



UNIVERSITAT DE  
BARCELONA

Lluís Garrido  
Barcelona, 2021

# TICQ

Versió 1.0

# Contents

<b>1</b>	<b>Informació</b>	<b>3</b>
1.1	Introducció . . . . .	3
1.1.1	Informació . . . . .	3
1.1.2	Comunicació (transmissió de informació) . . . . .	4
1.1.3	Computació (processament de informació) . . . . .	6
1.2	Informació clàssica. . . . .	8
1.2.1	Informació promig i entropia . . . . .	8
1.3	Comunicació clàssica: codificació de missatges . . . . .	9
1.3.1	Codi prefix . . . . .	10
1.3.2	Primer teorema de Shannon (cota inferior per a la longitud promig en un codi prefix) . . . . .	12
1.3.3	Implementacions de codi prefix amb mínima longitud promig (codi eficient) . . . . .	13
1.3.4	Implementació d'un codi prefix ineficient. . . . .	16
1.4	Comunicació clàssica: Criptografia . . . . .	18
1.4.1	Criptografia Moderna . . . . .	18
1.4.2	Protocol RSA . . . . .	18
1.5	Comunicació clàssica: Capacitat d'un canal de comunicació amb soroll . . . . .	19
1.5.1	Entropia Condicional . . . . .	19
1.5.2	Informació Mútua . . . . .	20
1.5.3	Entropia Relativa . . . . .	21
1.5.4	Canal de comunicació . . . . .	21
1.5.5	Segon Teorema de Shannon . . . . .	23
1.5.6	Detecció i correcció d'errors . . . . .	27
1.6	Computació clàssica . . . . .	29
1.6.1	Models de computació . . . . .	30
1.6.2	Programació d'ordinadors clàssics . . . . .	34
1.6.3	Classes de complexitat computacional . . . . .	34
1.7	Informació quàntica . . . . .	37
1.7.1	Qbits . . . . .	38
1.7.2	Entrellaçament . . . . .	39
1.7.3	Desigualtats de Bell . . . . .	43
1.7.4	La base de Bell de l'espai de Hilbert de dos qbits $C^2 \otimes C^2$ . . . . .	46
1.7.5	Descomposició de Schmidt . . . . .	46
1.7.6	Matriu densitat . . . . .	48
1.7.7	L'esfera de Bloch . . . . .	57
1.7.8	Informació quàntica: entropia de Von Neumann . . . . .	59
1.7.9	Teorema de no clonació . . . . .	62
1.7.10	Teleportació quàntica de la informació . . . . .	64
1.7.11	El teorema No-hiding . . . . .	65

1.8	Comunicació quàntica: Codificació quàntica . . . . .	67
1.9	Comunicació quàntica: Criptografia quàntica . . . . .	67
1.9.1	Protocol BB84 . . . . .	68
1.9.2	Ekert 91 . . . . .	68
1.10	Comunicació quàntica: Canals quàntics . . . . .	70
1.10.1	Modelització del soroll d'un canal . . . . .	70
1.10.2	Correcció d'errors . . . . .	71
1.11	Computació quàntica . . . . .	71
1.11.1	Portes quàntiques . . . . .	74
1.11.2	Circuits quàntics . . . . .	79
1.11.3	Algorismes quàntics . . . . .	83
1.11.4	Implementació física de computadores quàntiques . . . . .	93
1.12	Perspectives de la Computació Quàntica . . . . .	99

# Chapter 1

## Informació

En aquest capítol tractarem la informació i els seus conceptes derivats, tan des del punt de vista clàssic com des del punt de vista quàntic. Primer de tot tractarem la idea d'informació i la relacionarem amb el concepte que ja coneixem d'entropia, després parlarem de com intercanviar informació entre subjectes, fet que ens portarà als conceptes de comunicació i canal de informació, i per últim parlarem de computació com a mètode de manipulació de informació per solucionar problemes.

### 1.1 Introducció

La informació és un concepte lligat a la sorpresa que ens produeixi un determinat esdeveniment, com més inesperat sigui aquest més informació sembla aportar-nos. La comunicació és la transmissió de la informació i la computació és el processament de la mateixa.

#### 1.1.1 Informació

Escollit un sistema per estudiar, podríem preguntar-nos: “Com de complex és aquest sistema?”. Dos observadors del mateix sistema poden discrepar en les seves valoracions de la complexitat. Per exemple, en un cas molt senzill, el recompte del nombre de parts pot dependre de l’escala a la qual es visualitza el fenomen (comptar àtoms és diferent del recompte de molècules, cèl·lules, òrgans, etc.). Per tant, no hauríem d’esperar que poguèssim trobar una única mesura universal de complexitat. El millor que probablement tenim és un sistema de mesura útil per un determinat observador, en un context particular i per a un propòsit particular.

Aquí, per descriure la complexitat d'un sistema, ens centrarem en mesures relacionades amb el sorprendent o inesperat que sigui una observació d'un esdeveniment. Aquest plantejament ha estat descrit com a teoria de la informació. Podríem dir que la nostra observació de la cara que ens mostra una moneda ens proporciona informació sobre el món. Desenvoluparem una definició matemàtica formal del contingut d'informació d'un esdeveniment que ocorre amb una certa probabilitat.

La informació és un concepte lligat a la sorpresa que ens produeixi un determinat esdeveniment, com més inesperat sigui aquest, més informació sembla aportar-nos (ens aporta més informació el fet que una persona mossegui un gos que el fet, molt més habitual, que un gos mossegui una persona). Per això, hom pot esperar que la informació  $I$  estigui lligada amb la probabilitat  $p$  de que es produeixi l'esdeveniment i com menys improbable sigui aquest, més informació ens aporta la seva aparició. Si l'esdeveniment es produeix amb probabilitat 1 (o 0) aleshores és nul·la la informació que ens aporta (el resultat de tirar una moneda amb dues cares no es pot aportar cap informació doncs segur que dona cara),  $I(p = 1) = 0$ .

Suposem que tenim un dau i sigui  $I_5 = I(p_5)$  la informació que obtenim quant apareix un 5 que assumim té probabilitat  $p_5$  de produir-se. Aleshores suposem que tenim un altre dau amb una  $I'_5 = I(p'_5)$ . Si ara tirem els dos daus que assumirem independents, podem demanarem per simplicitat, que la informació de que surtin dos cincs sigui la suma  $I_5 + I'_5$ . Aleshores com la probabilitat, al ser independents, es multiplica, la relació entre informació i probabilitat ha de ser (exceptuant una constant)

$$I(p) = -\ln(p) \quad (1.1)$$

que direm mesura la informació en *nats* (si agafem  $\log_2$  en lloc de  $\ln$  direm aleshores que la informació es mesura en *bits*).

Per exemple, tirar un cop una moneda justa, es a dir, amb probabilitat  $1/2$  de que surti cara o que surti creu, ens proporcionarà  $-\log_2(1/2) = 1$  bits d'informació, tan si surt cara com si surt creu. Tirar una moneda justa  $n$  vegades (o, de manera equivalent, tirar  $n$  monedes justes) ens proporciona  $-\log_2((1/2)^n) = n$  bits d'informació per a cada possible resultat dels  $2^n$  possibles.

Des del punt de vista de la Mecànica Clàssica (MC) el concepte de probabilitat només pot estar lligat al fet de desconèixer les condicions inicials del sistema doncs l'evolució clàssica és determinista (no confondre determinisme amb absoluta predictibilitat doncs aquesta darrera només es podrà fer si coneixem totes les condicions inicials). Des del punt de vista Mecànica Quàntica (MQ) l'evolució d'un estat ve donada per l'equació d'Schrödinger que també és totalment determinista, però la mesura (collapse instantani de la funció d'ona) és un fet totalment probabilístic (i no local). Per tant, la probabilitat en MQ, a part de poder anar associada al desconeixement de l'estat (estats barreja, ja sigui per preparació o per considerar només una part d'un estat total que està entrellaçat), té una component intrínseca associada al fet de mesurar. Podem dir que en ambdós casos, tan en la MC com en la MQ (sense mesura), l'evolució és determinista i per tant coneixen l'estat inicial i el Hamiltonià ( $H$ ) del sistema, podem determinar l'estat en qualsevol instant de temps posterior: la informació es manté constant (al menys fins el moment de mesurar en MQ).

La informació que rebrà un estudiant durant un curs dependrà de la sorpresa del nous conceptes (si ja s'ho sap, no hi haurà sorpresa i per tant es transmetrà poca informació cap a ell) i de la seva capacitat d'entendre els nous conceptes. Per això les probabilitats que hem d'assignar poden ser subjectives segons el problema i en qualsevol cas, sempre serà millor parlar de increment de informació que de informació absoluta. Quantificarem aquest concepte quant parlem de l'entropia de Kullback.

### 1.1.2 Comunicació (transmissió de informació)

Per transmetre informació entre subjectes, hem d'aconseguir una comunicació eficaç, els missatges s'han de transmetre correctament en entorns sorollosos. Per a un entorn sorollós (canal), hi ha un límit d'eficiència de transmissió per aconseguir una transmissió sense errors. Això s'anomena límit de Shannon, presentat l'any 1948. Tot i així, la construcció del codi que ho implementa no és pràctica en el sentit que requereix una gran quantitat de temps. Des de llavors, la comunitat de teoria de la informació ha estat estudiant la construcció de codis pràctics dissenyats per assolir el límit de Shannon.

A la figura 1.1 es mostra un sistema de comunicació aconseguit mitjançant la transmissió de codis a través d'un canal. En aquesta figura, el remitent és una estació sense fils que envia missatges i el receptor és un usuari que té un telèfon intel·ligent. El codificador converteix un missatge  $M$  en un senyal  $X$  anomenat entrada de canal. Es transmet com una ona de ràdio modulada, on inevitablement afegeix soroll a la transmissió. El descodificador converteix la sortida  $Y$  del canal una reproducció  $M'$  del missatge original. La transmissió té èxit quan es compleix  $M = M'$  i la probabilitat d'error de descodificació es defineix com la probabilitat d'esdeveniments que satisfacin  $M \neq M'$ .

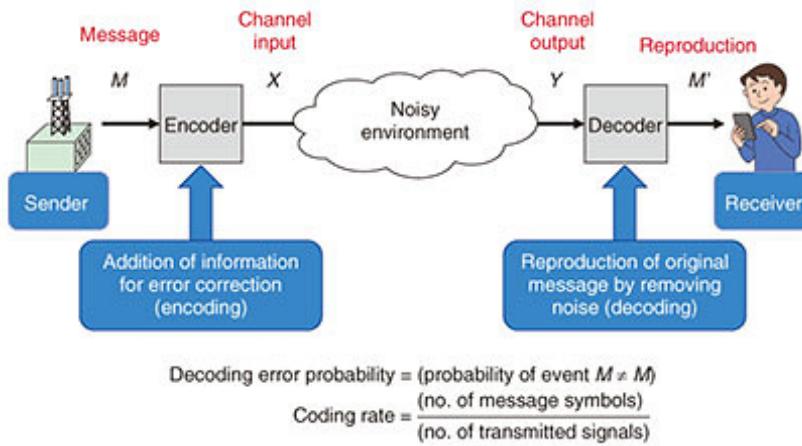


Figure 1.1: Canal típic de comunicació

La capacitat d'un canal (o eficiència de transmissió o també denominada velocitat de codificació) es defineix com el logaritme (en base dos per donar el resultat en bits) del nombre missatges distingibles que enviem a través del canal, dividit pel nombre de senyals transmesos en cada missatge.

Per exemple, si codifiquem cada missatge en 4 bits i no tenim soroll, el nombre màxim de missatges distingibles que podrem enviar serà  $2^4$  i per tant, la capacitat del canal serà  $\leq \log_2(2^4)/4 = 1$  bit (si en lloc d'enviar bits, utilitzem 4 símbols d'un codi  $d$ -nari, la cota eficiència de transmissió serà  $\leq \log_2(d^4)/4 = \log_2 d$  bits). Si hi ha soroll, la capacitat del canal serà evidentment menor, però coneugut el soroll, veurem que l'anomenat segon teorema de Shannon ens donarà un límit fonamental per aquesta capacitat. En aquests casos l'objectiu és construir un parell de codificadors i descodificadors on la probabilitat d'error de descodificació degut al soroll del canal sigui la menor possible de forma que la capacitat sigui propera al límit fonamental donat per Shannon.

En teoria de comunicació es molt comú anomenar els subjectes involucrats en una comunicació com Alice, Bob i Eve. A serà Alice el generador d'informació; B serà Bob el receptor i E serà Eve que intentarà interceptar-la:

- La font Alice (A) té una informació que vol enviar a Bob (B).
- Alice ha de codificar el missatge en un codi que els dos puguin entendre.
- Alice requerirà d'un canal físic per on comunicar la informació cap a Bob. Aquest canal generalment introduirà errors.
- És en aquest punt que Eve pot intervenir i interceptar el missatge.
- Bob haurà de descodificar el canal i seguidament el missatge.

Després de tot aquest procés s'haurà produït el fenomen de la comunicació entre Alice i Bob. En les següents seccions tractarem:

1. Com hem de codificar el missatge per a que tingui la mínima longitud? Vueren que la resposta és precisament  $l_i = \log_D(p_i)$  si utilitzem un codi D-nari.
2. Com podem evitar que Eve espii el missatge codificat? Xifrant-lo.

3. Com corregir els errors que segur ens introduceix qualsevol canal físic i que per tant no els podem evitar?. Estem buscant un parell de codificadors i descodificadors on la probabilitat d'error de descodificació sigui la menor possible de forma que la capacitat del canal sigui propera al límit fonamental donat per Shannon.

Tractarem els tres punts anteriors des del punt de vista clàssic i el segon, i parcialment el tercer des del punt de vista quàntic.

### 1.1.3 Computació (processament de informació)

En general anomenarem ordinador (del francès ordinateur) o computadora (del llatí computare, calcular) a qualsevol dispositiu físic que hem dissenyat per processar dades i convertir-les en informació útil per a nosaltres (figures 1.2 i 1.3).

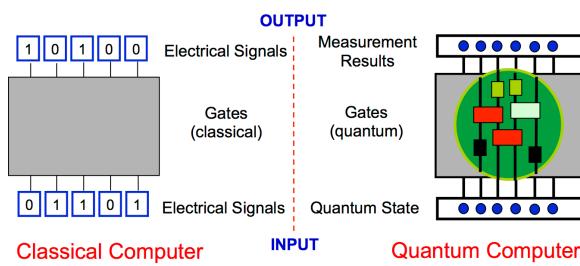


Figure 1.2: Ordinador.

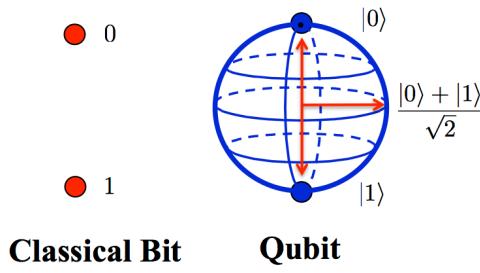


Figure 1.3: elements bàsics pel processament de la informació

Un ordinador clàssic és una màquina electrònica formada per un conjunt de circuits integrats que guarden i processen la informació. Es basen en l'àlgebra de Boole i el sistema binari, pren informació, normalment com a llista d'uns i zeros, és a dir, bits (com ara senyals elèctrics que poden estar a dos nivells de voltatge diferents) i utilitzar circuits electrònics per processar aquesta informació. Realitza un conjunt de càlculs predeterminats, que es poden desglossar en les anomenades "operacions de porta" en què l'estat d'alguns bits es modifica en funció del valor conegut d'altres bits. L'ordinador emet el resultat final, de nou com una cadena d'uns i zeros. El model en què es basen els ordinadors actuals és arquitectura de Von Neumann, és a dir, que utilitzen la memòria principal per emmagatzemar dades i instruccions alhora, característica que els permet executar programes diferents, convertint-la en una màquina de propòsit general.

Un ordinador quàntic agafarà informació codificada en un estat quàntic i després realitzarà "operacions de porta" predeterminades segons les lleis de la mecànica quàntica i produirà un nou estat quàntic que s'haurà de mesurar per poder determinar el resultat del càlcul.

La diferència clau amb l'ordinador clàssic rau en la capacitat d'un estat quàntic de representar molts estats possibles "clàssics" alhora. Mentre que un bit clàssic pot ser un "0" o un "1", un bit quàntic, o "qbit", és qualsevol combinació o "superposició" possible de "0" i "1", amb nombres complexos com a coeficients de la superposició. En la figura 1.3 podem veure que un "bit" d'informació en un ordinador clàssic es pot representar per dos punts ("0" o "1") mentre que un "qbit" es representa en un ordinador quàntic com qualsevol punt de la superfície d'un 3D esfera: l'"Esfera de Bloch".

La gran potència d'un ordinador quàntic rau en que pot processar tots els estats possibles en paral·lel. No només cada qbit pot estar en un estat de superposició, sinó que el sistema en general pot estar en una superposició de totes les combinacions d'estats diferents de tots els qbits. Aleshores el nombre d'estats possibles que poden estar presents a la superposició és enorme: si tenim  $N$  qbits, hi ha  $2^N$  estats possibles a la superposició mentre que un ordinador clàssic de  $N$  bits, només pot estar en una sola configuració. Un ordinador quàntic amb només

30 qbits tindria 1.073.741.824 estats possibles simultàniament i un ordinador quàntic amb 300 qbits tindria aproximadament el mateix nombre d'estats possibles que el nombre total d'àtoms de l'univers coneget.

Haurem de tenir en compte tres fets fonamentals en el disseny, construcció i utilització d'ordinador quàntics :

- Evolució unitària. L'evolució sempre unitària dels estats sota l'acció de qualsevol Hamiltoniana (exceptuant si decidim fer una mesura, evidentment) implica:
  - totes les "gates" quàntiques són forçosament reversibles. La unitarietat de la nostra dinàmica implica reversibilitat doncs totes les transformacions unitàries tenen una inversa .
  - no podem fer copies d'un estat quàntic arbitrari doncs veurem que cap transformació unitària ho pot fer (teorema de la no clonació).
- Existència d'estats entrellaçats. Podemaprofitar l'existència d'estats entrellaçats per processar la informació. La mesura d'una part col·lapse instantàniament l'estat de l'altra part, fet que podrem utilitzar per processar informació (però que recordem no permet enviar informació instantàniament).
- Mesura NO determinista. La mesura trenca l'evolució unitària amb un resultat probabilístic. Sempre haurem de fer una mesura per obtenir/observar el resultat. Haurem de conviure amb aquesta aleatorietat i aprofitar-la.

Aquest 3 fets anteriors fa que hi hagin dues grans dificultats amb els ordinadors quàntics: construir-ne un i determinar com programar aquest sistema .

