dello sistema combinato evolveranno in copie approssimate dell'originale. Nel 1996, Buzek e Hillery mostrarono che una macchina universale per clonare poteva realizzare una copia di uno stato sconosciuto con una precisione oltre l'80%[2].

2.5 Algoritmi Quantistici

Alcuni problemi interessanti sono impossibili da risolvere su computer classici, non per questioni di principio, ma per l'enorme quantità di risorse necessarie per la loro risoluzione.
Abbinati alla ridefinizione del bit e delle porte logiche, vi sono nuovi algoritmi che rendono affrontabili problemi impossibili per computer classici. Di questi nuovi algoritmi ne esistono fondamentalmente due tipi: il primo è la trasformata di Fourier quantistica di Shor, il secondo l'algoritmo di Grover per il quantum searching.
Questi ci interessano particolarmente perché entrambi si applicano alla crittografia: infatti l'algoritmo di Grover permette di accelerare la ricerca di chiavi per sistemi crittografici come l'ampiamente usata Data Encryption Standard, e la trasformata di Fourier quantistica fornisce soluzioni a problemi noi cari come quello del calcolo del logaritmo discreto e della fattorizzazione.
Il fatto che questo algoritmo possa non solo risolvere il problema della fattorizzazione, ma farlo esponenzialmente più velocemente del migliore conosciuto per computer classici, rende per un computer quantistico possibile rompere molti dei più diffusi sistemi di cifratura esistenti, tra cui l'RSA.

[2]Per ulteriori approfondimenti si rimanda a [1]

Tuttavia è bene sottolineare come non sia facile in generale formulare algoritmi, e a maggior ragione realizzarne di quantistici che funzionino meglio di quelli classici, soprattutto perchè le nostre intuizioni si adattano molto meglio al mondo classico rispetto a quello quantistico. Infatti se pensiamo ad un problema con la nostra intuizione nativa arriveremo inevitabilmente ad un algoritmo classico, per formularne di quantistici bisogna invece percorrere un'altra strada, ricorrendo a trucchi per abbandonare il ragionamento classico e pensare nella maniera corretta.

Quantum Fourier Transform e la Fattorizzazione

La trasformata di Fourier quantistica è un efficiente algoritmo quantistico per la risoluzione di molti problemi computazionali irrisolvibili con apparati classici.
Classicamente, la trasformata di Fourier discreta trova tantissime applicazioni e consiste in un cambio di base che ad un vettore complesso N-dimensionale x_0,...,x_{N-1} associa un nuovo vettore della stessa dimensione y_0,...,y_{N-1} definito come

$$y_k \equiv \frac{1}{\sqrt{N}} \Sigma_j x_j e^{2\pi i j k/N}.$$

La trasformata di Fourier quantistica ha espressione analoga, anche se si applica a vettori ket:

$$|j'\rangle \equiv \frac{1}{\sqrt{N}} \Sigma_k e^{2\pi i j' k/N} |k\rangle.$$

Anche se dalla definizione non è chiaro, questa è una trasformazione unitaria e quindi è possibile implementarla in un computer quantistico.
Di seguito descriveremo come questa si applica efficientemente al problema della fattorizzazione: il termine efficientemente non è utilizzato a sproposito ma significa che attraverso questo algoritmo è computazionalmente possibile fattorizzare un numero n molto grande. Infatti mentre per un computer classico la fattorizzazione di un intero di n-bit richiede $exp(\Theta(n^{1/3} \log^{2/3} n))$ operazioni usando il migliore algoritmo classico conosciuto, quello quantistico restituisce lo stesso risultato usando $O(n^2 \log n \log\log n)$ operazioni. Per paragonare le due scale di tempi basti pensare che per un input con n=100, classicamente impiega un tempo paragonabile all'età dell'universo, quantisticamente invece nell'ordine di giorni. Quindi un computer quantistico può fattorizzare un numero esponenzialmente più velocemente del migliore classico algoritmo conosciuto.
L'algoritmo di fattorizzazione è piuttosto complicato e fa uso del quantum order-finding, ovvero un altro algoritmo che sfrutta la trasformata di Fourier quantistica. Questo, dati x e N interi positivi con x < N, senza fattori in comune, determina l'ordine di x modulo N: ovvero il più piccolo intero positivo r tale che x^r=1(mod N). A sua volta questo ricorre ad un altro algoritmo per la quantum phase-estimation, che passa sempre per la quantum Fourier transform, per determinare la fase dell'autovalore relativo all'autovettore di un operatore unitario.