La programació d'un ordinador quàntic es fa més difícil per les lleis de la mesura en mecànica quàntica: quan mesurem el sistema no obtenim tots els resultats possibles, sinó que la mesura col·lapsa l'estat en un resultat dels possibles. Per això, no és fàcil dissenyar algoritmes que facin ús de la potència intrínseca d'un ordinador quàntic. També són difícils d'escriure, perquè són matemàticament més complexes i molt menys intuitives que els algorismes per a un ordinador clàssic. El primer algorisme per a una computadora quàntica va ser proposat per Peter Shor el 1994 i es podia utilitzar per trobar els factors d'un nombre de longitud  $n$  utilitzant un ordre  $\log(n)$  operacions. Aquest algorisme fa ús del fet que els ordinadors quàntics serien bons per trobar el període d'una funció periòdica, que es pot relacionar per la teoria de nombres amb el problema de factoritzar un nombre. Per fer una comparació amb els ordinadors clàssics, si suposem que un superordinador ràpid trigaria aproximadament un any a factoritzar un nombre de 150 díigits, el mateix ordinador requeriria aproximadament la vida útil de l'univers per factoritzar un nombre de 400 díigits. L'acceptació que aquests són els ordres de temps necessaris per factoritzar un número en els ordinadors clàssics, constitueix la base dels sistemes de xifratge actuals. Si es poguéss construir un ordinador quàntic que factoritzés un nombre de 150 díigits en un mes, llavors el mateix ordinador quàntic podria facturar un nombre de 400 díigits en aproximadament un any. Per descomptat, això podria ser una mala notícia per als esquemes de xifratge actuals, però la ciència de la informació quàntica també proporciona nous esquemes de xifratge de "substitució" per superar aquest problema. Des de l'algorisme de Shor, s'han desenvolupat molts algoritmes nous, que donen rapideses significatives (encara que no sempre exponencials) per a diferents problemes. No obstant això, el desenvolupament d'aquests algoritmes encara és un camp en creixement.

Construir un ordinador quàntic és un repte, ja que implica manipular un gran nombre d'objectes microscòpics segons les lleis de la mecànica quàntica. La principal dificultat és que l'objecte microscòpic sobre el qual emmagatzemem cada qbit ha d'estar aïllat del seu entorn tant com sigui possible per evitar la descoherència i que la informació passi a l'entorn (en un

sistema tancat la informació quàntica es manté i pot passar d'una part del sistema a l'altre però no queda el entrellaçament de les parts. Això és diferent de la informació clàssica on per exemple la informació d'un missatge xifrat no està ni en el missatge ni en la clau per xifrar, si no en les correlacions). Aquest transvasament d'informació quàntica entre l'ordinador quàntic i el seu entorn, també es pot veure com el soroll de l'entorn que randomitza l'estat del qbit.

L'exemple d'ordinador quàntic que tractarem amb més detall serà el dels Ions atrapats: en aquests sistemes, les cadenes d'ions s'emmagatzemen en trampes electromagnètiques i la informació es codifica en estats electrònics de llarga vida dels àtoms.

Ja el 1982 Feynman va discutir la construcció d'una màquina que funcionaria segons principis de mecànica quàntica per explorar altres efectes quàntics i fer simulacions. Però no va ser fins al 1994 quant Peter Shor va proposar el seu algoritme de factorització que el camp va veure un creixement exponencial i la raó pot trobar-se en el interès militar per desxifrar codis encriptats.

## 1.2 Informació clàssica.

Hem vist a la introducció que demanar que la informació sigui un concepte lligat a la sorpresa que ens produeixi un determinat esdeveniment (com més inesperat sigui aquest més informació sembla aportar-nos, es a dir, I ha de ser funció de la probabilitat) i el fet de demanar que la informació de dos esdeveniments independents sigui la suma de les informacions de cada un d'ells, aleshores hem deduït que la relació entre informació i probabilitat ha de ser (exceptuant una constant)

$$I(p) = -\ln(p) \quad (1.2)$$

que direm mesura la informació en *nats* (si agafem  $\log_2$  en llog de  $\ln$  direm aleshores que la informació es mesura en *bits*).

Per exemple, tirar un cop una moneda justa, es a dir, amb probabilitat  $1/2$  de que surti cara o que surti creu, ens proporcionarà  $-\log_2(1/2) = 1$  bits d'informació, tan si surt cara com si surt creu. Tirar una moneda justa  $n$  vegades (o, de manera equivalent, tirar  $n$  monedes justes) ens proporciona  $-\log_2((1/2)^n) = n$  bits d'informació per a cada possible resultat dels  $2^n$  possibles.

### 1.2.1 Informació promig i entropia

La informació promig quan tenim  $n$  possibles resultats d'un esdeveniment, on cadascun un d'ells té una certa probabilitat  $p_i$  de que es produeixi, serà

$$\bar{I} = -\sum_{i=1}^n p_i \log_2(p_i) \quad (1.3)$$

En el cas de tirar una moneda justa, hem vist que el fet de que surti cara ens aporta 1 bit de informació i el fet de que surti creu també un bit, per tant, la informació promig serà 1 bit (efectivament  $\bar{I} = 2 [ (1/2) \log_2(1/2) ] = 1$ ). En el cas de tirar  $n$  cops una moneda justa, hem vist que cada possible resultat ens aporta  $n$  bits de informació, per tant, la informació promig serà  $n$  bits (efectivament, tenim  $2^n$  resultats possibles, cada un amb la mateixa  $p = (1/2)^n$ , i per tant  $\bar{I} = -2^n [ (1/2)^n \log_2(1/2)^n ] = n$ ).

Si estem davant d'una col·lectivitat on la probabilitat de totes les  $n$  possibilitats són zero menys una determinada, que evidentment té  $p = 1$ , aleshores  $\bar{I} = 0$ . Per altra banda, si totes són equiprobables ( $p_i = 1/n$ ) aleshores

$$\bar{I} = -\sum_{i=1}^n p_i \log_2 p_i = -\sum_{i=1}^n \frac{1}{n} \log_2 \left( \frac{1}{n} \right) = \log_2(n) \quad (1.4)$$

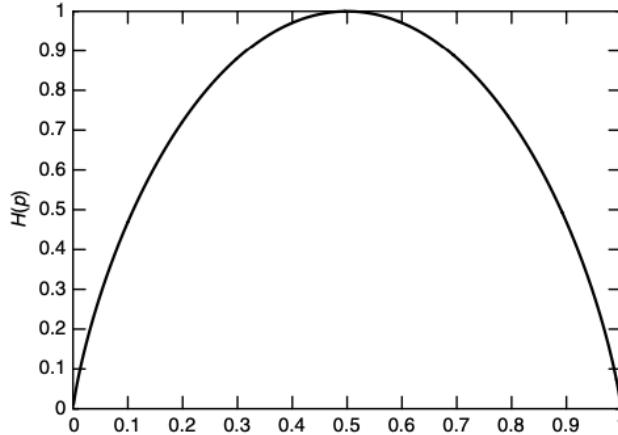


Figure 1.4: entropia per un sistema de dos possibles estats

Veiem que aquesta informació esperada (també anomenada informació de Shannon) dona també una idea del desordre esperat (en el primer cas on  $\bar{I} = 0$  no hi ha desordre ni informació esperats doncs el resultat sempre és el mateix, mentre que en el segon cas el desordre, com veurem, és màxim al igual que informació esperada). De fet, recordem que definim entropia per la col·lectivitat microcanònica com  $H(n) = \ln n$  o la seva generalització de Gibbs com  $H(\vec{p}) = -\sum_{i=1}^n p_i \ln p_i = \bar{I}$ . De fet sabem que  $0 \geq H(\vec{p}) \geq \ln(n)$  i per tan l'entropia, el desordre, és màxim quan totes les  $n$  probabilitats són iguals  $p_i = 1/n$ . En el cas  $n = 2$ , tindrem dues probabilitats  $p_0 \equiv p, p_1 = 1 - p$  i en la figura 1.4 podem veure (en base 2)

$$H(p_0, p_1) = -p \log_2 p - (1-p) \log_2(1-p) \quad (1.5)$$

en funció de  $p$ . Veiem, com esperaríem, que el màxim es troba a  $p = 1/2$ , o sigui  $p_0 = p_1 = 1/2$ , com era d'esperar.

Un exemple que il·lustra aquest resultat: quanta informació pot obtenir un estudiant d'una nota? Primer, la informació màxima es produeix si totes les notes tenen igual probabilitat (per exemple, en una classe de superació / fallada, la meitat hauria de passar si volem maximitzar la informació que dóna la nota). La informació màxima que l'alumne obté d'una nota serà:

- *Passar/fallar*: 1 bit.
- *A, B, C, D, F*: 2.3 bits.
- *A, A-, B+, ..., D-, F*: 3.16 bits.
- *0, 1, 2, ..., 10*: 3.46 bits

Així, utilitzar  $+/-$  en la classificació proporciona als estudiants quasi 1 bit més d'informació per nota que sense  $+/-$ , i uns 2.2 bits per nota més que no passen / fallen.

### 1.3 Comunicació clàssica: codificació de missatges

En aquesta secció ens basarem en les idees de Shannon per determinar quin és el millor codi (aquí millor voldrà dir el més curt) per enviar missatges on coneixem les probabilitats en que apareixen els diferents símbols utilitzats en els mateixos (per exemple les lletres de l'alfabet en llengua catalana).

Suposem que tenim un alfabet  $\chi$  discret que utilitzem per enviar missatges. Un codi  $s$ -nari (normalment utilitzarem el binari) és una aplicació dels símbols de l'alfabet  $\chi$  a un conjunt de paraules-codi  $s$ -naries que ens permeti codificar i descodificar de forma única qualsevol missatge.

**Exemple.** Sigui  $\chi = \{a, b, c\}$ , un codi binari del mateix pot ser

$$\begin{aligned} a &\rightarrow 00 \\ b &\rightarrow 10 \\ c &\rightarrow 11 \end{aligned} \tag{1.6}$$

**Exemple.** Sigui  $\chi = \{a, b, c\}$ , un codi binari del mateix pot ser

$$\begin{aligned} a &\rightarrow 0 \\ b &\rightarrow 10 \\ c &\rightarrow 11 \end{aligned} \tag{1.7}$$

**Exemple.** Sigui  $\chi = \{a, b, c\}$ , un codi binari del mateix pot ser

$$\begin{aligned} a &\rightarrow 0 \\ b &\rightarrow 01 \\ c &\rightarrow 10 \end{aligned} \tag{1.8}$$

Notem que es aquest darrer exemple quant estiguem descodificant i aparegui un 0 no podem dir si estem davant d'una  $a$  fins que veiem si el següent és un 0 o un 1. A més a més la descodificació no serà única, per exemple 0010 tan pot ser descodificat com  $aac$  com per  $aba$ . Per evitar això introduïm el concepte de codi prefix.

### 1.3.1 Codi prefix

Anomenarem "codi prefix" aquell codi, generalment de longitud variable, amb la propietat de "prefix", es a dir que cap paraula del codi es prefix de qualsevol altra paraula-codi (el darrer exemple no era "codi prefix" doncs el paraula-codi associada al símbol  $a$ , el 0, és prefix del paraula-codi associada al símbol  $b$ , el 01). Una forma de generar codi prefix el tenim en la figura 1.5 on les línies discontinues representen línies mortes, mentre que les contínues designen codis, en aquest cas  $\{0, 10, 110, 111\}$  (evidentment el podríem continuar desdoblat 111 a 1110 i 1111, i així successivament)

Designem per  $L(x)$  la longitud de la paraula-codi associada al símbol  $x$  d'un cert alfabet  $\chi$  (en l'exemple de 1.7 tenim  $L(a) = 1, L(b) = 2, L(c) = 2$ ). Aleshores, si utilitzen un codi prefix  $r$ -nari, es compleix

$$\sum_{x \in \chi} r^{-L(x)} \leq 1 \tag{1.9}$$

Anem a demostra-ho. Suposem que tenim  $s$  símbols en un cert alfabet  $\chi$  i volem trobar un codi prefix  $r - nari$  del mateix. Sense pèrdua de generalització, les longituds  $L(x)$  de les  $s$  paraules-codi estaran ordenades

$$n_1 \leq n_2 \leq \dots \leq n_s \tag{1.10}$$

Anem a construir el codi prefix en ordre creixent  $i = 1, \dots, s$ . Existirà un codi prefix si en cada pas  $j$  podem trobar una paraula-codi que no conté cap de les  $j - 1$  paraules-codi anteriors com a prefix. En un principi, per escollir la paraula codi  $j$  tindríem  $r^{n_j}$  possibilitats però degut a que, per exemple, la paraula-codi  $k < j$  no pot ser prefix per a  $j$ , tenim  $r^{n_j - n_k}$  possibilitats prohibides (són totes aquelles que tenen el mateix prefix donat per la paraula-codi  $k$ ). A més a més els conjunts prohibits per a cada paraula-codi amb  $k < j$  són excloents (si dos paraules-codi

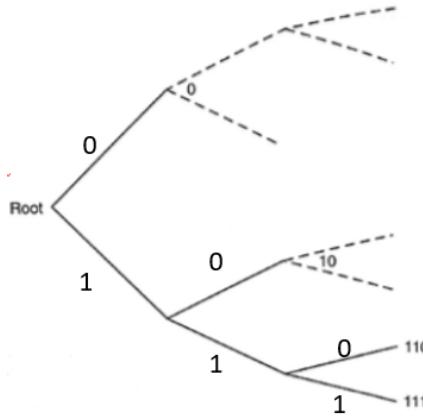


Figure 1.5: Esquema per generar un exemple de codi prefix

amb  $k < j$  prohibeixen una mateixa combinació en el pas  $j$  voldria dir que la més petita de les dues paraules-codi seria prefix de l'altra, fet que contradiu la nostre hipòtesis) i per tant el nombre de prohibicions total pel pas  $j$  és

$$\sum_{i=1}^{j-1} r^{n_j - n_i} \quad (1.11)$$

Aleshores per poder construir un codi prefix, en cada pas  $j$  només podrem escollir una paraula-codi si

$$r^{n_j} > \sum_{i=1}^{j-1} r^{n_j - n_i}, \quad \forall j = 2, \dots, s \quad (1.12)$$

Com les relacions anteriors són enters podem escriure

$$r^{n_j} \geq \sum_{i=1}^{j-1} r^{n_j - n_i} + 1 = \sum_{i=1}^j r^{n_j - n_i}, \quad \forall j = 1, \dots, s \quad (1.13)$$

i dividint ara per  $r^{n_j}$  obtenim

$$1 \geq \sum_{i=1}^j r^{-n_i}, \quad \forall j = 1, \dots, s \quad (1.14)$$

i en particular per  $j = s$  obtenim la desigualtat buscada

$$\sum_{i=1}^s r^{-n_i} \leq 1 \quad (1.15)$$

que es coneix com la desigualtat de Kraft. Com l'argumentari va en les dues direccions la desigualtat anterior és necessària i suficient per poder trobar un codi prefix. Evidentment de tots els codis prefix que podem associar a un cert alfabet, es convenient escollir un que utilitzi el mínim de longitud per enviar un missatge. Això ho tractarem en les properes subseccions. A més a més a partir d'ara, per simplicitat, només considerarem codis binaris però l'estensió a codis  $r$ -narís és trivial (substituint el 2 associat a la base binaria per  $r$ ).

### 1.3.2 Primer teorema de Shannon (cota inferior per a la longitud promig en un codi prefix)

Sigui un alfabet  $\chi$  on els símbols  $x \in \chi$  apareixen amb probabilitats  $p(x)$ . Donat un codi (binari) prefix amb longituds  $L(x)$  (per la paraula-codi associada al símbol  $x$ ), el valor promig de la longitud de les paraules-codi serà

$$\bar{L} = E_p(L) = \sum_{x \in \chi} p(x)L(x) \quad (1.16)$$

Demostrarem que aquesta longitud està acotada inferiorment de la forma

$$\bar{L} \geq -\sum_{x \in \chi} p(x) \log_2(p(x)) \equiv H(p) \quad (1.17)$$

on la igualtat es satisfà quant hem escollit un codi tal que  $L(x) = -\log_2(p(x))$  i  $H(p)$  és l'entropia de l'alfabet donat segons la seva probabilitat  $p(x)$ .

De la desigualtat de Kraft obtenim

$$C_L = \sum_{x \in \chi} 2^{-L(x)} \leq 1 \quad (1.18)$$

i definim  $Q(x) \equiv 2^{-L(x)}/C_L$  que la podem interpretar com una distribució de probabilitat doncs  $\sum_{x \in \chi} Q(x) = 1$ . Aleshores, com  $L(x) = -\log_2(Q(x)C_L)$ , tenim

$$\bar{L} = E_p(L(x)) = -E_p(\log_2 Q(x)) - E_p(\log_2 C_L) \geq -E_p(\log(Q(x))) \quad (1.19)$$

doncs  $-\log_2 C_L \geq 0$  i per tant

$$E_p(L) - H(p) = E_p(L(x)) + E_p(\log_2(p(x))) \geq E_p(\log_2(p(x))) - E_p(\log(Q(x))) = -E_p(\log_2 \left( \frac{Q(x)}{p(x)} \right)) \quad (1.20)$$

Tenint en compte que  $\log$  és una funció concava aleshores es compleix la desigualtat de Jensen  $\log(E_p(Y)) \geq E_p(\log(Y))$  (veure figura 1.6, com exemple pensem en una distribució plana per a Y on la desigualtat és evident) i per tant, com  $-E_p(\log(Y)) \geq -\log(E_p(Y))$ , tenim

$$E_p(L) - H(p) \geq -\log_2 E_p \left( \frac{Q(x)}{p(x)} \right) = -\log_2 1 = 0 \quad (1.21)$$

i per tant trobem la cota buscada

$$\bar{L} \geq H(p) = -\sum_{x \in \chi} p(x) \log_2(p(x)) \quad (1.22)$$

per a qualsevol codi prefix que escollim. Només quant escollim  $L(x) = -\log_2(p(x))$  tindrem la igualtat i estarem escollint el codi prefix que minimitza els bits necessaris per enviar missatges. Evidentment veiem que els símbols més freqüents estaran representats per paraules-codi més curtes.

També podem obtenir el mateix resultat amb multiplicadors de Lagrange: és un problema de minimitzar  $\bar{L}$  en el que utilitzarem els multiplicadors de Lagrange per posar com a restricció la desigualtat de Kraft (ho farem en general, es a dir per un codi D-nari)

$$J = \bar{L} + \lambda(C_L - 1) = \sum_{x \in \chi} p(x)L(x) + \lambda \left( \sum_{x \in \chi} 2^{-L(x)} - 1 \right) \quad (1.23)$$

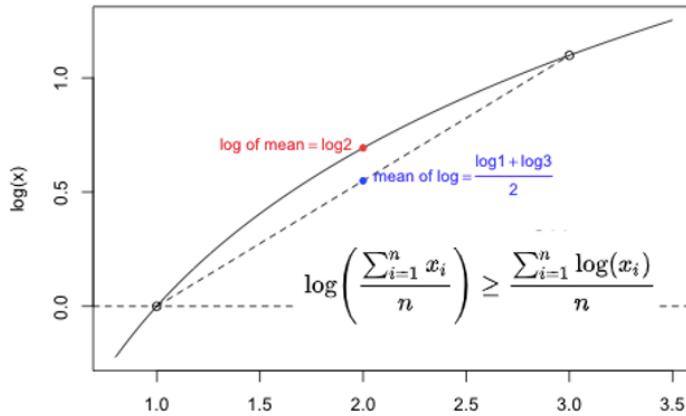


Figure 1.6: Desigualtat de Jensen

Si derivem respecte de la variable  $L(x)$  i igualem a 0 el resultat obtenim:  $D^{-L(x)} = p(x)/(\lambda \ln D)$  que sumant ens porta a

$$\sum_{x \in \chi} D^{-L(x)} = \frac{1}{\lambda \ln D} \sum_{x \in \chi} p(x) \quad (1.24)$$

com  $\sum_{x \in \chi} D^{-L(x)} = 1$  tenim que  $1 = \frac{1}{\lambda \ln D}$  i que ens porta a la solució òptima

$$D^{-L(x)} = p(x) \Rightarrow L(x) = -\log_D p(x) \quad (1.25)$$

i recuperem el resultat òptim ja conegit.

### 1.3.3 Implementacions de codi prefix amb mínima longitud promig (codi eficient)

Hem vist que una implementació de codi prefix amb mínima longitud promig s'aconsegueix quant

$$L^*(x) = [-\log_2(p(x))] \quad (1.26)$$

on per  $[ ]$  volem indicar l'enter més petit igual o major a  $-\log_2(p(x))$  (doncs  $L^*(x)$  ha de ser un nombre enter). Com

$$\sum_{x \in \chi} 2^{-L^*(x)} = \sum_{x \in \chi} 2^{-[-\log_2(p(x))]} \leq \sum_{x \in \chi} 2^{\log_2(p(x))} = 1 \quad (1.27)$$

i com es compleix la desigualtat de Kraft, estem davant un codi prefix. En aquest cas, com l'aproximació anterior de fer  $[ ]$  introduceix un error de com a màxim un bit, tenim que

$$H(p) \leq \bar{L}^* \leq H(p) + 1 \quad (1.28)$$

#### Algorisme de Shannon

Ara donarem un algoritme per trobar un d'aquest codi prefix amb mínima longitud promig. Suposem un alfabet  $\chi$  on els seus símbols tenen probabilitats  $p(x_1) \geq p(x_2) \geq \dots \geq p(x_s)$  i definim la funció acumulativa  $F(x_i) = \sum_{j=1}^{i-1} p(x_j)$ . Una implementació és l'anomenat codi de Shannon:

$$x_i \rightarrow F_i \text{ arrodonit a } [-\log_2(p(x_i))] \text{ bits} \quad (1.29)$$

Table 1.1: Donat l'alfabet  $\chi = \{a, b, c, d\}$  amb les seves probabilitats  $p(x)$ , aquesta taula mostra el codi prefix més eficient (en aquest exemple  $C_L$  val exactament 1 degut a que  $[-\log_2(p(x))]$  són enters, sinó seria lleugerament inferior)

$x$	$p(x)$	$L^*(x) = [-\log_2(p(x))]$	$2^{-L^*(x)}$	F(x)	codi	$p(x)L^*(x)$
a	0.5	1	0.5	0	0	0.5
b	0.25	2	0.25	0.5	10	0.5
c	0.125	3	0.125	0.75	110	0.375
d	0.125	3	0.125	0.875	111	0.375
-	$\sum_x p(x) = 1$	-	$C_L = \sum_x 2^{-L^*(x)} = 1$	-	-	$\bar{L}^* = \sum_x p(x)L^*(x) = 1.75$

(evident per aquest codi  $L^*(x) = [-\log_2(p(x))]$ )

**Exemple.** Sigui l'alfabet  $\chi = \{x_1, x_2, x_3\}$  on sabem que aquest símbols surten en els missatges amb probabilitats (ja ordenades de forma decreixent)  $p(x_1) = 11/20, p(x_2) = 1/4, p(x_3) = 1/5$ . La seva entropia val  $H(p) = \sum_{i=1}^3 p_i \log_2(p_i) = 1.429$  i per tant esperem que la longitud promig estigui entre  $1.429 \leq \bar{L}^* \leq 2.429$ . Aleshores

$$\begin{aligned} F(x_1) &= 0, \quad F(x_2) = 11/20, \quad F(x_3) = 11/20 + 1/4 = 4/5 \\ [-\log_2(p1)] &= [0.86] = 1, \quad [-\log_2(p2)] = [2] = 2, \quad [-\log_2(p3)] = [2.3] = 3 \end{aligned} \tag{1.30}$$

i per tant

$$\begin{aligned} x_1 &\rightarrow 0 \\ x_2 &\rightarrow 10 \\ x_3 &\rightarrow 110 \end{aligned} \tag{1.31}$$

(recordem que en binari  $1/2$  ve representat per 0.1) doncs per exemple  $F(x_2) = 11/20 = 1/2 + 1/20$  es representa fins a dos bits com 0.10 i  $F(x_3) = 4/5 = 1/2 + 1/4 + 1/20$  es representa fins a tres bits, com 0.110.

La longitud promig de les paraules-codi escollides és

$$\bar{L}^* = (11/20) \times 1 + (1/4) \times 2 + (1/5) \times 3 = 1.6 \text{ bits} \tag{1.32}$$

i que queda dins el marge esperat de  $[H(p), H(p) + 1]$ .

**Exemple.** Un altre exemple de construcció de codi prefix eficient el tenim en la taula 1.1 on el nombre de bits promig de les paraules-codi és 1.75. Qualsevol altra codi prefix del mateix alfabet tindrà un nombre de bits promig superior, tal i com veurem en la següent secció.

### Algorisme de Huffman

L'algorisme Huffman12 (1951) porta el nom del descobridor de l'algorisme que ens permet construir un codi prefix òptim per a un alfabet qualsevol .

Supossem l'exemple de la figura 1.7 on tenim un text en el que només utilitzem cinc símbols i on cada un apareix amb la freqüència donada. Si assignem arbitràriament un codi de 3 bits a cada un dels símbols, la longitud total del text, un cop passat a codi binari, serà  $L_T = 34 \times 3 + 11 \times 3 + 26 \times 3 + 12 \times 3 + 4 \times 3 = 261$ . L'algorisme Huffman12 ens permet codificar el missatge en menys bits.

Símbol	Codi	Distància (l)	Freq.	Bits totals
A	000	3	34	102
S	001		11	33
E	010		26	78
T	100		12	36
P	101		4	12

Figure 1.7: Taula del nostre text exemple.

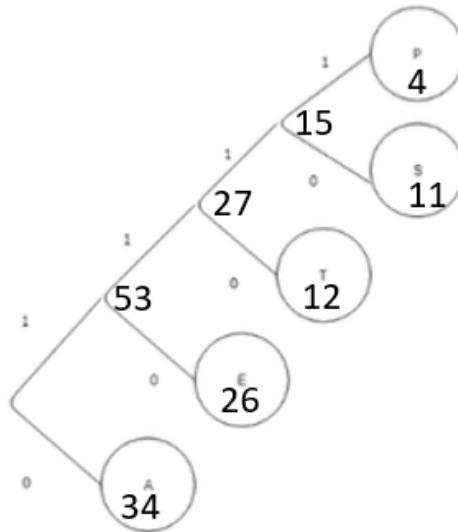


Figure 1.8: Algorisme Huffman

L'algorisme Huffman fa ús d'un estructura d'arbre com el de figura 1.8 on els símbols del nostre alfabet van a parar a les fulles i les branques ens determinen el codi que hem d'associar. El seu funcionament és el següent:

1. Identifiquem els dos símbols (dues fulles) que tinguin menys freqüència en el nostre text. Connectem les seves dues fulles a un node al que assignem la freqüència suma i ens oblidem dels dos símbols.
2. D'entre la resta de símbols i nodes que ens queden, identifiquem els dos que tinguin menys freqüència, els agrupem gràficament en un nou node al que assignem la freqüència suma i ens oblidem de les dues entitats que han participat en la seva definició.
3. Repetint el pas anterior fins que no ens quedi cap símbol, anirem construint una estructura d'arbre on els símbols utilitzats els anem posant en fulles i els nodes van creant l'estruatura en branques
4. Un cop creat l'arbre assignem 0 i 1 a cada branca que surti de cada node i la lectura de la seqüència des del node principal fins arribar al symbol corresponent, ens determina el codeword associat.

En l'exemple de la figura 1.7 identifiquem P i S com símbols amb menys freqüència amb 4 i 11 respectivament. Aleshores, agrupem les dues fulles P I S en un node al que assignem la freqüència 15. De les freqüències que ara ens queden (34, 26, 12 i 15) les dues més baixes són 15 (que és el node que hem creat abans) i 12 que és la T i els agrupem en un nou node. Si fem això fins a quedar-nos sense símbols, obtindrem l'arbre de la figura 1.8. Assignant 0 i 1 a cada branca que surti de cada node, podem veure els codewords associats : A = 0, E = 10, T = 110, S = 1110 i P = 1111. Si ara calculem el nombre de bits utilitzats per aquest text tenint en compte les noves longituds:  $34 + 52 + 36 + 44 + 16 = 182$ . Veiem que ens hem estalviat 79 bits.

### Algorisme addicional de Lempel-Ziv

Un cops utilitzats un dels dos algoritmes anteriors sobre un fitxer donat, s'ha creat un fitxer binari de longitud promig òptima. No obstant, pot passar i segur que passa, que en aquest nou

fitxer binari hi hagin estructures binaries de diferent longitud que es vagin repetint. Aquesta repetició és utilitzada per l'algorisme Lempel-Ziv per reduir la mida del fitxer.

L'algorisme Lempel-Ziv és també un codi compressor lossless com Huffman però utilitza un altre mètode per obtenir un arxiu amb la mínima quantitat de bits possible. El que fa és anar creant un diccionari de caràcters vistos anteriorment. Amb aquest mètode el que passa és que si estem codificant un text, començarem codificant estructures petites (lletres) però a mida que es vagi creant el diccionari acabarem emmagatzemant paraules o inclús frases.

Aquest algorisme es el que s'utilitza en compressor com el zip o el gzip.

### 1.3.4 Implementació d'un codi prefix ineficient.

En la secció anterior em vist que el codi prefix més eficient (el de mínima longitud promig de les paraules-codi) és aquell on  $L^*(x) = [-\log_2(p(x))]$ . Si utilitzem qualsevol altre codi prefix amb diferents longituds de les anteriors,  $L(x) \equiv -\log_2(Q(x)C_L)$  (on  $C_L = \sum_{x \in \chi} 2^{-L(x)} \leq 1$ ) i que defineix  $Q(x)$  com una densitat de probabilitat sabem de 1.21 que

$$\bar{L} - \bar{L}^* \geq \sum_{x \in \chi} p(x) \log_2 \left( \frac{p(x)}{Q(x)} \right) \geq 0 \quad (1.33)$$

on hem canviat  $H$  per  $\bar{L}^*$  (com  $H \leq \bar{L}^* \leq H+1$  la desigualtat sembla que s'hauria de complir-se amb  $\geq -1$  però com per a  $L$  també hem d'arrodonir a un nombre de bits enter, finalment ens queda la desigualtat donada). Per tant, necessitarem en promig més bits per les paraules-codi, exactament necessitarem  $-\sum_{x \in \chi} p(x) \log_2(Q(x)C_L) + \sum_{x \in \chi} p(x) \log_2(p(x))$  bits però una cota inferior la podem obtenir de

$$\Delta \bar{bits} \geq \sum_{x \in \chi} p(x) \log_2 \left( \frac{p(x)}{Q(x)} \right) \geq 0 \quad (1.34)$$

També podem procedir assumint una distribució fictícia  $Q(x)$  pels nostres símbols de l'alfabet i determinar un codi prefix amb longituds  $L(x) = [-\log_2(Q(x))]$ . En aquest cas l'anterior cota inferior serà directament una igualtat (sempre positiva)

$$\begin{aligned} \Delta \bar{bits} &= -\sum_{x \in \chi} p(x) \log_2(Q(x)) + \sum_{x \in \chi} p(x) \log_2(p(x)) = \sum_{x \in \chi} p(x) \log_2 \left( \frac{p(x)}{Q(x)} \right) \\ &= -E_p \left( \log_2 \left( \frac{Q(x)}{p(x)} \right) \right) \geq \log_2 \left( E_p \left( \frac{Q(x)}{p(x)} \right) \right) = 0 \end{aligned} \quad (1.35)$$

En la taula 1.2 hem assumit unes probabilitats  $Q(x)$  fictícies diferents a les reals  $p(x)$  donades en la taula 1.1. El resultat és que utilitzant  $Q(x)$  per definir el codi promig, utilitzarem (en promig) dos bits per codificar l'alfabet mentre que el codi prefix eficient donat en la taula 1.1 necessita només 1.75 bits. Tenim un  $\Delta \bar{bits} = -\sum_{x \in \chi} p(x) \log_2(Q(x)) + \sum_{x \in \chi} p(x) \log_2(p(x)) = 2 - 1.75 = 0.25$ .

En la taula 1.3 hem assumit un codi prefix donat per les paraules-codi de la segona columna en lloc del codi eficient donat en la taula 1.1. El resultat és que utilitzant les  $L(x)$  donades, utilitzarem (en promig) 3 bits per codificar l'alfabet quan el codi prefix eficient donat en la taula 1.1 necessita només 1.75 bits. Tenim un

$$\begin{aligned} \Delta \bar{bits} &= -\sum_{x \in \chi} p(x) \log_2(Q(x)C_L) + \sum_{x \in \chi} p(x) \log_2(p(x)) \\ &= -\sum_{x \in \chi} p(x)(\log_2(Q(x)) + \log_2(C_L)) + \sum_{x \in \chi} p(x) \log_2(p(x)) = (2 + 1) - 1.75 = 0.25 \end{aligned} \quad (1.36)$$

Table 1.2: Donat l'alfabet  $\chi = \{a, b, c, d\}$  amb les probabilitats  $p(x)$  donades en la taula 1.1, aquesta taula dona el codi prefix ineficient obtingut a l'assumir unes probabilitats fícties  $Q(x)$

$x$	$Q(x)$	$L(x) = [-\log_2(Q(x))]$	$2^{-L(x)}$	F(x)	codi	$p(x)L(x)$
a	0.25	2	0.25	0.5	10	1
b	0.5	1	0.5	0	0	0.25
c	0.125	3	0.125	0.75	110	0.375
d	0.125	3	0.125	0.875	111	0.375
-	$\sum_x Q(x) = 1$	-	$C_L = \sum_x 2^{-L(x)} = 1$	-	-	$\bar{L} = \sum_x p(x)L(x) = 2$

Table 1.3: Donat l'alfabet  $\chi = \{a, b, c, d\}$  amb les probabilitats  $p(x)$  donades en la taula 1.1, aquesta taula demostra que si escollim per codificar les paraules-codi de la segona columna, el codi prefix que estem definint és ineficient

$x$	codi	$L(x)$	$2^{-L(x)}$	$Q(x) = 2^{-L(x)}/C_L$	$p(x)L(x)$
a	000	3	0.13	0.25	1.5
b	100	3	0.13	0.25	0.75
c	110	3	0.13	0.25	0.375
d	111	3	0.13	0.25	0.375
-	-	-	$C_L = \sum_x 2^{-L(x)} = 0.5$	$\sum_x Q(x) = 1$	$\bar{L} = \sum_x p(x)L(x) = 3$

Finalment en la taula 1.4 hem assumit un codi prefix donat per les paraules-codi de la segona columna en lloc del codi eficient donat en la taula 1.1 que necessita només 1.75 bits. El resultat és que utilitzant les  $L(x)$  donades, utilitzarem (en promig) 2 bits per codificar l'alfabet quant el codi prefix eficient donat en la taula 1.1 necessita només 1.75 bits. Tenim un (com en aquest cas  $C_L = 1$ )

$$\begin{aligned} \Delta \bar{bits} &= - \sum_{x \in \chi} p(x) \log_2(Q(x)C_L) + \sum_{x \in \chi} p(x) \log_2(p(x)) \\ &= - \sum_{x \in \chi} p(x) \log_2(Q(x)) + \sum_{x \in \chi} p(x) \log_2(p(x)) = 2 - 1.75 = 0.25 \end{aligned} \quad (1.37)$$

Table 1.4: Donat l'alfabet  $\chi = \{a, b, c, d\}$  amb les probabilitats  $p(x)$  donades en la taula 1.1, aquesta taula demostra que si escollim per codificar les paraules-codi de la segona columna, el codi prefix que estem definint és ineficient

$x$	codi	$L(x)$	$2^{-L(x)}$	$Q(x) = 2^{-L(x)}/C_L$	$p(x)L(x)$
a	00	2	0.25	0.25	1
b	01	2	0.25	0.25	0.5
c	10	2	0.25	0.25	0.25
d	11	2	0.25	0.25	0.25
-	-	-	$C_L = \sum_x 2^{-L(x)} = 1$	$\sum_x Q(x) = 1$	$\bar{L} = \sum_x p(x)L(x) = 2$

## 1.4 Comunicació clàssica: Criptografia

La criptografia (clàssica) té com objectiu que Eve, encara que pugui fer-se amb una copia del missatge que Alice envia a Bob, no pugui desxifrar el contingut. La idea es que Alice i Bob comparteixen una clau/mètode per desxifrar el missatge. Evidentment aquest mètode crea un problema: com distribuïm la clau sense que ens la interceptin? O com fer que sigui pràcticament impossible deduir-la.

### 1.4.1 Criptografia Moderna

La encriptació moderna no es basa en fer que no es pugui interceptar/desxifrar la clau sinó en el fet de que desxifrar-la requereixi tal volum de càcul que un ordinador clàssic tardaria de l'ordre de desenes d'anys. Aquest mètode se l'anomena criptografia de clau pública o criptografia de clau asimètrica i requereix dues claus, una pública i una privada. El funcionament d'aquests protocols és el següent per a que Alice pugui enviar un missatge a Bob

1. Bob genera una clau privada i una clau pública.
2. Bob envia a Alice la clau pública.
3. Alice encripta el missatge utilitzant la clau pública de Bob.
4. Bob rep el missatge encriptat i el desencripta utilitzant la seva clau privada.

Encara que Eve intercepti tan el missatge com la clau pública, haurà de fer un esforç tan enorme per deduir quina és la clau privada de Bob que li resultarà impossible. Evidentment hi ha altres problemes que no considerarem com per exemple com autentiquem que Bob no hagi patit un atac de substitució o que Alice o Bob tinguin un hacker en el seu mateix entorn.

Quan Diffie i Hellman van proposar aquest mètode d'encriptació van generar una carrera per trobar possibles solucions i una d'aquestes va ser el protocol RSA.

### 1.4.2 Protocol RSA

Passem directament a veure com funciona l'algoritme del protocol RSA respecte a la generació de claus:

1. Escollim dos nombres primers de forma aleatòria i de longitud en bits semblant que anomenarem  $p$  i  $q$ . Fem  $n = pq$ .
2. CLAU PÚBLICA  $(e, n)$ . Escollim un nombre positiu  $e$  tal que  $1 < e < m$  on  $m = (p - 1)(q - 1)$ . El nombre  $e$ , juntament amb  $n$  es donen a conèixer públicament.
3. CLAU PRIVADA  $(d, n)$ . Ara trobem un nombre  $d$  tal que compleixi  $de \equiv 1 \pmod{m}$  ( $(de - 1)/m$  sigui un nombre enter). Hem obtingut la clau privada  $(d, n)$ .

Aleshores qualsevol missatge  $M$  es xifra com  $C = M^e [mod n]$  i es desxifra fent  $M = C^d [mod n]$ .

Per exemple, Alice vol enviar un missatge xifrat a Bob. Primer de tot Bob ha de crear tan la clau privada com la pública. Per això Bob escolleix  $p = 17, q = 11$  ( $n = pq = 187$ ), i agafa  $e = 7 < (p - 1)(q - 1) = 160$  i  $d = 23$  doncs és co-primer amb  $(p - 1)(q - 1) = 160$  ( $(ed - 1) \equiv 1 \pmod{160}$ ). Finalment tenim doncs: CLAU PÚBLICA  $(e = 7, n = 187)$  i CLAU PRIVADA  $(d = 23, n = 187)$ ). Veiem ara com s'aplica si Alice vol enviar un missatge xifrat a Bob

1. Alice rep la clau pública ( $e = 7, n = 187$ ) de Bob i xifra el seu missatge "x" (per exemple en codi ascii suposem que sigui  $M = 88$ ) com  $C = M^e \pmod{n} = 11$
2. Alice fa públic el seu missatge xifrat  $C = 11$
3. Bob utilitza la seva clau privada per desxifrar el missatge.  $M = C^d \pmod{n} = 88$

Com hem dit anteriorment, amb aquest algorisme la seguretat resideix en la complexitat de càlcul que requereix calcular la clau privada a partir de la clau pública per poder accedir a la informació. Amb el millor algorisme de factorització el temps va com  $t \sim \exp\left(\sqrt[3]{64/9}n^{1/3} \log^{3/2} n\right)$ . Si tenim  $n = 512$  aleshores  $t \sim 1h$ ,  $n = 1024 \rightarrow t \sim 11500$  anys i si  $n = 2048 \rightarrow t > 10^{17}s$ .