Questi algoritmi risolvono polinomialmente problemi computazionalmente difficili da affrontare classicamente, tuttavia lo fanno solo con buona probabilità. Infatti fanno parte di una classe di algoritmi, classici e quantistici, detti probabilistici o randomizzati, di cui esistono due tipologie: quelli che danno un risultato sicuramente corretto in tempi probabilmente brevi (di tipo Las Vegas) e quelli, di cui stiamo parlando, che danno un risultato probabilmente corretto in tempi sicuramente brevi (di tipo Monte Carlo). Il fatto di avere un risultato eventualmente sbagliato è tollerabile in linea di principio a patto che la probabilità con cui questo sbagli possa essere ridotta a piacere e che sia minore di quella che si verifichi un altro genere di errore dovuto a malfunzionamento del calcolatore.

Capitolo 3

Crittografia Quantistica

Nell'ultima parte abbiamo visto che con computer quantistici è possibile rompere alcuni dei migliori sistemi di cifratura a chiave pubblica esistenti. Per fortuna però ciò che la meccanica quantistica toglie con una mano, dà indietro con l'altra: infatti vi è una procedura chiamata *quantum cryptography* o *quantum key distribution* che sfrutta i principi della meccanica quantistica per permettere la distribuzione di chiavi riservate in maniera certamente sicura senza passare per canali privati: tali chiavi poi devono essere sfruttate a dovere in cifrari perfetti a private key di tipo One Time Pad.
Dal momento che la meccanica quantistica, come ogni teoria fisica, funziona finchè non se ne dimostra sperimentalmente l'invalidità, effettivamente la situazione è simile a quella relativa ai cifrari a public key, fondati su problemi matematici che si postulano essere impossibili da risolvere. Tuttavia, fino a quando non se ne dimostri l'invalidità, essa funziona e i sistemi di cifratura che ne derivano sono certamente sicuri.
Questo significa che i crittoanalisti della nostra epoca dovranno essere ottimi fisici sperimentali che riescano ad osservare eventuali situazioni in cui cadono i postulati della meccanica quantistica, o in alternativa darsi per vinti, ponendo fine una volta per tutte alla lotta millenaria tra crittografi e crittoanalisti.
In questa ultima sezione quindi andremo a descrivere la procedura di distribuzione quantistica della chaive, ne daremo alcuni esempi pratici e ne discuteremo la sicurezza. Infine parleremo di Qnet:la prima rete a crittografia quantistica funzionante.

3.1 Quantum Key Distribution

Il QKD è un protocollo la cui sicurezza è provabile, attraverso cui i bit di una chiave privata possono essere creati tra due parti tramite un canale pubblico e poi sfruttate in cifrari perfetti. Come abbiamo visto precedentemente, questi sistemi crittografici sono in assoluto i più sicuri, tuttavia peccano in efficienza dal momento che la chiave deve essere scambiata di volta in volta attraverso un canale protetto, eventualmente anche co-

stoso (come un incontro di persona). La meccanica quantistica risolve questo problema fornendo un metodo sicuro a cui è sufficiente la comunicazione attraverso un canale non protetto. L'unica richiesta per il protocollo QKD è che i qubits possano essere comunicati attraverso il canale pubblico con un tasso di errore limitato.
L'idea alla base della sicurezza del QKD sta nel fatto che un crittoanalista non può in alcun modo estrarre informazione dal qubit trasmesso da Paola a Chiara senza disturbarne lo stato. Prima di tutto per il teorema di no-cloning, non è possibile copiare il qubit di Paola senza modificarlo, inoltre, più in generale, qualsiasi estrazione di informazione implica un disturbo: in particolare questo vale in ogni tentativo di compiere una distinzione tra due stati non ortogonali. Se ne da di seguito una prova:
Prendiamo due stati non ortogonali $|\psi\rangle$ e $|\phi\rangle$ riguardo cui il crittoanalista vuole ottenere informazione. Senza perdere generalità possiamo assumere che il processo usato per questo scopo si basi sul fare interagire gli stati attraverso trasformazione unitaria con un dispositivo preparato in uno stato standard $|u\rangle$. Assumendo che questo processo non disturbi gli stati $|\psi\rangle$ e $|\phi\rangle$ di partenza, otteniamo:

$$\begin{aligned} |\psi\rangle|u\rangle &\longmapsto |\psi\rangle|v\rangle \\ |\phi\rangle|u\rangle &\longmapsto |\phi\rangle|v'\rangle \end{aligned}$$

dove la condizione $|v\rangle \neq |v'\rangle$ è indispensabile al crittoanalista per estrarre informazione sull'identità degli stati. Tuttavia dal momento che i prodotti interni sono conservati sotto trasformazioni unitarie, si deve avere che

$$\begin{aligned} \langle v|v'\rangle\langle\psi|\phi\rangle &= \langle u|u\rangle\langle\psi|\phi\rangle \\ \langle v|v'\rangle &= \langle u|u\rangle = 1, \end{aligned}$$

che implica l'uguaglianza tra $|v\rangle$ e $|v'\rangle$. Quindi per distinguere tra $|\psi\rangle$ e $|\phi\rangle$ è necessario disturbare almeno uno dei due stati.