## 1.5 Comunicació clàssica: Capacitat d'un canal de comunicació amb soroll

En les dues seccions anteriors sobre comunicació clàssica, no ha estat necessari parlar de informació doncs tot el procés és determinista i per tant la informació es manté constant (de fet podem dir que constant a zero dons només tenim una possibilitat amb  $p = 1$ ). En el cas real, quant intentem enviar un missatge codificat a través d'un canal físic, sempre tindrem soroll i és aquesta aleatorietat, i el fet de voler corregir aquests errors, que ens迫a a parlar de informació o entropia.

Per a un canal sorollós, hi ha un límit d'eficiència de transmissió per aconseguir una transmissió sense errors. Això es coneix com el segon teorema de Shannon que va presentar aquest límit el 1948. Tot i així, la seva construcció de codi no és pràctica en el sentit que requereix una gran quantitat de temps. La comunitat de teoria de la informació ha estat estudiant la construcció de codis pràctics dissenyats per assolir el límit de Shannon durant anys.

El problema que volem resoldre és el següent: coneugut el soroll  $P(X|Y)$  d'un canal típic com el donat en la figura 1.1, quin és el número màxim  $M_{max}$  de símbols/missatges d'un alfabet que podem transmetre sense (o quasi sense) errors si els codifiquem utilitzant paraules de  $n$  símbols r-naris?. Considerem dos casos extrems:

- Si no hi ha soroll (exemple  $P(X = i|Y = j) = \delta_{ij}$ ), aleshores evidentment  $M_{max} = r^n$ .
- Si  $X$  i  $Y$  són independents ( $P(X|Y) = P(X)$ ), aleshores és un canal totalment inútil doncs  $M_{max} = 1$  (envia o no envia).

### 1.5.1 Entropia Condicional

Recordem que  $P(X, Y) = P(Y)P(X|Y)$ , es a dir, la probabilitat conjunta de  $X, Y$  és la probabilitat d'obtenir  $Y$  i després  $X$  condicionat al valor obtingut de  $Y$ . Per tant, la probabilitat condicional és

$$P(X|Y) = \frac{P(X, Y)}{P(Y)} \quad (1.38)$$

En la figura 1.9 es pot veure tant la distribució conjunta de les variables  $X, Y$ ,  $P(X, Y)$  en blau i normalitzada doncs el volum sota la superficie val 1, així com la corba proporcional a la distribució de la  $X$  quant sabem que la  $Y$  ha pres un valor entre  $[-2.3, -2]$ ,  $P(X|Y \in [-2.3, -2])/N$  en vermell on hem de posar el factor de normalització  $N = \int dx P(X, Y \in [-2.3, -2]) = P(Y \in [-2.3, -2])$  per assegurar-nos que l'àrea que determina  $P(X|Y \in [-2.3, -2])$  sigui 1. Per tant podem escriure

$$P(X|Y \in [-2.3, -2]) = \frac{P(X, Y \in [-2.3, -2])}{P(Y \in [-2.3, -2])} \quad (1.39)$$

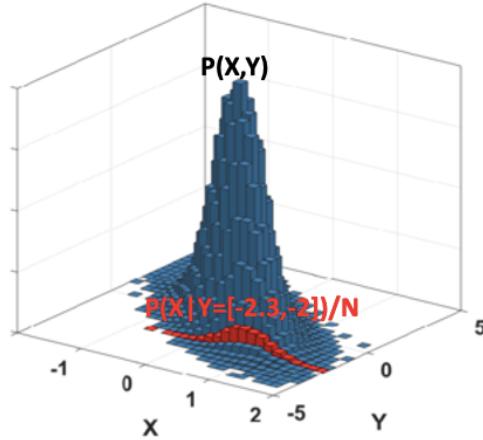


Figure 1.9:  $P(X, Y)$  en blau i  $P(X|Y \in [-2.3, -2])/N$  en vermell

Si tenim una distribució de probabilitat de dues variables,  $P(X, Y)$  la seva entropia serà evidentment

$$H(X, Y) = - \sum_{i,j} P(x_i, y_j) \ln P(x_i, y_j) \quad (1.40)$$

i l'entropia de X condicionada a conèixer que  $Y = y_j$  serà

$$H(X|Y = y_j) = - \sum_i P(x_i|y_j) \ln P(x_i|y_j) \quad (1.41)$$

Aquesta darrera expressió ens permet definir l'entropia de X condicionada a Y, com el valor esperat de la mateixa

$$\begin{aligned} H(X|Y) &= \bar{H}(X|Y = y_j) = \sum_j P(y_j) H(X|Y = y_j) = - \sum_j P(y_j) \sum_i P(x_i|y_j) \ln P(x_i|y_j) \\ &= - \sum_{i,j} P(x_i, y_j) \ln P(x_i|y_j) \end{aligned} \quad (1.42)$$

Si les dues variables són independents tenim

$$H(X|Y) = - \sum_{i,j} P(x_i, y_j) \ln P(x_i|y_j) = - \sum_{i,j} P(x_i)P(y_j) \ln P(x_i) = \sum_j P(y_j) \left( \sum_i P(x_i) \ln P(x_i) \right) = H(X) \quad (1.43)$$

com era d'esperar doncs al ser independents Y no pot condicionar de cap forma a X (el desordre de X és independent de Y)

En l'altre extrem ( $Y = X$ ) tenim que

$$H(X|X) = - \sum_{i,j} p(x_i, x_j) \ln P(x_i|x_j) = - \sum_{i,j} p(x_i)\delta_{ij} \ln \delta_{ij} = 0 \quad (1.44)$$

com era d'esperar doncs X condiciona totalment ella mateixa.

### 1.5.2 Informació Mútua

Com que  $H(X)$  és la informació promig de X i  $H(X|Y)$  és la informació promig de X condicionada a conèixer Y, podem definir la informació mútua com:

$$I(X, Y) \equiv H(X) - H(X|Y) \quad (1.45)$$

que mesura la reducció de la informació promig de X degut al coneixement de Y, o en paraules d'entropia, mesura la reducció de la incertesa o desordre de X degut al coneixement de Y. Veiem que en el cas de dues variables independents  $I(X, Y) = 0$  com era d'esperar doncs cap de les dues variables té influència amb l'altre i per tant no pot haver-hi "reducció" del desordre de X si coneixem Y.

És fàcil demostrar que

$$I(X, Y) = H(X) - H(X|Y) = \sum_{i,j} P(x_i, y_j) \ln \frac{P(x_i, y_j)}{P(x_i)P(y_j)} \quad (1.46)$$

Efectivament

$$\begin{aligned} & \sum_{i,j} P(x_i, y_j) \ln \frac{P(x_i, y_j)}{P(x_i)P(y_j)} = \sum_{i,j} P(x_i, y_j) \ln \frac{P(x_i|y_j)P(y_j)}{P(x_i)P(y_j)} \\ &= -\sum_{i,j} P(x_i, y_j) \ln P(x_i) + \sum_{i,j} P(x_i, y_j) \ln P(x_i|y_j) = -\sum_{i,j} P(x_i)P(y_j|x_i) \ln P(x_i) - H(X, Y) \\ &= -\sum_i P(x_i) \ln p(x_i) (\sum_j P(y_j|x_i)) - H(X, Y) = -\sum_i P(x_i) \ln P(x_i) - H(X, Y) = H(X) - H(X|Y) \end{aligned} \quad (1.47)$$

També són fàcils de provar les propietats de  $I(X, Y)$ :

- $I(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = I(Y, X)$
- $I(X, Y) = H(X) + H(Y) - H(X, Y)$
- $I(X, X) = H(X) - H(X|X) = H(X)$

aquesta darrera propietat ens diu que la informació de X sobre ella mateixa és precisament l'entropia  $H(X)$ , en altres paraules, la informació mútua de X sobre X és precisament tota la informació promig de X,  $H(X)$ , com era d'esperar. Per aquesta raó, l'entropia o informació promig de X,  $H(X)$ , s'anomena també com auto-informació.

### 1.5.3 Entropia Relativa

Definim l'entropia relativa, també anomenada distància de Kullback-Leibler (veure implementació d'un codi ineficient 1.35), entre dues distribucions p i q de la mateixa variable aleatòria X com:

$$D(p||q) = \sum_i p(x_i) \ln \frac{p(x_i)}{q(x_i)} \quad (1.48)$$

Notem que generalment tindrem  $D(p||q) \neq D(q||p)$  doncs l'expressió no és simètrica entre p i q.

### 1.5.4 Canal de comunicació

Què volem dir quan diem que A es comunica amb B? Volem dir que els processos físics de A han induït un estat físic en B. Aquesta transferència d'informació és un procés físic i per tant està sotmés al soroll ambiental. La comunicació direm que té èxit si el receptor B i l'emissor A coincideixen en allò que s'ha enviat.

Un sistema típic de comunicació es mostra a la figura 1.10. Suposem que tenim un alfabet amb  $M$  símbols que volem codificar en paraules d'una longitud típica de  $n$  símbols d'un codi D-nari  $D = \{x_1, x_2, \dots, x_D\}$ . Aleshores, per poder codificar tots els  $M$  símbols de l'alfabet,

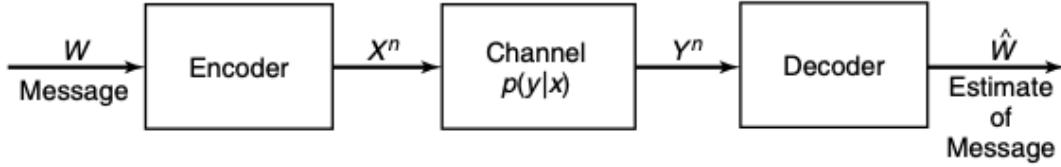


Figure 1.10: Sistema de comunicació

evidentment s'haurà de complir que  $M \leq D^n$  (per exemple, si utilitzem  $n$  bits del codi binari  $D_2 = \{0, 1\}$ , podrem codificar/enviar tots els  $M$  símbols d'un cert alfabet sempre i quant  $M \leq 2^n$ ). Aquesta seqüència de  $n$  símbols (exemple  $X^n = (X_1^n, X_2^n, \dots, X_n^n)$ , on  $X_i^n \in D$ ) es enviada pel canal, la qual cosa produeix la seqüència final del canal generalment de la mateixa longitud  $n$  i també D-nari, es a dir  $Y^n = (Y_1^n, Y_2^n, \dots, Y_n^n)$  amb  $Y_i^n \in D$ . Aquesta seqüència de sortida, degut al soroll del canal, és aleatòria, però té una distribució que depèn de la seqüència d'entrada (coneixem  $P(Y|X)$ ). Des de la seqüència de sortida, intentem recuperar el missatge transmès.

### Capacitat operativa d'un canal

**Pregunta:** conegit el soroll del canal ( $P(Y|X)$ ), quin és el número màxim  $M_{max}$  de símbols de l'alfabet que podem transmetre sense error si els codifiquem utilitzant paraules de  $n$  símbols D-naris?

Cadascuna de les seqüències d'entrada possibles al canal de la figura 1.10 (recordem que assumim que estan formades per  $n$  d-bits) induceix una distribució de probabilitats a les seqüències de sortida. Com, degut al soroll, dues seqüències d'entrada diferents poden donar lloc a la mateixa seqüència de sortida, les entrades són confuses. Demostrarem que podem triar un subconjunt "no confusable" de seqüències d'entrada de manera que amb una alta probabilitat només hi hagi una entrada molt probable que pugui haver provocat la sortida observada, es a dir, podem reconstruir aquest subconjunt de seqüències d'entrada a partir de les seqüències observades a la sortida amb una probabilitat d'error insignificant. El nombre  $M_{max}$  de seqüències d'entrada "distingibles", evidentment ha de créixer exponencialment amb  $n$ ,  $M_{max}(n) = D^{C'n} = 2^{C'n}$  (evidentment  $C' \leq 1$ ).

Aleshores, definim la capacitat operativa del canal (en bits) com

$$C_{op} = \log_2(M_{max})/n \quad (1.49)$$

### Capacitat de informació d'un canal (discret i sense memòria)

Definim un canal discret com un sistema format per un codi D-nari d'entrada  $X$  i un codi D'-nari de sortida  $Y$  (generalment  $D' = D$ ) i una matriu de transició de probabilitats  $p(y|x)$  que expressa la probabilitat d'observar el símbol de sortida  $y$  si el d'entrada és  $x$ . Es diu que el canal no té memòria si la distribució de probabilitats de la sortida depèn només de l'entrada en aquell moment i és independent de les entrades o sortides anteriors en el canal.

Per aquests tipus de canals discrets i sense memòria (és una molt bona aproximació als canals habituals), definim la capacitat de "informació" del canal com:

$$C_I = \max_{p(x)} I(X; Y) \quad (1.50)$$

on el màxim és sobre totes les possibles distribucions d'entrada  $p(x)$ .

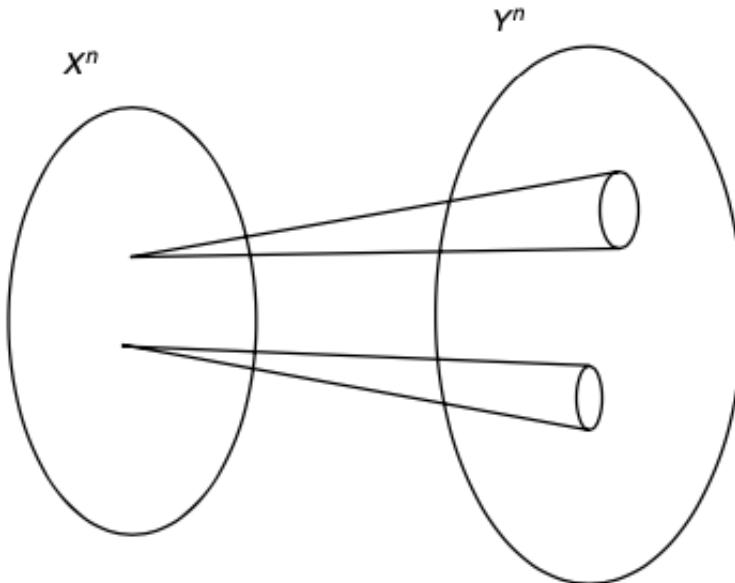


Figure 1.11: Canal amb soroll on enviem codewords formats per  $n$  D-bits

Precisament el segon teorema de Shannon estableix que la capacitat d'informació del canal és igual a la capacitat operativa del canal. El fet de que el logaritme del nombre de senyals distingibles (per d-bit, es a dir, el nombre màxim de bits que podem transmetre en cada d-bit de manera fiable i que anomenen capacitat operativa del canal) sigui igual a la informació mütua màxima (capacitat de informació d'un canal) és l'èxit central i més famós de la teoria de la informació. És el conegut segon teorema de Shannon de la capacitat d'un canal.

### 1.5.5 Segon Teorema de Shannon

Ara demostrarem el segon teorema de Shannon, que dóna un sentit pràctic a la definició de capacitat de informació com el nombre màxim de bits que podem transmetre (per cada d-bit) de manera fiable a través del canal, es a dir, la capacitat operativa (per simplicitat considerarem canals binaris però la seva generalització és trivial).

Suposem que utilitzem un codi binari sense soroll amb probabilitats  $p_0$  i  $p_1$ . Si enviem  $n$  bits, el nombre de paraules típiques (entenem per típiques les més probables, es a dir, aquelles que tinguin  $np_0$  zeros i  $np_1$  uns) seran:

$$\binom{n}{np_0} = \frac{n!}{np_0!np_1!} \quad (1.51)$$

i que utilitzant l'aproximació d'Stirling  $\ln n! \approx n \ln n - n$ , tenim

$$\ln \binom{n}{np_0} \approx -n(p_0 \ln p_0 + p_1 \ln p_1) = nH(\vec{p}) \quad (1.52)$$

i per tant en nombre de paraules típiques serà aproximadament de  $2^{nH}$ .

Consideren ara que tenim soroll al canal. Aleshores, per a cada seqüència (típica) d'entrada de longitud  $n$ , hi ha aproximadament  $2^{nH(Y|X)}$  seqüències Y possibles, totes igualment probables (figura 1.11). Com  $\approx 2^{nH(Y)}$  és el nombre total de seqüències Y (típiques) possibles a la sortida, aleshores podem dir que el nombre màxim de seqüències d'entrada que ens portin a conjunt

disjunts de sortida és inferior o igual a

$$M \leq \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y) - H(Y|X))} = 2^{nI(X;Y)} \quad (1.53)$$

Per tant, podem enviar com a màxim  $2^{nI(X;Y)}$  seqüències "distingibles" de longitud  $n$  o sigui  $nI(X;Y)$  bits.

Escolllint adequadament la distribució de les  $p(x)$  podem fer que  $I(X, Y)$  sigui màxim i per tant acabem d'obtenir que la capacitat de informació del canal  $C_I = \max_{p(x)} I(X; Y)$  és una cota superior a la capacitat d'operació del canal  $C_{op} = \log_2(M_{max})/n$  (el nombre de bits que podem enviar per un canal sense que hi hagi error). Tot i que la derivació anterior dona un límit superior de la capacitat operativa, es pot demostrar que aquesta taxa es pot aconseguir amb una probabilitat d'error arbitràriament baixa. És el que es coneix com el segon teorema de Shannon.

Anem a expressar aquest segon teorema de Shannon d'una forma més formal. Per això definim:

- Un canal discret el caracteritzarem com  $(X, p(y|x), Y)$ . X i Y són dos conjunts finits de símbols (per exemple  $X = Y = \{0, 1\}$ ) i  $p(x|y)$  la probabilitat de que l'entrada hagi estat x quant hem observat a la sortida del canal y .
- Un codi  $(M, n)$  és un conjunt format per M seqüències, cada una d'elles formada per n símbols de X. Els seus elements s'anomenen codewords (per exemple, donat un canal binari X, un possible codi  $(26, 5)$  per les 26 lletres de l'alfabet català, seria codificar cada una de les 26 lletres en 5 bits que indiquin, per exemple, la seva posició dins l'alfabet).
- La proporció  $R$  d'un codi  $(M, n)$  es defineix (en bits) com

$$R \equiv \frac{\log_2 M}{n} \quad (1.54)$$

- Una proporció R es diu que es pot aconseguir si existeix una seqüència de  $(2^{nR}, n)$  codis de tal manera que la màxima probabilitat d'error tendeix a 0 quan n tendeix a  $\infty$ .

Hem vist que la capacitat de informació del canal  $C_I = \max_{p(x)} I(X; Y)$  dona una cota màxima de totes les R assolibles. Així, per taxes R inferiors a la capacitat sempre podrem "seleccionar"  $2^{nR}$  codis d'entrada de tal manera que tinguem una probabilitat d'error arbitràriament petita per a longituds de bloc prou grans.