Quindi sfruttiamo questo risultato trasmettendo qubits di stati non ortogonali tra Paola e Chiara. Controllando il disturbo negli stati trasmessi, stabiliscono un limite massimo all'errore, causato da naturale malfunzionamento dei dispositivi o intercettazione, nel loro canale di comunicazione. In seguito applicano una procedura che permette loro di passare dalla situazione in cui loro possiedono chiavi leggermente diverse di cui il crittoanalista è riuscito ad estrarre alcune informazioni, a quella, favorevole, in cui hanno una chiave identica, più corta dell'originale, ma riducendo a piacere l'informazione in mano all'intercettatore. Di seguito descriviamo questo protocollo che si applica in generale a sistemi di tipo private key, chiamato *Privacy amplification and information reconciliation.*

3.1.1 Privacy amplification and information reconciliation

Abbiamo visto come la parte fondamentale dei sistemi a Private Key sia lo scambio della chiave, ma cosa succede se questa è imperfetta?

Mettiamo caso che Paola e Chiara condividano una chiave *quasi* identica, e che abbiano una stima della discrepanza tra le due. Questa differenza tra le due chiavi può essere causata da un'intercettazione ma anche da imperfezioni nella linea di trasmissione e, visto che è impossibile distinguere tra le due eventualità, per garantire la sicurezza è opportuno assumere che tutti gli errori siano dovuti a intercettazione.
Noto che il tasso di errore tra le chiavi è sotto una certa soglia, può essere attuata una procedura nota col nome di *Privacy amplification and information reconciliation* per eliminare i bits sbagliati e poi ridurre a piacere l'informazione che l'intercettatore possiede sulla chiave.

Information reconciliation

Questa è una tecnica di correzione dell'errore che si attua tra le chiavi di Paola e Chiara per renderle identiche. Il processo viene fatto attraverso un canale pubblico, per cui è fondamentale minimizzare l'informazione inviata riguardo a ciascuna chiave dal momento che qualunque messaggio potrebbe essere soggetta a intercettazione.
Un protocollo comune è quello del metodo a cascata, proposto nel 1994. Esso agisce in diversi turni, ogni volta le chiavi sono raggruppate in blocchi su cui viene fatto un controllo della parità: questo è una delle più semplici forme di error detecting e consiste nel determinare se i bit accesi sono in numero pari o dispari. Nel caso in cui questo test mostri delle differenze di parità in blocchi corrispondenti, viene effettuato un controllo più specifico per trovare ed eliminare l'errore. Inoltre se in un turno successivo si trova una discrepanza di parità dove precedentemente non vi era errore, allora significa che ve ne sono in realtà almeno due. Il nome deriva dalla ricorsività con cui si compie il controllo procedendo ricorsivamente e considerando sempre divisioni diverse delle chiavi. Dopo un certo numero di controlli Paola e Chiara hanno con buona probabilità chiavi identiche, tuttavia il crittoanalista in ascolto sul canale pubblico ha acquisito ulteriore informazione riguardo le chiavi tramite la conoscenza di queste informazioni sulla parità.

Privacy amplification

Complementare all'*information reconciliation* vi è la tecnica di *privacy amplification* che come suggerisce il nome si occupa di ridurre, fino ad eliminare, l'informazione parziale ottenuta dal crittoanalista sulla chiave: informazione che si suppone dovuta sia a intercettazione, che inevitabilmente introduce un errore misurabile, sia dall'ascolto dell'*information reconciliation*, dove assumiamo che il crittoanalista ottenga tutta l'informazione possibile sulla parità. Questo metodo consiste semplicemente nel prendere la chiave di Paola e Chiara, e produrne una più breve ma di cui il crittoanalista possieda solo informazione trascurabile.
Ciò è possibile usando la classe di *universal hash functions*: queste sono una classe di funzioni che mappano un insieme di stringhe di n-bit (le vecchie chiavi) in uno di m-bit

(le nuove chiavi) in modo tale che, presa una di queste a caso, sia poco probabile che due stringhe diverse di n bit corrispondano alla stessa stringa di m bit. È possibile dimostrare che, nota quanta informazione il crittoanalista ha raccolto con l'intercettazione e i dati sulla parità, m può essere preso sufficientemente piccolo da rendere arbitrariamente grande l'incertezza del crittoanalista sulla nuova chiave, rendendola sicura a piacere.

3.2 Esempi di QKD

Esponiamo di seguito tre procedure che sfruttano la Quantum Key Distribution.

Protocollo BB84

Paola parte con a e b, due stringhe ciascuna di 4n bits classici casuali. Quindi codifica queste stringhe in un blocco di 4n qubits,

$$|\psi\rangle = |\psi_{a_1 b_1}\rangle \otimes |\psi_{a_2 b_2}\rangle \otimes ... \otimes |\psi_{a_{4n} b_{4n}}\rangle,$$

dove a_i e b_i sono l'i-esimo bit di a e di b, e ciascun qubit è uno dei quattro stati

$$\begin{aligned} |\psi_{00}\rangle &= |0\rangle, \\ |\psi_{10}\rangle &= |1\rangle, \\ |\psi_{01}\rangle &= |+\rangle, \\ |\psi_{11}\rangle &= |-\rangle.^{1} \end{aligned}$$

L'effetto di questa procedura è quello di codificare a, nella base X o Z a seconda del valore di b. Si noti che i quattro stati non sono tutti ortogonali tra loro, quindi nessuna misura li può distinguere tutti con certezza.
Paola invia $|\psi\rangle$ a Chiara attraverso il loro canale di comunicazione quantistico.
Chiara riceve $E(|\psi\rangle\langle\psi|)$, dove E descrive l'operazione quantistica causata dall'effetto combinato di intercettazione e rumore del canale, e annuncia questo fatto pubblicamente. A questo punto le tre parti in gioco (Paola, Chiara e l'intercettatore) hanno i loro stati descritti da separate matrici densità. Si noti inoltre che finchè Paola non rivela b, l'intercettatore non ha idea di quale base avrebbe dovuto misurare per intercettare la comunicazione, egli può al massimo provare a indovinare, e nel caso in cui sbagliasse, disturberebbe lo stato ricevuto da Chiara.
Certamente però anche Chiara non sa in che base leggere i qubits non sapendo b, tuttavia anche lei ne prende atto e misura ciascun qubit nella base X o Z, a seconda del valore di una stringa di 4n bits casuali b' che ha creato per conto suo. Chiara quindi troverà una stringa a' in generale diversa da a. A questo punto Paola annuncia pubblicamente b, e

[1] possiamo pensare la base Z = $\{|0\rangle, |1\rangle\}$ come ad eventuale base su cui misurare $\hat{Z}$ (lo spin lungo l'asse $\hat{z}$), mentre X = $\{|+\rangle, |-\rangle\}$ come ad quella relativa all'operatore $\hat{X}$ (rispetto all'asse $\hat{x}$).

tramite una discussione con Chiara attraverso un canale pubblico scartano tutte quelle coppie $\{a'_i, a_i\}$ eccetto quelle per i quali i bits corrispondenti b'$_i$ e b$_i$ sono uguali.
I bits rimanenti quindi soddisfano $a' = a$, dal momento che questi bits sono quelli che Chiara ha accidentalmente misurato nella stessa base in cui Paola li aveva preparati. Si noti che b non rivela nulla su a ma è fondamentale che Paola la renda pubblica solo dopo che Chiara ha annunciato di aver ricevuto i qubits da Paola.
A questo punto mediamente Paola e Chiara hanno una stringa comune di 2n bits, ora effettuano qualche test per determinare quanto rumore o intercettazione ci siano stati durante la comunicazione. Paola sceglie casualmente n dei 2n bits e annuncia pubblicamente la selezione, quindi, sempre attraverso un canale pubblico insieme pubblicano i valori di questi n bits e ne discutono: se più di t bits sono diversi, allora è bene abortire e ricominciare il protocollo daccapo. Il valore di t è scelto in modo tale che se il test passa, allora possono applicare la procedura di *information reconciliation e privacy amplification*, per ottenere una chiave condivisa accettabilmente sicura di m bits dagli n rimanenti.