### Exemples de capacitat d'un canal

#### exemple 1: canal binari sense soroll

Imaginem un canal ideal on Alice envia 1 bit ( $X = \{0, 1\}$ ) i Bob rep exactament el que envia Alice. Aleshores  $Y = \{0, 1\}$  i  $P(Y = 0|X = 0) = P(Y = 1|X = 1) = 1$ ,  $P(Y = 0|X = 1) = P(Y = 1|X = 0) = 0$ . Aleshores podem calcular la informació mútua (ho farem en bits i per tant utilitzarem  $\log_2$ )

$$\begin{aligned} I(X, Y) &= \sum_{i,j=0}^1 P(X = i, Y = j) \log \frac{P(X = i, Y = j)}{P(X = i)P(Y = j)} \\ &= \sum_{i,j=0}^1 P(Y = j|X = i)P(X = i) \log \frac{P(X = i|Y = j)P(Y = j)}{P(X = i)P(Y = j)} \\ &= -\sum_{i=0}^1 P(X = i) \log P(X = i) \end{aligned} \quad (1.55)$$

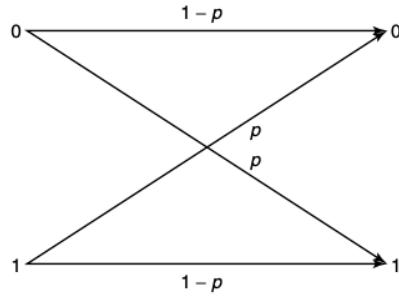


Figure 1.12: Canal binari simètric amb soroll

i que sabem és màxima quant  $P(X = 0) = P(X = 1) = 1/2$ . Per tan la capacitat de informació del canal és

$$C = \max_{p(x)} I(X, Y) = -2(1/2) \log(1/2) = 1 \quad (1.56)$$

1 bit. La taxa de transmissió de dades d'aquest canal és un bit.

Una altre forma de calcular la capacitat és recordant que  $I(X, Y) = H(X) - H(X|Y)$ , i com en aquest cas  $H(X|Y) = H(X|X) = 0$ , tenim que  $I(X, Y) = H(X)$  i sabem que per a 1 codi binari el màxim d'entropia s'asseoleix quan  $p_0 = p_1 = 1/2$ .

Per altra banda, com no hi ha soroll, el nombre  $M_{max}$  de símbols que podem distingir d'un alfabet amb aquest canal seran 2 i per tant, la capacitat operativa del canal serà  $\log_2 2 = 1$  que coincideix amb la capacitat de informació.

Si en lloc d'enviar un bit, els missatges pel canal són de  $n$  bits, com no hi ha soroll, el nombre  $M_{max}$  de símbols que podem distingir d'un alfabet amb aquest canal seran  $2^n$  i per tant, la capacitat operativa del canal ( $C_{op} = \log M_{max}/n$ ) dona també 1 com esperaríem doncs recordem  $C_{op}$  dona el nombre efectiu de bits transmesos per d-bit transmès.

### **exemple 2: canal binari simètric amb soroll**

Imaginem un canal on Alice envia bits  $X = \{0, 1\}$  i Bob rep el contrari del bit enviat per Alice amb una certa probabilitat  $p$ , tal i com es mostra en la figura 1.12. Es d'esperar que, amb comparació amb el cas anterior, la capacitat ara sigui menor que 1 ( $C < 1$ ). En aquest cas tenim

$$\begin{aligned} I(X, Y) &= H(Y) - H(Y|X) = H(Y) - \sum_{i=0}^1 p(X=i)H(Y|X=i) = H(Y) - \sum_{i=0}^1 p(X=i)H(p) \\ &= H(Y) - H(p) \leq 1 - H(p) \end{aligned} \quad (1.57)$$

on en el darrer pas hem utilitzat el fet que  $H(Y) \leq 1$  al ser  $Y$  una distribució binària. L'igualtat es complirà quan la distribució de les  $X$  sigui uniforma i per tant

$$C = \max_{p(x)} I(X, Y) = 1 - H(p) \quad (1.58)$$

Veiem que quan  $p = 0$  recuperem el resultat anterior. Per altra part, també tenim  $C = 1$  quan  $p = 1$  doncs correspon al flip del bit al passar pel canal però això ens permet reconstruir el bit enviat per Alice sense cap tipus d'ambigüïtat.

Aquest resultat ens esta dient que si en lloc d'enviar un bit, els missatges pel canal són de  $n$  bits, puc escollir hàbilment  $2^{n(1-H(p))}$  missatges dels  $2^n$  possibles de tal forma que Bob els pugui distingir-los amb una alta probabilitat. Només haig d'escollir aquests missatges d'entrada suficientment separats (suficients bits diferents entre ells) per a que, encara que alguns bits canviïn degut al soroll, el resultat final encara estigui més proper al missatge original que a qualsevol dels altres escollits.

$n=$	10			
$p$	$h(p)$	$2^n$	$2^{n(1-H)}$	$n(1-H)$
0	0	1024	1024	10
0.05	0.28639696	1024	140,656306	7,13603043
0.1	0.46899559	1024	39,6718581	5,31004406
0.15	0.6098403	1024	14,9450617	3,90159695
0.2	0.72192809	1024	6,87194767	2,78071905
0.25	0.81127812	1024	3,69921398	1,88721876
0.3	0.8812909	1024	2,27693169	1,18709101
0.35	0.93406806	1024	1,57933744	0,65931945
0.4	0.97095059	1024	1,22305905	0,29049406
0.45	0.99277445	1024	1,05135906	0,07225546
0.5	1	1024	1	0
0.55	0.99277445	1024	1,05135906	0,07225546
0.6	0.97095059	1024	1,22305905	0,29049406
0.65	0.93406806	1024	1,57933744	0,65931945
0.7	0.8812909	1024	2,27693169	1,18709101
0.75	0.81127812	1024	3,69921398	1,88721876
0.8	0.72192809	1024	6,87194767	2,78071905
0.85	0.6098403	1024	14,9450617	3,90159695
0.9	0.46899559	1024	39,6718581	5,31004406
0.95	0.28639696	1024	140,656306	7,13603043
1	0	1024	1024	10

Figure 1.13: efecte del soroll sobre el nombre efectiu de bits de informació que enviem

En la figura 1.13 hem considerat el cas on enviem  $n = 10$  bits i estudiem que passa en funció d'un soroll donat per  $p$ . Si no hi ha soroll podem enviar sense problemes 1024 missatges sense ambigüïtat, però aquest nombre queda reduït en funció de  $p$  com es veu en la quarta columna o el seu equivalent en bits en la cinquena columna. Per exemple si  $p = 0.1$  només podrem distingir uns 40 missatges o el que és equivalent 5 bits de informació. Es a dir, enviem 10 bits físics, però en realitat només podem utilitzar 5 bits de informació, dels 1024 possibles missatges només podem enviar sense problemes uns 40. Això es pot aconseguir escollint adequadament 40 combinacions d'entre les 1024 inicials que tenim de 10 bits, de forma que estiguin molt separades entre elles, de forma que si canvia un bit durant la transmissió, encara puguem inferir (amb una certa probabilitat) quin dels 40 possibles s'ha intentat enviar. Per exemple (no sé si és possible) si els 40 escollits són tals que la "distància" (entesa com el nombre de bits diferents) entre ells és sempre igual o major a 3, la probabilitat de confusió només apareixerà quant el soroll del canal canviï 2 o més bits (o sigui 1 menys la probabilitat de de que no canviï cap bit, menys la probabilitat de que només canvi un bit

$$Prob = 1 - (1 - p)^n - \frac{n!}{1!(n-1)!} p(1-p)^{n-1} = 1 - 0.9^{10} - 10 * 0.1 * 0.9^9 = 0.26 \quad (1.59)$$

Quant  $n$  augmenta podem demanar que la distància entre els missatges que hem d'escollir augmenti, fet que ens portarà a que la probabilitat de confusió baixi, tendint cap a zero quan  $n \rightarrow \infty$ .

### exemple 3: Canal binari amb soroll sense overlap

Considerem ara el cas on Alice utilitza un codi binari i Bob un codi 4-nari i que el soroll del canal produueix:

$$\begin{aligned} 0 &\rightarrow 1 \text{ o } 2 \\ 1 &\rightarrow 3 \text{ o } 4 \end{aligned} \quad (1.60)$$

A primera vista pot semblar un problema però com les possibles sortides no estan entrellaçades, podem determinar a partir d'elles quin era l'input sense ambigüïtat. Esta clar que podem enviar 2 missatges sense ambigüïtat i per tant la capacitat operativa del canal és  $\log_2 2 = 1$ .

Per calcular la capacitat de informació del canal ho podem fer com en el primer cas,

$$I(X, Y) = \sum_{i=0}^1 \sum_{j=1}^4 P(X = i, Y = j) \log \frac{P(X = i, Y = j)}{P(X = i)P(Y = j)}$$

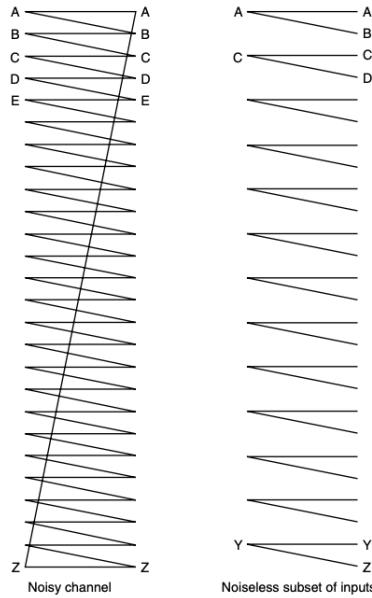


Figure 1.14: Canal amb soroll i la solució

$$\begin{aligned}
 &= \sum_{i=0}^1 \sum_{j=1}^4 P(Y = j | X = i) P(X = i) \log \frac{P(X = i | Y = j) P(Y = j)}{P(X = i) P(Y = j)} \\
 &= - \sum_{i=0}^1 P(X = i) \log P(X = i)
 \end{aligned} \tag{1.61}$$

Per tant, igual que abans per  $p_0 = p_1 = 1/2$ , I es maximitzarà i tindrem  $C = 1$ .

#### **exemple 4: Canal amb soroll sense overlap**

Considerem el cas on Alive i Bob utilitzim un codi 26-nari. Si no hi ha soroll sabem que la capacitat d'aquest canal és de  $\log_2(26)$  bits.

Ara estudiem un cas on hi ha soroll. Un cop tenim un input tenim probabilitat  $1/2$  de que surti igual o que passi a ser la següent lletra tal com es mostra a la figura 1.14. Com podem veure gràcies a la figura, només podrem enviar 13 caràcters (dels 26 que té l'alfabet) sense error. Alice podrà enviar A,C,E,...,Y i Bob els podrà descodificar i, per tant, la capacitat operativa del canal serà  $C = \log_2(13)$  bits.

Calculem ara la capacitat de informació

$$C = \max_{p(x)} I(X; Y) = \max_{p(x)} [H(Y) - H(Y|X)] = H(Y) - 1 = \log \frac{26}{2} = \log(13) \tag{1.62}$$

On en la primera igualtat hem utilitzat la definició de capacitat, en la segona la definició de la informació mútua, en la tercera hem aplicat l'operador  $\max_{p(x)}$  que ens ha fet determinar que  $p(x)$  ha d'estar distribuït uniformement sobre tots els possibles inputs per tal de fer  $H(Y|X) = 1$ . Per últim en la darrera igualtat hem aplicat la definició d'entropia i propietats dels logaritmes.

Veiem que la capacitat de informació coincideix amb la capacitat operativa.

#### 1.5.6 Detecció i correcció d'errors

Des que Shannon va publicar el seu famós article on prometia la existència d'un codi que ens permetria transmetre informació amb un rate proper a la capacitat del canal i amb una probabilitat d'error arbitràriament petita, la comunitat científica l'ha intentat trobar.

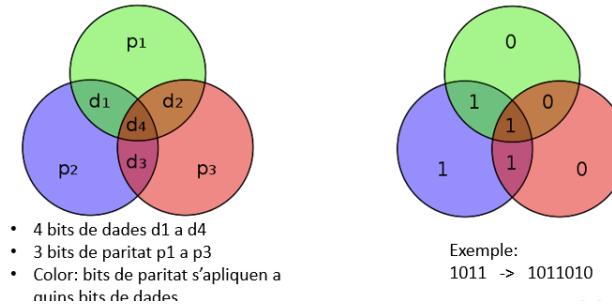


Figure 1.15: Codi Hamming(7,4)

Un possible mètode és afegir redundància ja que d'aquesta manera si part de la informació es perd o es corromp podríem igualment reproduir el missatge original. Suposem, per exemple, un canal binari (sabem que si no té soroll la seva capacitat és 1 i per tant, per a qualsevol codi que implementem tindrem  $R < 1$ ) i com alfabet a transmetre considerem únicament dos símbols: {0,1} (per tant  $M = 2$ ). Aleshores podem utilitzar pels dos símbols de l'alfabet codewords de longitud  $n = 5$  00000 per 0 i 11111 per 1 (el receptor pot descodificar-los agafant el bit "més votat"). Per tant

$$R = \frac{\log M}{n} = \frac{1}{5} \quad (1.63)$$

que es clarament més petit que 1 però no negligible. De totes maneres, si hi ha una probabilitat  $p > 0$  de que un bit del canal canviï, hi ha una probabilitat no nul·la de que la descodificació sigui incorrecta. Si féssim la mida del codeword enviat més gran reduiríem la probabilitat d'error arbitràriament ja que cada cop seria més difícil que es corrompessin tants símbols. El problema és que també fariem  $R \rightarrow 0$  mentre que el que ens interessa és que R s'apropi a C.

Un altre mètode consisteix en fer paritychecks: enviar els bits en blocs i afegir a cada bloc un bit de paritat que el receptor pot utilitzar per validar i determinar si hi ha hagut error. El problema és que aquest codi no detecta un nombre parell d'errors i, quan són imparells, no sabem quin s'ha corromput i quin no. La potència dels codis de Hamming que veurem a continuació resideix en que utilitzant l'àlgebra es pot aconseguir que puguem detectar en alguns casos quin bit s'ha corromput i per tant arreglar-ho..

**NOTA:** Hi ha una dualitat entre els problemes de compressió de dades i transmissió de dades. Durant la compressió, eliminem tota la redundància de les dades per formar la versió més comprimida possible, mentre que durant la transmissió de dades, afegim una redundància de manera controlada per combatre els errors del canal. Es pot demostrar que els problemes de compressió de dades i de transmissió de dades es poden considerar per separat.

### Codi de Hamming

El codi Hamming és un codi detector i corrector d'errors que porta el nom del seu inventor, Richard Hamming. En les dades codificades en Hamming es poden detectar errors en un bit i corregir-los, però no es distingeix entre errors de dos bits i d'un bit (pel que es fa servir Hamming estès). Això representa una millora respecte als codis amb bit de paritat, que poden detectar errors en només un bit, però no poden corregir-lo.

Un exemple és el Hamming(7,4) que afegeix 3 bits de verificació addicionals a cada quatre bit de dades (veure figura 1.15).

## 1.6 Computació clàssica

Entendrem computació com el processament de informació per obtenir uns resultats desitjats. Per això utilitzarem el que coneixem com algoritmes. Un algorisme és una recepta precisa per realitzar alguna tasca, com ara l'algorisme per calcular l'arrel quadrada d'un número (que ara ja ningú coneix!!!).

Els algorismes han estat utilitzats per la humanitat des de fa segles, com per exemple l'algorisme de dos mil anys d'Euclides per trobar el màxim divisor comú de dos enters positius.

Per obtenir el  $MCD(a, b)$ , l'algorisme d'Euclides diu que mentre  $b \neq 0$  repetiu les tres instruccions següents:

- $r = \text{residu de } a/b$
- $a = b$
- $b = r$

i el resultat és a

A principis del segle XX, Hilbert va preguntar si existia o no algun algorisme que es pogués utilitzar, en principi, per resoldre tots els problemes de les matemàtiques. Hilbert esperava que la resposta a aquesta pregunta seria sí. Sorprenentment, la resposta va resultar ser no: no hi ha cap algorisme per resoldre tots els problemes matemàtics. Per demostrar-ho, Church i Turing van haver de resoldre el profund problema de captar en una definició matemàtica del que volem dir quan fem servir el concepte intuïtiu d'algorisme. En fer-ho, van establir les bases de la teoria moderna dels algorismes i, en conseqüència, de la teoria moderna de la computació (clàssica).

Com ha resultat de la pregunta de Hilbert, Turing va definir una classe de màquines, ara conegudes com a màquines de Turing, per tal de captar la noció d'un algorisme per realitzar una tasca computacional. Un segon model de càlcul equivalent al de Turing és el que es coneix com el model de circuit de càlcul i que és especialment útil com a preparació per al nostre estudi posterior d'ordinadors quàntics.

Davant un problema del que estem buscant la solució ens podem formular dues preguntes:

- quins problemes són realment implementables com a tasques computacionals ? Es a dir, en quin casos podrem proporcionant algoritmes explícits per resoldre problemes específics
- podem donar un límit inferior de recursos computacionals necessaris per fer la tasca?. Per exemple, es poden donar límits inferiors per al nombre d'operacions que ha de realitzar qualsevol algorisme.

La situació ideal seria que la recerca d'algoritmes per resoldre problemes computacionals i les limitacions conegudes, coincidissin perfectament. A la pràctica, sovint existeix una bretxa significativa entre les millors tècniques conegudes per resoldre un problema computacional i les limitacions més estrictes conegudes de la solució.

En les properes subseccions tractarem primer els models computacionals de Turing i el de circuit, incloent la resposta negativa a la pregunta de Hilbert, i finalment la classificació dels problemes en funció del límit inferior de recursos necessaris per solucionar-los.

Evidentment, les idees de Turing o el model de circuit es poden implementar en màquines amb pocs elements diferents repetint-se moltes vegades i en una arquitectura adequada. Avui en dia, la implementació de la coneguda màquina Universal de Turing, són els ordinadors clàssics basant en l'electrònica binària.

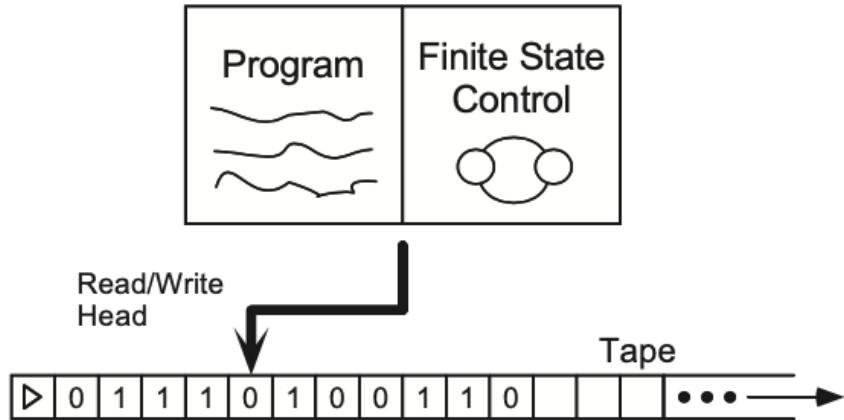


Figure 1.16: Principals elements d'una màquina de Turing. Al text, els espais en blanc de la cinta es denoten amb una 'b'. Tingueu en compte que  $\triangleright$  marca l'extrem esquerre de la cinta.