Protocollo B92

Il protocollo BB84 può essere generalizzato per usare altri stati e basi, e forniscono simili conclusioni. Infatti esiste un protocollo particolarmente semplice nel quale si usano solamente due stati: prendiamo come riferimento un singolo bit alla volta, la descrizione si generalizza poi facilmente a stringhe come nel caso del BB84.
Paola prepara un bit casuale classico a, e a seconda del suo valore invia a Chiara un qubit $|\psi\rangle = |0\rangle$ se a=0, e $|\psi\rangle = |+\rangle$ se a=1. Chiara, analogamente, a seconda del valore a' del suo bit random misura il qubit ricevuto nella base Z $|0\rangle, |1\rangle$ se a'=0, o nella base X $|+\rangle, |-\rangle$ se a'=1. Da questa misura ottiene b che assume valori 0 e 1 in corrispondenza del valore -1 o +1 degli autostati di $\hat{X}$ e $\hat{Z}$. Chiara quindi annuncia b pubblicamente, tenendo a' segreto, e attraverso una discussione pubblica tengono solo quelle coppie a',a per cui b=1 (si noti che quando a=a', b=0 sempre). Solo se a'=1-a Chiara otterà il valore b=1, e avviene con probabilità del 50%. La chiave finale, ripetuta la procedura per un numero consistente di qubits sarà a per Paola e $1 - a$ per Chiara.

Questo protocollo sottolinea come l'impossibilità di perfetta distinzione tra stati non ortogonali stia alla base della crittografia quantistica. Poichè come nel BB84 un intercettatore non riesce a fare distinzioni tra gli stati comunicati senza disturbarli, Paola e Chiara riescono nel loro intento di creare una chiave di bits e contemporaneamente limitare il rumore e l'intercettazione durante la loro comunicazione. Possono quindi applicare i protocolli di *information reconciliation e privacy amplification* per estrarre una chiave sicura a piacere dalle loro stringhe di bits casuali correlati.

Protocollo EPR

Le chiavi che emergono dai protocolli BB84 e B92 possono apparire come originati da Paola, in realtà però esse emergono da un fondamentale processo casuale che coinvolge le proprietà dell'entanglement.
Supponiamo che Paola e Chiara condividano un insieme di n coppie di qubits entangled nello stato $|+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, che abbiamo chiamato coppie EPR. Questa condivisione può avvenire in diversi modi: Paola può creare una coppia e inviarne una metà a Chiara, o viceversa, oppure una terza parte può prepararla e inviarne una metà ciascuna, o ancora potrebbero averla creata tempo addietro e conservarla gelosamente. Paola e Chiara quindi selezionano un insieme casuale di queste coppie a loro disposizione e verificano se violano la disuguaglianza di Bell, o qualche altro test appropriato. Passare il test certifica che loro continuano a possedere sufficientemente puri stati entangled, convenendo un limite alla fedeltà delle rimanenti coppie EPR. Quando misurano queste in una base casuale scelta insieme, Paola e Chiara ottengono stringhe correlate di bits classici, da cui possono ottenere una chiave di bit segreta come nei protocolli BB84 e B92.
Ma da dove derivano questi bits nel protocollo EPR? Essendo simmetrico, ed effettuando misure identiche a distanze di tipo spazio[2], non si può dire che una delle due generi la chiave, invece questa è veramente casuale.
Supponiamo che Paola prepari un bit classico casuale b, e a seconda del suo valore, misuri la sua metà della coppia EPR o nella base X, o in quella Z, ottenendo il bit a, e che faccia lo stesso Chiara a seconda del bit casuale b' misurando il bit a'. Ora confrontano i valori di b e b' su un canale classico e tengono solo le coppie a',a per cui b'=b. Questa chiave è indeterminata finchè Paola e Chiara non effettuano le misure, per cui la crittografia quantistica è pensata non come un protocollo per lo scambio di chiavi segrete, quanto piuttosto come un metodo per la loro generazione, dal momento che né Paola né Chiara possono predeterminare la chiave finale fincè non avranno portato a termine il protocollo.

Dal momento che a causa dell'inevitabile disturbo portato sullo stato comunicato, in seguito ad un guadagno di informazione da parte di un intercettatore, abbiamo motivo di credere nella sicurezza del QKD. Ciò di cui abbiamo bisogno però, per concludere che il protocollo è sicuro, è una quantificabile definizione di sicurezza che esplicitamente vincoli la conoscenza dell'intercettatore riguardo alla chiave finale, in base alle misure effettuate da Paola e Chiara.
Si ritiene accettabile il seguente criterio: un protocollo QKD è definito sicuro se per alcuni parametri positivi sicuri s e l, scelti da Paola e Chiara, e per qualsiasi strategia di intercettazione, o lo schema abortisce, o ha successo con probabilità almeno di 1-O(2^{-s}), e garantisce che l'informazione recepita dall'intercettatore sulla chiave finale sia meno di 2^{-l}.

[2]Nel senso relativistico del termine, per cui le due misure avvengono senza possibilità di influenzarsi a vicenda.

Esempio pratico sul funzionamento

Vediamo un semplice esempio pratico di quantum key distribution.
Paola inizialmente genera stati coerenti $|\alpha\rangle$ usando un diodo laser che emette luce con lunghezza d'onda $\lambda = 1.3\mu m$ e li trasmette attraverso una fibra ottica a Chiara, che ne attenua l'intensità fino ad ottenere approssimativamente un singolo fotone. Inoltre polarizza il fotone in uno dei quattro stati del protocollo BB84, prendendo come stati di polarizzazione orizzontali e verticali $|0\rangle$ e $|1\rangle$. Quindi restituisce il fotone a Paola, che lo misura attraverso un dispositivo sperimentale in una base casuale. Usando questa configurazione speciale in cui il fotone percorre la stessa strada due volte, l'apparato può essere realizzato in modo tale da autocompensare le imperfezioni lungo la fibra ottica. Paola e Chiara quindi selezionano il sottoinsieme di risultati in cui hanno usato la stessa base per orientare e misurare il qubit, e attuano il solito protocollo di *information reconciliation e privacy amplification*, comunicando sopra un canale pubblico con fotoni (attraverso la stessa fibra) con $\lambda = 1.55\mu m$. I bit della chiave possono essere scambiati ad un tasso di alcune centinaia al secondo.

La prima rete di computer che comunicassero in maniera sicura attraverso la crittografia quantistica fu realizzata nel 2004 a Cambridge in Massachussetts. La prima transazione elettronica di denaro fu realizzata precedentemente il 21 aprile dello stesso anno tra due istituti bancari austriaci utilizzando fotoni entangled per creare un codice di comunicazione indecifrabile, ma Qnet è la prima rete consistente in più di due nodi basata sulla crittografia quantistica.
Dispositivi crittografici che permettano la Quantum Key Distribution sono sul mercato[3] per scopi amministrativi e commerciali.

[3]vedi Quantum-Safe Crypto su *www.idquantique.com/quantum-safe-crypto/*

Appendice A

Algebra Lineare

Uno spazio vettoriale V[1] definito su un campo K è un insieme di elementi $|v\rangle$, detti vettori ket, chiuso rispetto all'operazione di somma tra vettori e di prodotto per scalari del campo K: questo significa che dati $|v\rangle$ e $|w\rangle$ elementi di V e a,b elementi di K, un qualsiasi $|z\rangle$ di V si può esprimere come:

$$|z\rangle = \mathrm{a}|v\rangle + \mathrm{b}|w\rangle$$

$|z\rangle$ si dice quindi combinazione lineare dei vettori $|v\rangle$ e $|w\rangle$.
Si definisce sottospazio vettoriale W, sottoinsieme di V, che sia chiuso rispetto alle stesse operazioni di somma e di prodotto per scalari di V.
Un insieme di n vettori si dice linearmente indipendente se un qualsiasi vettore ha una rappresentazione unica, cioè sono unici gli a_i con i=1,2,...,n , rispetto a quell'insieme: in pratica significa che è possibile introdurre un set di coordinate univoche che associ ad ogni vettore una n-upla di numeri $(a_1, a_2, \ldots, a_n)$.
Si definisce lo span di n vettori $|i\rangle$ lo spazio vettoriale W generato da essi, ovvero l'insieme di $|z\rangle$ che si possono esprimere come combinazione lineare degli $|i\rangle$ vettori, che si dicono generatori di W.
Una base $|i\rangle$ di n elementi di V un insieme massimale di vettori linearmente indipendenti e minimale di generatori: infatti estendendo l'insieme di vettori aggiungendone un altro, questi non saranno pi linearmente indipendenti, mentre togliendone uno qualsiasi lo span degli n-1 $|i\rangle$ non genererà pi tutto V, n viene detta dimensione dello spazio vettoriale.
La questione sarà complicata da problemi di convergenza quando si estenderà la dimensione all'infinito, tuttavia per i nostri scopi basterà rimanere a dimensioni finite.
Una qualsiasi funzione A tra due spazi vettoriali V e W generici che sia lineare nel suo argomento si dice applicazione lineare.
Quando un'applicazione lineare A:V$\longmapsto$V si dice che è un operatore lineare, in particolare fanno parte degli operatori lineare tutte quelle funzioni che collegano una base di

[1]Per questa sezione è stata seguita la trattazione di [2].