### 1.6.1 Models de computació

#### Màquina de Turing

Alan Turing va definir una màquina hipotètica que treballa de forma molt simple i que pot implementar qualsevol algorisme. La màquina consisteix en (veure figura 1.16):

- Cinta: Es divideix en cel·les individuals on en cada una hi ha un símbol d'un alfabet finit. Per simplicitat i sense pèrdua de generalitat, suposarem que l'alfabet conté quatre símbols, que denotem per 0, 1, b (el símbol 'en blanc') i  $\triangleright$ , per marcar la vora esquerra de la cinta. Inicialment, la cinta conté una  $\triangleright$  a l'extrem esquerre, un nombre finit de 0s i 1s, i la resta de la cinta conté espais en blanc.
- Capçal: Que llegeix la cinta i té la capacitat d'esborrar-ne símbols i escriure'n. També pot moure-la una cel·la (cap a l'esquerra o la dreta) a l'hora.
- Registre: Emmagatzema l'estat de la màquina de Turing. Aquest estat pot ser un entre  $\{q_1, q_2, \dots, q_m\}$  on inclús es permet variar el nombre m però sense pèrdua de generalitat podem suposar que m és una constant fixa. A més dels estats anteriors també hi ha dos estats interns especials, etiquetats  $q_s$  i  $q_h$ . Els anomenem estat inicial i estat de parada, respectivament. La idea és que al principi del càlcul, la màquina de Turing es troba en l'estat  $q_s$  inicial. L'execució del càlcul fa que l'estat intern de la màquina de Turing canviï. Si el càlcul s'acaba en algun moment, la màquina de Turing acaba en l'estat  $q_h$  per indicar que la màquina ha finalitzat el seu funcionament.
- Programa: és una taula finita d'accions on hi consten els diferents estats en els que pot estar la màquina i, dependent d'aquests juntament amb l'input que la que la màquina rep de la cinta, com ha d'actuar. En concret, un programa per a una màquina de Turing és una llista finita ordenada de línies de programa de la forma  $\langle q, x, q', x', s \rangle$ . La manera com funciona el programa és que, a cada cicle de la màquina de Turing on suposem que està el l'estat  $q$  i el capçal llegeix  $x$ , aleshores examina la llista de línies de programa en ordre, cercant una línia  $\langle q, x, \cdot, \cdot, \cdot \rangle$ . Si no troba aquesta línia de programa, l'estat intern de la màquina es canvia a  $q_h$  i la màquina atura el funcionament. Si es troba una línia així, s'executa aquesta línia de programa. L'execució d'una línia de programa implica els passos següents: l'estat intern de la màquina es canvia a  $q'$ ; el símbol  $x$  de la cinta es sobreescriu

amb el símbol  $x'$  i el cap de la cinta es mou cap a l'esquerra, cap a la dreta o es queda quiet, segons si  $s$  sigui  $-1, +1, 0$  respectivament. L'única excepció a aquesta regla és si el cap de cinta es troba al quadrat de la cinta més a l'esquerra i  $s = -1$ , en aquest cas el cap de cinta es manté.

Un exemple de programa por ser

- $\langle q_s, \triangleright, q_1, \triangleright, +1 \rangle$
- $\langle q_1, 0, q_1, b, +1 \rangle$
- $\langle q_1, 1, q_1, b, +1 \rangle$
- $\langle q_1, b, q_2, b, -1 \rangle$
- $\langle q_2, b, q_2, b, -1 \rangle$
- $\langle q_2, \triangleright, q_3, \triangleright, +1 \rangle$
- $\langle q_3, b, q_h, 1, 0 \rangle$

Podem comprovar que el programa anterior calcula  $f(x) = 1$  per a qualsevol  $x$  inicial de la cinta. El  $x$  pot estar expressat en forma binaria i on hem posat una  $b$  quant  $x$  s'acaba d'expressar. El 1 apareixerà en la segona casella de la cinta, immediatament després de la primera on tenim  $\triangleright$  que no hem modificat durant el programa.

Podem calcular coses més complicades?. La hipòtesis de Church-Turing ens diu que: la classe de funcions computables per una màquina de Turing correspon exactament a la classe de funcions que consideraríem com a computables mitjançant un algorisme. A priori no és obvi que totes les funcions que intuïtivament consideraríem computables mitjançant un algorisme es poden calcular mitjançant una màquina de Turing però en quasi 100 anys no s'han trobat proves en contra.

El fet de que a qualsevol màquina de Turing se li pot assignar un número natural  $X$  ens permet introduir el concepte de màquina de Turing universal (UTM). Assumint  $E$  possibles estats  $q_i$ , el nombre de possibles instruccions  $\langle q, x, q', x', s \rangle$  és  $I = E \times 4 \times E \times 4 \times 3$  i per tant, qualsevol programa es pot descriure amb  $I$  bits que l'identifiquen unívocament només indicant amb 0 o 1 si la instrucció associada apareix o no en el codi. El nombre natural  $M$  que representa aquesta llista de bits, l'associarem a la màquina de Turing que anomenarem  $T_M$  (evidentment alguns nombres  $M$  correspondran a màquines que no fan res).

La màquina Universal Turing (figura 1.17) actua de la forma següent. Sigui  $M$  qualsevol màquina de Turing i  $T_M$  sigui el número de Turing associat a la màquina  $M$ . A continuació, introduceix la representació binària per al  $T_M$  seguit del banc, seguit de qualsevol cadena de símbols  $x$  a la resta de la cinta, la Màquina Universal de Turing dóna com a sortida qualsevol màquina que  $M$  tingui a l'entrada de  $x$ . Per tant, la màquina universal de Turing és capaç de simular qualsevol altra màquina de Turing. La màquina Universal Turing té un esperit similar a un ordinador programable modern, en el qual l'acció que ha de fer l'ordinador (el “programa”) s’emmagatzema a la memòria, de manera anàloga a la cadena de bits  $T_M$  emmagatzemada al començament de la cinta Màquina Universal de Turing. Les dades que el programa ha de processar s’emmagatzemen en una part independent de la memòria, de manera anàloga al paper de  $x$  a la màquina universal de Turing. A continuació, s'utilitza alguna maquinaria fixa per executar el programa, produint la sortida. Aquest maquinaria fixa és anàleg als estats interns i el programa (fix) que executa la màquina Universal Turing.

Resulta que els ordinadors quàntics també obereixen la tesi Church-Turing. és a dir, els ordinadors quàntics poden calcular la mateixa classe de funcions que la màquina de Turing. La

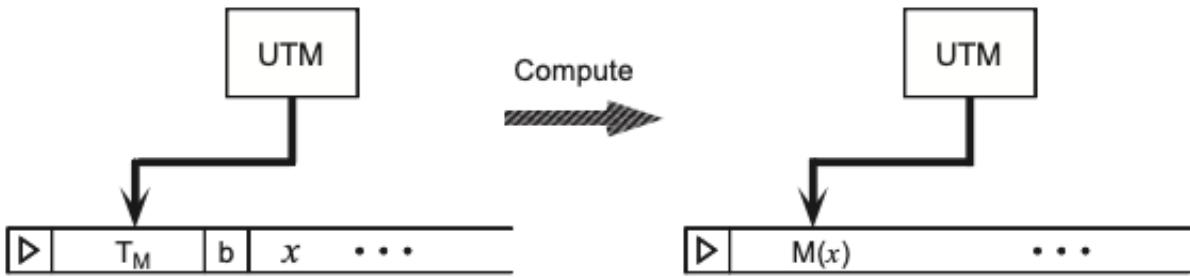


Figure 1.17: Màquina de Turing Uninversal. Actua com un ordinador actual.

diferència entre els ordinadors quàntics i les màquines de Turing resulta de la eficiència amb què es pot realitzar el càlcul de la funció; hi ha funcions que es poden calcular de manera molt més eficient en un ordinador quàntic del que es creu que seria possible amb un ordinador clàssic.

Tornem al problema de Hilbert que va originar tot el camp de la computació: hi ha algun algorisme per decidir tots els problemes de les matemàtiques? Church i Turing van demostrar que la resposta a aquesta pregunta era negativa i està lligada al concepte de indecidibilitat (per exemple, avui es sap que el problema de decidir si dos espais topològics són equivalents topològicament ("homeomorf") és indecidible).

Per resoldre el problema de Hilbert, Turing va utilitzar la numeració  $T_M$  de les possibles màquines per plantear el problema de la detenció: la màquina amb el número Turing  $x$  s'atura quan s'introdueix el número  $y$ ? Aquest és un problema matemàtic ben plantejat i interessant. Al cap i a la fi, és una qüestió d'interès considerable si els nostres algoritmes s'aturen o no. Tot i així, resulta que no hi ha cap algorisme capaç de resoldre el problema de la parada. Per veure això, Turing es va preguntar si hi ha un algorisme per resoldre un problema encara més especialitzat: la màquina amb el número Turing  $x$  s'atura quan s'introdueix el mateix nombre  $x$ ? Turing va definir la funció

$$h(x) = \begin{cases} 0 & \text{si el número de màquina } x \text{ no s'atura en introduir } x \\ 1 & \text{si el número de màquina } x \text{ s'atura en introduir } x \end{cases} \quad (1.64)$$

Si hi ha un algorisme per resoldre el problema de detenció, segur que hi ha un algoritme per avaluar  $h(x)$ . Intentarem arribar a una contradicció suposant que existeix tal algorisme, denotat per  $\text{HALT}(x)$ . Penseu en un algorisme que calcula la funció  $\text{TURING}(x)$ , amb pseudocodi

$\text{TURING}(x)$

- $y = \text{HALT}(x)$
- if  $y = 0$  then
- halt
- else
- loop forever
- end if

Com que  $\text{HALT}$  és un programa vàlid,  $\text{TURING}$  també ha de ser un programa vàlid, amb algun número de Turing,  $t$ . Per definició de la funció d'aturada,  $h(t) = 1$  si i només si  $\text{TURING}$  s'atura a l'entrada de  $t$ . Però en inspeccionar el programa per a  $\text{TURING}$ , veiem que  $\text{TURING}$  s'atura a l'entrada de  $t$  si i només si  $h(t) = 0$ . Així,  $h(t) = 1$  si i només si  $h(t) = 0$ , una

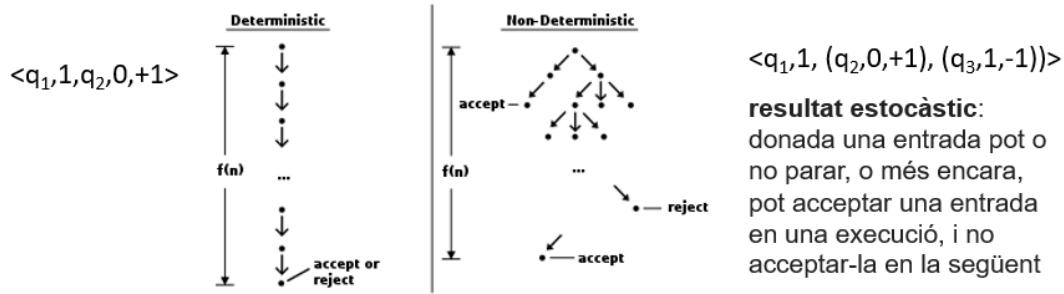


Figure 1.18: Comparació màquines de Turing deterministes versus probabilístiques

NOT	AND	NAND	OR	NOR	XOR	XNOR
$\overline{A}$	$AB$	$\overline{AB}$	$A+B$	$\overline{A+B}$	$A \oplus B$	$\overline{A \oplus B}$
$A$	$A$	$B$	$X$	$A$	$B$	$X$
$0$	$1$	$0$	$0$	$0$	$1$	$0$
$1$	$0$	$1$	$0$	$1$	$0$	$1$
			$1$	$0$	$1$	$0$
			$1$	$1$	$0$	$1$

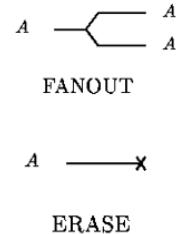


Figure 1.19: portes lògiques elementals.

contradicció. Per tant, la nostra suposició inicial que existeix un algorisme per avaluar  $h(x)$  deu haver estat incorrecta. Concloem que no hi ha cap algorisme que ens permeti resoldre el problema de la parada.

## Màquina de Turing probabilística

Una Màquina de Turing probabilística és una màquina de Turing que selecciona aleatoriament entre les transicions disponibles a cada punt amb idèntica probabilitat per cada alternativa com es pot veure en la figura 1.18. El resultat és estocàstic: donada una entrada pot o no parar, o més encara, pot acceptar una entrada en una execució, i no acceptar-la en la següent.

Circuits

El model de circuit, que és equivalent a la màquina de Turing (determinista) en termes de potència computacional, però és més còmode i realista per a moltes aplicacions. En particular, el model de circuit de càlcul és especialment important com a preparació per a la nostra investigació d'ordinadors quàntics.

Un circuit està format per cables i portes (generalment electròniques), que transporten informació al voltant i realitzen tasques de càlcul senzilles, respectivament. En generalment, un circuit pot implicar molts bits d'entrada i sortida, molts cables i moltes portes lògiques. Una porta lògica és una funció  $f : \{0, 1\}^k \rightarrow \{0, 1\}^l$  d'un nombre fix  $k$  de bits d'entrada a un nombre fix  $l$  de bits de sortida. També és habitual fer la convenció que no es permeten bucles al circuit, per evitar possibles inestabilitats (com pot passar si la sortida de la porta que inverteix el bit el reconnectem a la seva entrada). Diem que un circuit d'aquest tipus és acíclic i ens adherim a la convenció que els circuits del model de circuit de càlcul són acíclics.

Hi ha moltes altres portes lògiques elementals que són útils per al càlcul. Una llista parcial es

dona en la figura 1.19. Als circuits sovint permetem que els bits es "divideixin", substituint un bit per dues còpies de si mateix, una operació coneguda com a FANOUT. També permetem el CROSSOVER de bits, és a dir, a partir dues sequències de n bits, les tallen en una certa posició i en fem dues sequències noves de n bits, intercanviant els dos darrers fragments. (NOTA: en computadors quàntics no tenim l'equivalent del FANOUT degut al teorema de no-clonació).

Un resultat teòric important és que qualsevol funció de bits es pot calcular a partir de la composició de portes NAND, i per això es coneix com la porta UNIVERSAL. Per contra, XOR sol o fins i tot junt amb NOT, no és universal. Una manera de veure-ho és observar que aplicar una porta XOR no canvia la paritat total dels bits. Com a resultat, qualsevol circuit que impliqui només portes XOR mantindrà la paritat entre l'entrada i la sortida, restringint així la classe de funcions que es poden calcular i, per tant, excloent la universalitat.

Es diu que una família de circuits  $\{C_n\}$  s'anomena família de circuits uniforme si hi ha algun algorisme que s'executa en una màquina de Turing que, després de l'entrada de n, genera una descripció de  $\{C_n\}$ . és a dir, l'algoritme proporciona una descripció de quines portes hi ha al circuit  $\{C_n\}$ , de com es connecten aquestes portes per formar un circuit, de qualsevol bit auxiliar que necessiti el circuit i de les operacions, i on s'ha de llegir la sortida del circuit .

Es pot demostrar que la classe de funcions computables per famílies de circuits anomenats "uniformes" és exactament la mateixa que la classe de funcions que es poden calcular en una màquina de Turing. Amb aquesta restricció d'uniformitat, els resultats en el model de càlcul de la màquina de Turing se solen donar una traducció directa al model de circuits de càlcul i viceversa. Més tard, prestem una atenció similar a les qüestions de la uniformitat en el model de càlcul de circuits quàntics.

### 1.6.2 Programació d'ordinadors clàssics

Evidentment en els inicis dels primer computadors, no hi havien llenguatges d'alt nivell i per "programar-los" havien d'utilitzar llenguatges de baix nivell, com pot ser l'ensamblador on les seves instruccions són molt simples (moure bits, copiar-los, operar-los amb les portes bàsiques,...). Però fins i tot abans de l'assemblar la implementació dels algoritmes s'havia de fer pensant en quines portes lògiques elementals que podem implementar fàcilment amb electrònica ( figura 1.19) s'han de combinar per obtenir la sortida desitjada. Per exemple en la figura 1.20 es pot veure com implementar la suma de números de 8 bits utilitzant portes elementals (primer la suma de dos bits utilitzant una porta XOR i una AND, després suma d'un tercer i així fins a 8).

Avui en dia ja disposem de llenguatges d'alt nivell que ens faciliten la implementació de qualsevol algorisme en els ordinadors convencionals. Generalment abans d'escriure el codi, es descriu l'algorisme en pseudo-codi o es fa un diagrama de flux com es pot veure en la figura 1.21 pel cas de calcular  $N!$ . En la mateixa figura es mostra el codi en llenguatge C i també com seria una funció que implementés  $N!$  que admetés crides recursives a la mateixa funció (legant però poc eficient).

### 1.6.3 Classes de complexitat computacional

Avui en dia dins del marc de la teoria de complexitat computacional podem trobar diferents classificacions segons la mida o el temps requerit per resoldre'ls problemes de decisió.