V ad un'altra per cui sono tutte equivalenti a meno dell'applicazione di una funzione lineare detto cambio di base. L'operatore che manda una base in sé stessa è banalmente l'operatore identit I.
Per le applicazioni è possibile introdurre operazioni binarie di somma e moltiplicazione per scalari proprio come per i vettori, inoltre si introduce l'operazione di moltiplicazione (o composizione) di due operatori considerandoli come applicazioni composte che agiscono su vettori in ordine da destra a sinistra.
Tuttavia a contrario della moltiplicazione tra numeri che è sempre commutativa (dalla definizione di campo) quella tra operatori non sempre lo è: infatti è non banale l'operatore [-,-] commutatore, che dati due operatori A e B:V$\longmapsto$V [A,B] che agisce su un ket $|v\rangle$ in modo che

$$[\mathrm{A,B}]|v\rangle = (\mathrm{AB - BA})|v\rangle = \mathrm{AB}|v\rangle - \mathrm{BA}|v\rangle$$

In campo complesso si definisce applicazione sesquilineare una funzione f:W$\times$V$\longmapsto$C lineare nell'argomento di destra e antilineare in quello di sinistra. Si può inoltre definire da qui un funzionale lineare da V in C fissando il primo vettore di un'applicazione sesquilineare, questo insieme di funzioni si dimostra essere uno spazio vettoriale a sua volta, detto spazio duale V^* i cui elementi denomineremo con $\langle v|$ vettori bra, ovvero dei vettori che messi in relazione con qualsiasi elemento di V danno come risultato un numero.
Il prodotto interno in uno spazio complesso di due vettori è un'applicazione sesquilineare

$$\begin{gathered} \mathrm{V}\times V \longmapsto \mathrm{K} \\ (|w\rangle, |v\rangle) \longmapsto \langle w|v\rangle \end{gathered}$$

con la proprietà che $\langle v|v\rangle \geq 0$ dove l'uguaglianza vale solo per $|v\rangle = 0$ vettore nullo, inoltre viste le proprietà dei vettori bra e ket risulta chiaramente che $\langle v|w\rangle = \langle w|v\rangle^*$.
Dati due vettori $|v\rangle$ e $|w\rangle$ appartenenti a due spazi vettoriali V e W si definisce prodotto tensoriale dei due vettori come un nuovo vettore appartenente allo spazio V$\otimes$W $|v\rangle|w\rangle$ (o anche $|v, w\rangle$) tale che l'applicazione di un bra $\langle v| \in V^*$ a $|v, w\rangle$ dà come risultato un vettore $\langle v|v, w\rangle \in W$.
Il prodotto scalare fornisce automaticamente la norma di un vettore

$$\| \, |v\rangle \, \| = \sqrt{\langle v|v\rangle}$$

che per quanto appena detto è un numero reale e positivo. È spesso utile normalizzare i vettori di base, ovvero dividere ciascuno di essi per la propria norma, in modo da avere un set di $|i\rangle$ con

$$\| \, |i\rangle \, \| = 1$$

Uno spazio vettoriale dotato di prodotto scalare si chiama spazio di Hilbert. Attraverso l'operazione di prodotto scalare possiamo introdurre il concetto di ortogonalità: diremo che due vettori non nulli sono ortogonali se il loro prodotto scalare è nullo. Risulta ora facile da intuire che generalmente la base più comoda che possa essere presa è una base ortonormale, ovvero in cui si hanno vettori ortogonali e normalizzati.
Questo processo è sempre possibile attraverso l'algoritmo di Gram-Schmidt, che data una base, ne costruisce, attraverso una combinazione lineare dei suoi argomenti, un'altra equivalente di vettori ortonormali. Da qui è automatica l'estensione del concetto di ortogonalit a sottospazi vettoriali: si dicono ortogonali due sottospazi di uno stesso spazio vettoriale V se i loro vettori sono tutto ortogonali a coppie, ossia che

$$\langle v|w\rangle = 0$$

per qualsiasi v di V e w di W.
Con la nozione di prodotto scalare possiamo allargare la classe di operatori introducendo la definizione di operatore aggiunto: dati $|v\rangle$ e $|w\rangle$ vettori e A operatore, si definisce A^+ aggiunto di A come quell'operatore che risolve

$$\langle v|Aw\rangle=\langle A^+v|w\rangle$$

In termini matriciali A^+ equivalente alla matrice trasposta e complessa coniugata di A.
Un problema strettamente collegato con la fisica è la ricerca degli autovalori e autovettori di un operatore: si definisce autovettore $|v\rangle$ di un operatore A un vettore non nullo che soddisfa la propriet