De problemes de decisió en tenim de dos tipus:

- Existeix una solució? Exemple: donat el conjunt  $\{-2, -3, 8, 15, -10\}$  existeix algun subconjunt que doni 0? (si)
- És solució?. Exemple:  $\{-2, -3, -10, 15\}$  és solució de l'anterior problema? (si)

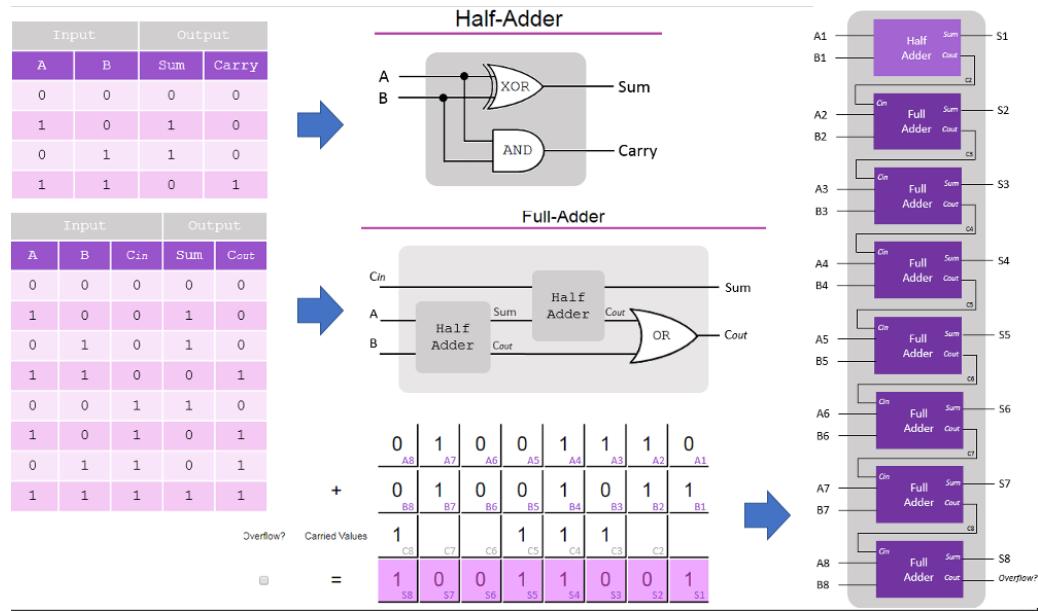
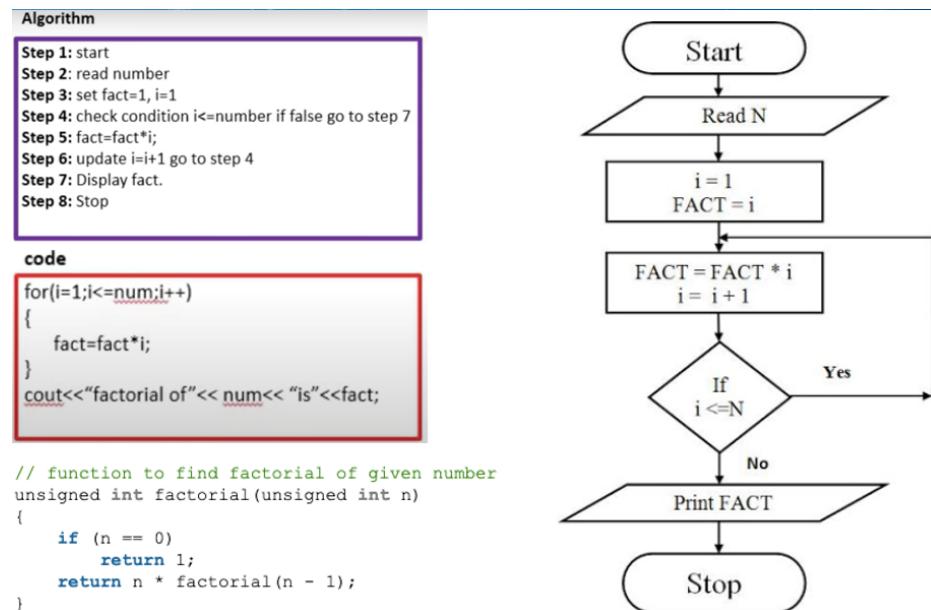


Figure 1.20: Implementació de la suma binària utilitzant portes lògiques

Figure 1.21: Algorisme per calcular  $n!$  i la seva implementació en un programa d'alt nivell per ser executat en una computadora clàssica

**Important complexity classes** [edit]

Many important complexity classes can be defined by bounding the time or space used by the algorithm. Some important complexity classes of decision problems defined in this manner are the following:

Complexity class	Model of computation	Resource constraint	Complexity class	Model of computation	Resource constraint
Deterministic time			Deterministic space		
<b>DTIME(<math>f(n)</math>)</b>	Deterministic Turing machine	Time $O(f(n))$	<b>DSPACE(<math>f(n)</math>)</b>	Deterministic Turing machine	Space $O(f(n))$
			<b>L</b>	Deterministic Turing machine	Space $O(\log n)$
<b>P</b>	Deterministic Turing machine	Time $O(\text{poly}(n))$	<b>PSPACE</b>	Deterministic Turing machine	Space $O(\text{poly}(n))$
<b>EXPTIME</b>	Deterministic Turing machine	Time $O(2^{\text{poly}(n)})$	<b>EXPSPACE</b>	Deterministic Turing machine	Space $O(2^{\text{poly}(n)})$
Non-deterministic time			Non-deterministic space		
<b>NTIME(<math>f(n)</math>)</b>	Non-deterministic Turing machine	Time $O(f(n))$	<b>NSPACE(<math>f(n)</math>)</b>	Non-deterministic Turing machine	Space $O(f(n))$
			<b>NL</b>	Non-deterministic Turing machine	Space $O(\log n)$
<b>NP</b>	Non-deterministic Turing machine	Time $O(\text{poly}(n))$	<b>NPSPACE</b>	Non-deterministic Turing machine	Space $O(\text{poly}(n))$
<b>NEXPTIME</b>	Non-deterministic Turing machine	Time $O(2^{\text{poly}(n)})$	<b>NEXPSPACE</b>	Non-deterministic Turing machine	Space $O(2^{\text{poly}(n)})$

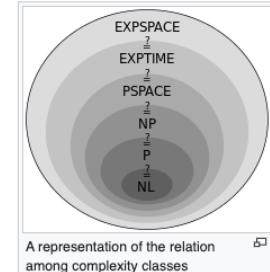


Figure 1.22: classes de complexitat

Dins dels que s'encarreguen del temps trobem els problemes P i els problemes NP.

- classe P: consisteix en tots els problemes de decisió que poden ser resolts en un ordinador clàssic determinista en un temps polinòmic respecte la mida de les entrades.
- classe NP: la formen aquells problemes de decisió en els quals verificar una solució en màquines deterministes es pot fer en un temps polinòmic però no així trobar les seves solucions (o el que es creu equivalentment, les seves solucions poden ser trobades en un temps polinòmic en una màquina no determinista) .

Un dels problemes que no s'ha resolt en computer science és saber si  $P = NP?$ , es a dir, si una solució positiva per un problema de SI/NO pot ser verificada ràpidament, poden calcular-se igual de ràpid les respuestes?.

La consideració de màquines de Turing NO deterministes ens permet introduir una nova classe BPP que consisteix en tots els problemes de decisió que poden ser resolts per una màquina de Turing probabilística en temps polinòmic (respecte la mida de les entrades) amb una probabilitat d'error menor del 1/3. Igual que abans ens podem preguntar si  $P = BPP?$ , es a dir, hi ha un problema que es pugui resoldre en temps polinòmic mitjançant una màquina de Turing probabilista però no una màquina de Turing determinista? O, les màquines de Turing deterministes poden simular de manera eficient totes les màquines probabilístiques de Turing amb, com a màxim, una desacceleració polinòmica? NO HO SABEM.

El cas de la factorització, que és el que ens interessa ja que és el requerit per la encriptació, no el podem col·locar a cap dels dos grups. Amb els coneixements d'avui en dia no es sap si podem resoldre el problema amb temps polinòmic però de moment no es pot. Amb els ordinadors quàntics la factorització esdevé un problema del tipus P.

Un resum de les possibles categories que actualment es consideran tan en temps com en recursos es pot veure a la figura 1.22

## 1.7 Informació quàntica

La informació quàntica difereix fortament de la informació clàssica. La unitat fonamental d'informació clàssica és el bit, la unitat més bàsica d'informació quàntica és el qbit. La informació clàssica es mesura mitjançant l'entropia de Shannon  $S = -\sum_i p_i \ln p_i$ , mentre que l'anàlogic de la mecànica quàntica és l'entropia de Von Neumann  $S = -Tr(\rho \ln \rho)$

La informació d'un estat pur donat és zero, com ho és en el cas clàssic al especificar l'estat exactament, per exemple, estar en l'estat 0. Això no s'ha de confondre amb la informació que pot emmagatzemar un bit o un qbit. Un bit, sinó queda identificat el seu estat, pot estar en 0 o 1, i si els considerem equiprobables aleshores la informació promig o entropia serà 1 bit. Igualment passa amb un qbit, si no especificuem el seu estat i considerem equiprobables dos estats seus ortogonals, aleshores la informació promig o entropia serà 1 qbit. Com veiem, al parlar de informació, hem de tenir molt clar a quina distribució de probabilitats ens estem referint. Per exemple en el cas d'un bit, especificar el seu estat com 0 vol dir  $p_0 = 1, p_1 = 0$  i per tant  $I = 0$ , per contra, no especificar-lo i assumir que  $p_0 = 1/2, p_1 = 1/2$  ens porta a  $I = 1$  bit.

La informació quàntica es pot desplaçar per un canal quàntic, anàlogament al concepte de canal de comunicacions clàssic. Els missatges quàntics tenen una mida finita, mesurada en qbits; els canals quàntics tenen una capacitat de canal finita, mesurada en qbits.

Degut a l'evolució unitària dels estats en mecànica quàntica, tenim cinc famosos teoremes que descriuen els límits de la manipulació de la informació quàntica:

- teorema de no teleportació, que afirma que un qbit no es pot convertir (totalment) en bits clàssics; és a dir, no es pot "llegir". Dit d'una altra manera, afirma que la unitat d'informació quàntica, el qbit, no es pot convertir exactament en bits d'informació clàssica (Això no s'ha de confondre amb la teleportació quàntica, que permet destruir un estat quàntic en un lloc i crear una rèplica exacta en un lloc diferent). El teorema és conseqüència dels efectes probabilístics d'una mesura sobre un estat quàntic. Un cop convertida la mesura en informació clàssica, la informació quàntica no es pot recuperar (En canvi, si és possible convertir la informació clàssica en informació quàntica i tornar a la informació clàssica, només hem de codificar els bits clàssics en estats quàntics ortogonals, que sempre es poden distingir).
- teorema de no clonació, que impedeix fer una copia d'un qbit arbitrari. El teorema de la no teleportació implica el teorema de la no clonació: si fos possible convertir un qbit en bits clàssics, llavors un qbit seria fàcil de copiar (ja que els bits clàssics es poden copiar trivialment). El teorema de la no-clonació (tal com s'entén generalment) es refereix només als estats purs, mentre que l'enunciat generalitzat sobre estats barreja es coneix com el teorema de no difusió. El teorema de no clonació impedeix l'ús de tècniques clàssiques de correcció d'errors en estats quàntics. Per exemple, les còpies de seguretat d'un estat enmig d'un càlcul quàntic no es poden crear i utilitzar per corregir errors posteriors. La correcció d'errors és vital per a la pràctica de la computació quàntica, i durant algun temps es va pensar que això era una limitació fatal. El 1995, Shor i Steane van reviure les perspectives de la informàtica quàntica dissenyant independentment els primers codis de correcció d'errors quàntics que no es basen en la copia (impossible) de qbits.
- teorema de no supressió, que evita que es suprimeixi un qbit arbitrari utilitzant una transformació unitària. Això és conseqüència que, en general, donades dues còpies d'algun estat quàntic arbitrari, és impossible eliminar una de les còpies de forma unitària al ser un procés dual invertit en el temps del teorema de la no clonació. Per eliminar-lo l'hem de mesurar.

- teorema de no emissió. Tot i que es pot transportar un qbit únic d'un lloc a un altre (per exemple, mitjançant la teleportació quàntica), no es pot lliurar a diversos destinataris. És un corol·lari del teorema de la no clonació: els estats quàntics no es poden copiar en general, no es poden emetre.
- teorema de no hiding, que demostra que si es perd informació d'un sistema mitjançant la decoherència que li produceix l'entorn, aquesta es trasllada al subespai de l'entorn i no pot romandre en la correlació entre el sistema i l'entorn. Aquesta és una conseqüència fonamental de la linealitat i la unitaritat de la mecànica quàntica. El teorema demostra que si l'amplitud de probabilitat desapareix d'un sistema, reapareixerà en un altre sistema. Atès que la funció d'ona conté tota la informació rellevant sobre un sistema físic, aquesta conservació de la funció d'ona equival a la conservació de la informació quàntica.

### 1.7.1 Qbits

El qbit és la unitat mínima i per tant constitutiva de la informació en la teoria quàntica, és a dir, l'anàleg al bit en la teoria clàssica. En què consisteix? El qbit és qualsevol estat d'un sistema quàntic amb dos únics nivells, que anomenarem  $|0\rangle$  i  $|1\rangle$ , o sigui  $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ .

Els qbits experimentals que s'utilitzen majoritàriament són:

- Els ions considerant només els dos estats possibles que té aquest sistema físic:  $|groundstate\rangle \rightarrow |0\rangle$  o  $|excited\rangle \rightarrow |1\rangle$  (on hem posat les etiquetes 0 i 1 arbitràriament).
- Els electrons on els dos estats possibles d'spin amunt o avall:  $|+\rangle \rightarrow |0\rangle$  o  $|-\rangle \rightarrow |1\rangle$
- Polarització dels fotons.

Ara ens centrarem en el segon cas però tot és exactament idèntic per la resta. Podem agafar com a base dels qbits  $|0\rangle = |Sz, +\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ,  $|1\rangle = |Sz, -\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , vectors que són propis de l'operador

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (1.65)$$

amb valors propis  $\pm\hbar/2$ . Recordem que  $S_z$  és l'operador associat a fer la mesura de l'spin de l'electró utilitzant un Stern-Gerlach en la direcció  $\hat{z}$ .

Recordem que l'operador associat a mesurar de l'spin de l'electró utilitzant un Stern-Gerlach en la direcció arbitrària  $\hat{d} = (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta)$  de coordenades esfèriques  $(\theta, \phi)$  és

$$S_{\hat{d}} = \vec{S} \cdot \hat{d} = \frac{\hbar}{2} \vec{\sigma} \cdot \hat{d} = \frac{1}{2} \begin{pmatrix} d_3 & d_1 - id_2 \\ d_1 + id_2 & -d_3 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} \cos\theta & e^{-i\phi} \sin\theta \\ e^{i\phi} \sin\theta & -\cos\theta \end{pmatrix} \quad (1.66)$$

que té com a vectors propis

$$\begin{aligned} |S_{\hat{d}}, +\rangle &= \begin{pmatrix} \cos(\theta/2) \\ e^{i\phi} \sin(\theta/2) \end{pmatrix} = \cos \frac{\theta}{2} |Sz, +\rangle + e^{i\varphi} \sin \frac{\theta}{2} |Sz, -\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \\ |S_{\hat{d}}, -\rangle &= \begin{pmatrix} -e^{-i\phi} \sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix} = \cos \frac{\theta}{2} |Sz, -\rangle - e^{-i\varphi} \sin \frac{\theta}{2} |Sz, -\rangle = \cos \frac{\theta}{2} |1\rangle - e^{-i\varphi} \sin \frac{\theta}{2} |0\rangle \end{aligned} \quad (1.67)$$

amb valors propis  $\pm\hbar/2$ .

L'estat més general possible és qualsevol combinació  $|\psi\rangle = \alpha|Sz, -\rangle + \beta|Sz, +\rangle$  i que sempre podem escriure com  $|S_d, +\rangle$  (exceptuant una fase global). Per tant, podem determinar els valors de  $(\theta, \phi)$ . Això ens permet identificar qualsevol estat d'spin 1/2 amb una direcció en l'espai (per spins més grans de 1/2 no es pot fer doncs tenim més de dos estats propis associats a la mesura de un SG).

A més a més sabem com manipular aquests spins i portar-los de qualsevol direcció a qualsevol altre. Això ho podem fer utilitzant camps magnètics, tan uniformes com dependents del temps, durant un cert interval de temps i que sabem venen implementats per transformacions unitàries sobre els estats

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle \quad (1.68)$$

un  $U(t, t_0)$  és l'operador d'evolució temporal i que en el cas que  $H$  no depengui del temps, val  $U(t, t_0) = \exp(-i/\hbar H(t - t_0))$ .

Per exemple, el moviment d'una partícula de spin 1/2, dins un camp magnètic uniforme (si no estem interessats en la descripció de la part espacial, generalment, per estar davant un sistema de comportament espacial clàssic), ve descrit pel hamiltonià  $H$

$$H = -\vec{\mu} \cdot \vec{B} = g \frac{\mu_B}{\hbar} \vec{S} \cdot \vec{B} = \frac{e}{m} \frac{\hbar}{2} B \vec{\sigma} \cdot \hat{n} \quad (1.69)$$

on  $\hat{n}$  és la direcció del camp magnètic. Per tant, l'operador d'evolució temporal és

$$U(t, 0) = \exp\left(-\frac{i}{\hbar} H t\right) = \exp\left(-\frac{i}{\hbar} \frac{e}{m} \frac{\hbar}{2} B \vec{\sigma} \cdot \hat{n} t\right) = \exp\left(-\frac{i}{2} \varphi(t) \hat{n} \cdot \vec{\sigma}\right) \quad (1.70)$$

on  $\varphi(t) = eBt/m$ .

Per fer el calcul de la darrera expressió, recordem que per spin 1/2, la rotació  $\mathcal{R}(\varphi, \hat{n})$  queda implementada en aquest espai per l'operador unitari

$$\begin{aligned} U_{\mathcal{R}}(\varphi, \hat{n}) &= \exp\left(-\frac{i}{\hbar} \varphi \hat{n} \cdot \vec{S}\right) = \exp\left(-\frac{i}{2} \varphi \hat{n} \cdot \vec{\sigma}\right) = I \cos \frac{\varphi}{2} - i \hat{n} \cdot \vec{\sigma} \sin \frac{\varphi}{2} \\ &= \begin{pmatrix} \cos \frac{\varphi}{2} - i n_3 \sin \frac{\varphi}{2} & (-in_1 - n_2) \sin \frac{\varphi}{2} \\ (-in_1 + n_2) \sin \frac{\varphi}{2} & \cos \frac{\varphi}{2} + i n_3 \sin \frac{\varphi}{2} \end{pmatrix} \end{aligned} \quad (1.71)$$

Aleshores si inicialment l'estat és, per exemple,  $|\psi(t=0)\rangle = |Sz, +\rangle$ , l'estat evoluciona com

$$\begin{aligned} |\psi(t)\rangle = U(t, 0)|\psi(0)\rangle &= \begin{pmatrix} \cos \frac{\varphi(t)}{2} - i n_3 \sin \frac{\varphi(t)}{2} & (-in_1 - n_2) \sin \frac{\varphi(t)}{2} \\ (-in_1 + n_2) \sin \frac{\varphi(t)}{2} & \cos \frac{\varphi(t)}{2} + i n_3 \sin \frac{\varphi(t)}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos \frac{\varphi(t)}{2} - i n_3 \sin \frac{\varphi(t)}{2} \\ (-in_1 + n_2) \sin \frac{\varphi(t)}{2} \end{pmatrix} \end{aligned} \quad (1.72)$$

on veiem que escollint adequadament la direcció  $\hat{n}$  del camp magnètic i el temps ( $\varphi(t) = eBt/m$ ) que està actuant, podem portar l'estat inicial que apunta en la direcció  $\hat{z}$  a qualsevol altre direcció.

### 1.7.2 Entrellaçament

Els estats entrellaçats juguen un paper molt rellevant en la computació quàntica i per aquest motiu els hi dediquem aquesta secció.

Un estat  $|\psi\rangle$  direm que és un estat entrellaçat entre dos dels seus subsistemes si l'estat de cada un dels dos subsistemes no queda ben definit, es a dir, l'estat d'un subsistema està correlat quànticament (direm entrellaçat) amb el de l'altre subsistema.

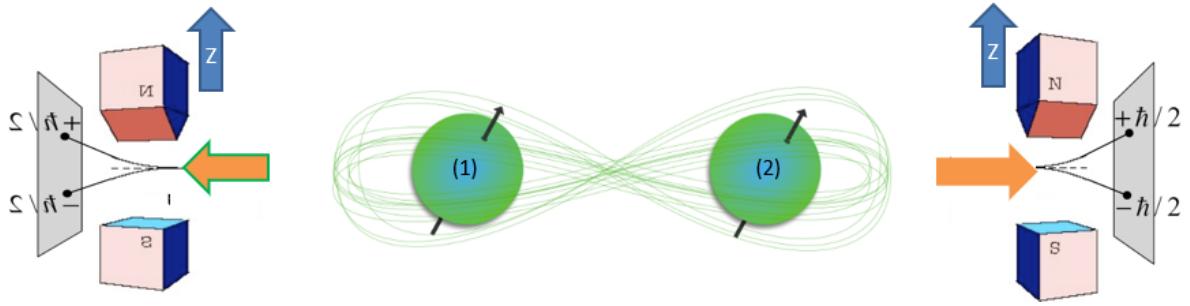


Figure 1.23: Mesura simultània de l'spin en direcció  $z$  de les dues partícules

Un exemple el tenim en l'estat de dues parícules d'spin 1/2

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} - |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)})$$

que és un estat entrellaçat doncs l'estat de la primera (o de la segona) partícula no queda ben definit, pot estar tant up o down i totalment correlat amb la segona: quan està up(down) la segona està down(up).

Una forma experimental de crear entrellaçament, per exemple entre spins, és aprofitant la interacció spin-spin quant estan aprop i que recordem és de la forma  $H = \frac{A}{2\hbar} \vec{S}_1 \cdot \vec{S}_2 = \frac{A}{\hbar} (\vec{S}^2 - \vec{S}_1^2 - \vec{S}_2^2)$ .