$$A|v\rangle = a|v\rangle$$

si noti che per la linearità di A se $|v\rangle$ è autovettore qualsiasi vettore $c|v\rangle$ lo è, quindi è una propriet della direzione di $|v\rangle$ mentre non ha particolare interesse il suo modulo, a invece rappresenta il fattore numerico di cui cambia $|v\rangle$ in seguito a questa operazione.
In generale il problema agli autovettori e autovalori non ammette sempre soluzione, tuttavia esistono operatori particolari che ammettono un insieme completo di autovettori, ovvero n direzioni che rimangono invariate sotto l'applicazione dell'operatore A. In questa base quindi la matrice associata ad A risulta essere in forma diagonale, avendo come elementi tutti e soli gli a_i autovalori degli autoket di base. Questi operatori particolari sono detti autoaggiunti e rispettano la proprietà che $A=A^+$: attraverso il Teorema Spettrale si dimostra che un operatore così fatto ammette sempre una base ortonormale di autoket, inoltre visto che

$$a\langle v|v\rangle=\langle v|Av\rangle=\langle A^+v|v\rangle=\langle Av|v\rangle=a^*\langle v|v\rangle$$

risulta che

$$a=a^*$$

per cui gli autovalori sono tutti reali. Si definisce inoltre autospazio V_a il sottospazio vettoriale generato da autoket di A con lo stesso autovalore a, è facile verificare che qualsiasi combinazione lineare di m autoket di autovalore a ancora un autoket di autovalore a.
La base di un operatore autoaggiunto quindi è un insieme di vettori che rispetta le proprietà di ortonormalità

$$\langle i|j\rangle{=}\delta_{ij}$$

e completezza

$$\Sigma_i|i\rangle\langle i| = \mathrm{I}$$

Da qua è facile introdurre il concetto di operatore di proiezione su un sottospazio

$$\mathrm{P}_i{=}|i\rangle\langle i|$$

infatti la sua azione su un generico ket $|v\rangle$ dà come risultato $|i\rangle\langle i|v\rangle$ dove $|i\rangle$ rappresenta la direzione del vettore proiettato e $\langle i|v\rangle$ la componente di $|v\rangle$ su $|i\rangle$
Di conseguenza si rappresenterà l'espansione del generico vettore $|v\rangle$ attraverso la relazione di completezza:

$$|v\rangle = \mathrm{I}|v\rangle = \Sigma_i \mathrm{P}_i|v\rangle = \Sigma_i|i\rangle\ \langle i|v\rangle$$

Con questa definizione risulta che qualsiasi operatore autoaggiunto A si può esprimere in seguito alla diagonalizzazione come somma di operatori di proiezione sui sottospazi generati dagli autoket di A appartenenti allo stesso autospazio moltiplicati ognuno per l'autovalore corrispondente.
Dati due operatori A e B autoaggiunti con autoket $|a\rangle$ e $|b\rangle$, ha senso chiedersi se esistano basi ortonormali comuni $|a, b\rangle$ in cui sia A che B sono in forma diagonale: risulta che questo è possibile solo se il loro commutatore è nullo, e si dice che gli operatori sono diagonalizzabili attraverso un set di autoket simultanei di A e B.
Attraverso il concetto di lunghezza di un vettore ereditata dalla definizione di prodotto scalare, si possono ora studiare un particolare gruppo di operatori detti unitari: U unitario

$$\langle Uv|Uw\rangle = \langle u|v\rangle$$

tali applicazioni non modificano la lunghezza dei vettori su cui agiscono. L'esempio classico di applicazione unitaria la rotazione attorno ad un asse, questa infatti agendo su un vettore qualsiasi al più ne modifica la direzione, senza però cambiarne la norma.

Bibliografia

[1] Buzek V., Hillery M., *Quantum copying: Beyond the no-cloning theorem*, Phys. Rev. A54 (1996) 1844

[2] Dirac P. A. M., *The Principles of Quantum Mechanics*, Clarendon Press, Oxford, 1948

[3] Ferragina P., Luccio F., *Crittografia - Principi, Algoritmi, Applicazioni*, Bollati Boringhieri Editori, Torino, 2001

[4] Nielsen M. A., Chuang I. L., *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2002

[5] Rivest R., *Handbook of Theoretical Computer Science*, Elsevier Publisher, Cambridge, 1990

[6] Simonite T., *NSA Says It Must Act Now Against the Quantum Computing Threat*, https://www.technologyreview.com/s/600715/nsa-says-it-must-act-now-against-the-quantum-computing-threat/