Per exemple, considerem dos partícules d'spin 1/2 que inicialment es troben en l'estat separat:

$$|\psi(0)\rangle = |S_x, +\rangle \otimes |S_z, -\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)|-\rangle = \frac{1}{\sqrt{2}}|+-\rangle + \frac{1}{\sqrt{2}}|--\rangle = \frac{1}{2}(|10\rangle + |00\rangle) + \frac{1}{\sqrt{2}}|1-1\rangle \quad (1.73)$$

expressat en la base acoblada. Com en aquest cas  $H = A/\hbar(\vec{S}^2 - 3\hbar^2/2)$ , l'evolució de l'estat serà (obviant la fase global irrellevant  $\exp(-3At/2)$ )

$$\begin{aligned} |\psi(t)\rangle &= \exp(-\frac{i}{\hbar}Ht)|\psi(0)\rangle = \exp(-iAs(s+1))t|\psi(0)\rangle = \frac{1}{2}(e^{-i2At}|10\rangle + |00\rangle) + \frac{e^{-i2At}}{\sqrt{2}}|1-1\rangle \\ &= e^{-i2At}\left(\frac{1}{2}|10\rangle + \frac{1}{\sqrt{2}}|1-1\rangle\right) + \frac{1}{2}|00\rangle \\ &= e^{-i2At}\left(\frac{1}{2\sqrt{2}}(|+-\rangle + |-+\rangle) + \frac{1}{\sqrt{2}}|--\rangle\right) + \frac{1}{2\sqrt{2}}(|+-\rangle - |-+\rangle) \\ &= \frac{1}{2\sqrt{2}}(e^{-i2At} + 1)|+-\rangle + \frac{1}{2\sqrt{2}}(e^{-i2At} - 1)|-+\rangle + \frac{e^{-i2At}}{\sqrt{2}}|--\rangle \end{aligned} \quad (1.74)$$

que és un estat entrellaçat (per la majoria de  $t$ ) .

Anem a veure alguns estats entrellaçats i els resultats esperats si fem mesures sobre ells.

**Estat entrellaçat:**  $|\psi\rangle = \frac{1}{\sqrt{2}}(|S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} - |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)})$

Fent, sobre aquest estat, les mesures (en l'eix  $z$ ) de la figura 1.23, obtenim:

$$P_{|\psi\rangle} \left( S_z^{(1)} : +\frac{\hbar}{2}; S_z^{(2)} : -\frac{\hbar}{2} \right) = \left| \left( \langle S_z, +|^{(1)} \otimes \langle S_z, -|^{(2)} \right) |\psi\rangle \right|^2 = \frac{1}{2}$$

$$P_{|\psi\rangle} \left( S_z^{(1)} : -\frac{\hbar}{2}; S_z^{(2)} : +\frac{\hbar}{2} \right) = \left| \left( \langle S_z, -|^{(1)} \otimes \langle S_z, +|^{(2)} \right) |\psi\rangle \right|^2 = \frac{1}{2}$$

$$P_{|\psi\rangle} \left( S_z^{(1)} : +\frac{\hbar}{2}; S_z^{(2)} : +\frac{\hbar}{2} \right) = \left| (\langle S_z, +|^{(1)} \otimes \langle S_z, +|^{(2)}) |\psi\rangle \right|^2 = 0$$

$$P_{|\psi\rangle} \left( S_z^{(1)} : -\frac{\hbar}{2}; S_z^{(2)} : -\frac{\hbar}{2} \right) = \left| (\langle S_z, -|^{(1)} \otimes \langle S_z, -|^{(2)}) |\psi\rangle \right|^2 = 0$$

i per tant

- la suma dels dos resultats (en l'eix z) és SEMPRE zero
- els producte del signes dels dos resultats (en l'eix z) SEMPRE és negatiu

El primer resultat reflecteix el fet que l'estat  $|\psi\rangle$  és propi, amb valor propi 0, de l'operador  $S_z$  associat a mesurar el spin TOTAL en la direcció  $z$ .

$$S_z |\psi\rangle = (S_z^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_z^{(2)}) |\psi\rangle = 0$$

$$P_{|\psi\rangle}(S_z : 0) = \langle \psi | \Pi_0 | \psi \rangle = \langle \psi | \Pi_+^{(1)} \otimes \Pi_-^{(2)} + \Pi_-^{(1)} \otimes \Pi_+^{(2)} | \psi \rangle = 1$$

És fàcil comprovar que

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} - |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)}) = \frac{1}{\sqrt{2}} (|S_{\hat{n}}, +\rangle^{(1)} \otimes |S_{\hat{n}}, -\rangle^{(2)} - |S_{\hat{n}}, -\rangle^{(1)} \otimes |S_{\hat{n}}, +\rangle^{(2)})$$

per a qualsevol direcció  $\hat{n}$  i per tant

- la suma dels dos resultats és SEMPRE zero EN QUALESEVOL DIRECCIÓ
- els producte del signes dels dos resultats SEMPRE és negatiu EN QUALESEVOL DIRECCIÓ

Una forma alternativa de veureu és observant que l'estat considerat és (en la base acoplada)  $|l=0, m=0\rangle$  i per tant, al ser un escalar, s'ha de comportar de la mateixa manera en qualsevol direcció.

**Estat entrellaçat:**  $|\psi\rangle = \frac{1}{\sqrt{2}} (|S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} + |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)})$

Fent, sobre aquest estat, les mesures (en l'eix z) segons la figura 1.32, obtenim:

$$P_{|\psi\rangle} \left( S_z^{(1)} : +\frac{\hbar}{2}; S_z^{(2)} : -\frac{\hbar}{2} \right) = \left| (\langle S_z, +|^{(1)} \otimes \langle S_z, -|^{(2)}) |\psi\rangle \right|^2 = \frac{1}{2}$$

$$P_{|\psi\rangle} \left( S_z^{(1)} : -\frac{\hbar}{2}; S_z^{(2)} : +\frac{\hbar}{2} \right) = \left| (\langle S_z, -|^{(1)} \otimes \langle S_z, +|^{(2)}) |\psi\rangle \right|^2 = \frac{1}{2}$$

$$P_{|\psi\rangle} \left( S_z^{(1)} : +\frac{\hbar}{2}; S_z^{(2)} : +\frac{\hbar}{2} \right) = \left| (\langle S_z, +|^{(1)} \otimes \langle S_z, +|^{(2)}) |\psi\rangle \right|^2 = 0$$

$$P_{|\psi\rangle} \left( S_z^{(1)} : -\frac{\hbar}{2}; S_z^{(2)} : -\frac{\hbar}{2} \right) = \left| (\langle S_z, -|^{(1)} \otimes \langle S_z, -|^{(2)}) |\psi\rangle \right|^2 = 0$$

i per tant

- la suma dels dos resultats (en l'eix z) és SEMPRE zero
- els producte del signes dels dos resultats (en l'eix z) SEMPRE és negatiu

El primer resultat reflecteix el fet que l'estat  $|\psi\rangle$  és propi, amb valor propi 0, de l'operador  $S_z$  associat a mesurar el spin TOTAL en la direcció  $z$ .

$$S_z|\psi\rangle = (S_z^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_z^{(2)})|\psi\rangle = 0$$

$$P_{|\psi\rangle}(S_z : 0) = \langle\psi|\Pi_0|\psi\rangle = \langle\psi|\Pi_+^{(1)} \otimes \Pi_-^{(2)} + \Pi_-^{(1)} \otimes \Pi_+^{(2)}|\psi\rangle = 1$$

Ara bé, en aquest cas

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} + |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)}) \neq \frac{1}{\sqrt{2}}(|S_{\hat{n}}, +\rangle^{(1)} \otimes |S_{\hat{n}}, -\rangle^{(2)} + |S_{\hat{n}}, -\rangle^{(1)} \otimes |S_{\hat{n}}, +\rangle^{(2)})$$

i per tant els resultats poden dependre de la direcció de la mesura. Anem a calcular primer el valor de  $\langle S_{\hat{n}} \rangle_{|\psi\rangle}$

$$\begin{aligned} S_{\hat{n}}|\psi\rangle &= (S_{\hat{n}}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_{\hat{n}}^{(2)}) \frac{1}{\sqrt{2}}(|S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} + |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)}) \\ &= \left[ \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{\hbar}{2} \begin{pmatrix} \cos \theta & \sin \theta e^{-i\phi} \\ \sin \theta e^{i\phi} & -\cos \theta \end{pmatrix} \right] \\ &\quad \frac{1}{\sqrt{2}} \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \\ &= \frac{1}{\sqrt{2}} \left[ \begin{pmatrix} \cos \theta \\ \sin \theta e^{i\phi} \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} \sin \theta e^{i\phi} \\ -\cos \theta \end{pmatrix} + \begin{pmatrix} \sin \theta e^{i\phi} \\ -\cos \theta \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} \cos \theta \\ \sin \theta e^{i\phi} \end{pmatrix} \right] \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \sin \theta e^{i\phi} \\ \cos \theta \\ -\cos \theta \\ \sin \theta e^{i\phi} \end{pmatrix} \end{aligned} \tag{1.75}$$

i per tant

$$\langle S_{\hat{n}} \rangle_{|\psi\rangle} = \langle\psi|S_{\hat{n}}|\psi\rangle = \frac{1}{\sqrt{2}}(0 \ 1 \ 1 \ 0) \frac{1}{\sqrt{2}} \begin{pmatrix} \sin \theta e^{i\phi} \\ \cos \theta \\ -\cos \theta \\ \sin \theta e^{i\phi} \end{pmatrix} = 0$$

es a dir que per aquest estat també la suma dels dos resultats és SEMPRE zero EN QUALESEVOL DIRECCIÓ.

Ara bé, com

$$|S_{\hat{n}}, +\rangle|S_{\hat{n}}, +\rangle = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos^2 \frac{\theta}{2} \\ \cos \frac{\theta}{2} e^{i\phi} \sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} e^{i\phi} \sin \frac{\theta}{2} \\ e^{i2\phi} \sin^2 \frac{\theta}{2} \end{pmatrix}$$

aleshores

$$\begin{aligned} P_{|\psi\rangle} \left( S_{\hat{n}}^{(1)} : +\frac{\hbar}{2}; S_{\hat{n}}^{(2)} : +\frac{\hbar}{2} \right) &= \left| \left( \langle S_{\hat{n}}, +|^{(1)} \otimes \langle S_{\hat{n}}, +|^{(2)} \right) |\psi\rangle \right|^2 \\ &= \left| \left( \cos^2 \frac{\theta}{2} \quad \cos \frac{\theta}{2} e^{-i\phi} \sin \frac{\theta}{2} \quad \cos \frac{\theta}{2} e^{-i\phi} \sin \frac{\theta}{2} \quad e^{-i2\phi} \sin^2 \frac{\theta}{2} \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right|^2 \\ &= \frac{1}{2} \left| \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{-i\phi} + \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{-i\phi} \right|^2 = 2 \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2} \end{aligned} \tag{1.76}$$

i com  $\langle S_{\hat{n}} \rangle_{|\psi\rangle} = 0$  aleshores s'ha de complir que

$$P_{|\psi\rangle} \left( S_{\hat{n}}^{(1)} : -\frac{\hbar}{2}; S_{\hat{n}}^{(2)} : +\frac{\hbar}{2} \right) = P_{|\psi\rangle} \left( S_{\hat{n}}^{(1)} : +\frac{\hbar}{2}; S_{\hat{n}}^{(2)} : +\frac{\hbar}{2} \right)$$

En resum, com que

$$\begin{aligned} \langle S_{\hat{n}} \rangle &= (+\hbar)P(++) + (-\hbar)P(--) + (0\hbar)P(+-) + (0\hbar)P(-+) = (+\hbar)P(++) + (-\hbar)P(--) \\ \langle \text{signe} \rangle &= (+1)[P(++) + P(--)] + (-1)[P(+-) + P(-+)] \end{aligned} \quad (1.77)$$

tenim que

$$P(++) = P(--) = 2 \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2}, \Rightarrow \langle S_{\hat{n}} \rangle = 0$$

$$\langle \text{signe} \rangle = (+1)P_+ + (-1)P_- = P_+ - (1 - P_+) = 2P_+ - 1 = 8 \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2} - 1$$

o sigui que

- la suma dels dos resultats és SEMPRE zero EN QUALSEVOL DIRECCIÓ
- el producte del signes dels dos resultats depèn de la DIRECCIÓ

Una forma alternativa de veureu és observant que l'estat considerat és (en la base acoplada)  $|l=1, m=0\rangle$  i per tant, encara que el spin total ha de ser 0 en qualsevol direcció, al NO ser un escalar, el producte dels signes NO s'ha de comportar de la mateixa manera en qualsevol direcció.

### 1.7.3 Desigualtats de Bell

Ara compararem les prediccions que fa la MQ dels resultats de les mesures sobre estats entrelaçats (i que coincideixen amb els observats experimentalment) amb les prediccions que podríem esperar de teories locals.

#### Prediccions de la MQ

En certs processos es produueixen parelles de partícules en l'estat

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( |S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} - |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)} \right) \quad (1.78)$$

i que hem vist tenen spin 0 en qualsevol direcció:

$$\left. \begin{array}{l} S_z |\psi\rangle = (S_z^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_z^{(2)}) |\psi\rangle = 0 |\psi\rangle \\ S_x |\psi\rangle = (S_x^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_x^{(2)}) |\psi\rangle = 0 |\psi\rangle \\ S_y |\psi\rangle = (S_y^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_y^{(2)}) |\psi\rangle = 0 |\psi\rangle \end{array} \right\} \Rightarrow S_{\vec{n}} |\phi\rangle = \vec{n} \cdot \vec{S} |\psi\rangle = 0, \forall \vec{n} \quad (1.79)$$

Si en aquests estats *entrellaçats* mesurem l'spin de la primera partícula en direcció  $z$  i ens dóna  $+\hbar/2$ , aleshores el col-lapse instantani de la funció ens diu que la segona estarà forçosament en l'estat  $|S_z, -\rangle$ , encara que aquesta segona partícula estigui molt allunyada de la primera.

Suposem ara que mesurem simultàniament l'spin de la partícula 1 en la direcció  $\vec{a}$  i el de la partícula 2 en la direcció  $\vec{b}$  (figura 1.24). Les dues mesures es fan de tal manera que no pot haver-hi connexió casual entre elles i repetim l'experiment diverses vegades per obtenir el valor del producte dels signes. Podem predir el valor esperat del producte dels signes dels seus spins com

$$E(\vec{a}, \vec{b}) = \langle \psi | \vec{\sigma} \cdot \vec{a} \otimes \vec{\sigma} \cdot \vec{b} | \psi \rangle \quad (1.80)$$

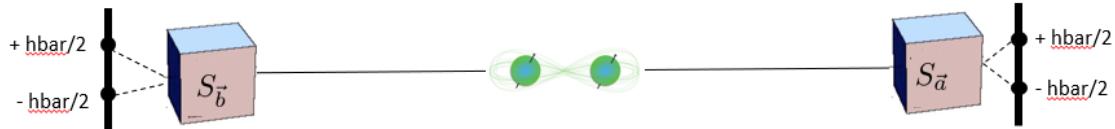


Figure 1.24: Mesura simultània de l'spin en direcció  $\vec{a}$  ( $i \vec{b}$ ) sobre la primera (i segona) partícula

Per calcular aquest valor esperat suposarem que primer fem la mesura en una de les partícules, per exemple la de la dreta, fet que porta al col·lapse de l'estat de l'altra partícula a un de ben definit, i que immediatament després fem la mesura sobre aquest. Suposem, doncs, que sobre la partícula de la dreta fem una mesura  $S_{\vec{a}}$ . El resultat d'aquesta mesura pot ser  $+\hbar/2$  o  $-\hbar/2$ , ambdós amb la mateixa probabilitat (0.5), segons hem deduït a (??). Si el resultat és  $+\hbar/2$  ( $-\hbar/2$ ), podrem assegurar que la partícula de l'esquerra és en l'estat  $|S_{\vec{a}}, -\rangle$  ( $|S_{\vec{a}}, +\rangle$ ), ja que hem dit que l'spin total és 0 en qualsevol direcció i ho hem vist explícitament en l'equació (??). Aleshores, si fem una mesura  $S_{\vec{b}}$  sobre aquesta partícula de l'esquerra, la probabilitat d'obtenir un resultat de signe contrari a l'obtingut sobre la partícula de la dreta serà:

$$\begin{aligned} P_{|S_{\vec{a}}, -\rangle} \left( S_{\vec{b}} : -\frac{\hbar}{2} \right) &= \cos^2 \frac{\phi}{2}, \text{ si } S_{\vec{a}} \text{ sobre la partícula de la dreta ha donat } +\hbar/2 \\ P_{|S_{\vec{a}}, +\rangle} \left( S_{\vec{b}} : +\frac{\hbar}{2} \right) &= \cos^2 \frac{\phi}{2}, \text{ si } S_{\vec{a}} \text{ sobre la partícula de la dreta ha donat } -\hbar/2 \end{aligned} \quad (1.81)$$

on  $\phi$  és l'angle entre els vectors  $\vec{a}$  i  $\vec{b}$  com ja hem dit abans. Aquest resultat s'obté considerant la direcció  $\vec{a}$  com la de l'eix  $z$  i  $\vec{b}$  com la direcció  $\vec{n}$  en (??), llavors  $P_{|S_{\vec{a}}, -\rangle} \left( S_{\vec{b}} : -\frac{\hbar}{2} \right) = |\langle S_{\vec{b}}, - | S_{\vec{a}}, - \rangle|^2 = \cos^2 \frac{\phi}{2}$ .

Observem que la probabilitat trobada és independent del resultat de la mesura  $S_{\vec{a}}$ . Aleshores, anomenant  $P_+$  la probabilitat que les mesures  $S_{\vec{a}}$  i  $S_{\vec{b}}$  donin el mateix signe, i  $P_-$  que donin signe contrari, com que hem vist que  $P_- = \cos^2 \frac{\phi}{2}$ , tindrem

$$E(\vec{a}, \vec{b}) = (+1)P_+ + (-1)P_- = 1 - 2P_- = 1 - 2\cos^2 \frac{\phi}{2} = -\cos \phi \quad (1.82)$$

### Teoria local

Observem que la natura es comporta com prediu la MQ. Ara bé, la MQ ens diu que hi ha un col·lapse instantani de l'estat que fa que la mesura d'una part del sistema tingui influència instantània en l'altra part, encara que no estiguin connectades casualment. Per això ens podem preguntar si aquesta "influència" NO local de la MQ i els seus resultats es podrien reproduir amb teories que mantinguessin la localitat, es a dir, teories on cada una de les dues partícules ja surt amb la informació necessària que determina el resultat de la mesura que es farà sobre ella, independentment de la mesura que es farà sobre l'altre (com per exemple, l'observació de que tenen spin oposats si fem les dues mesures en la mateixa direcció sobre l'estat estudiat anteriorment). Per això haurem de "compartir" algunes variables que les correlacionin i aquestes teories es coneixen com teories locals.

Suposem doncs que pot existir una teoria local que pot descriure els mateixos resultats que prediu l'MQ. Aquesta teoria haurà de tenir, com a mínim, una variable  $\lambda$  que especifiqui completament l'estat del sistema. Cada sistema generat tindrà una  $\lambda$  ben definida i la generació seguirà una certa densitat de probabilitat  $p(\lambda)$ . Indiquem, per mitjà d' $A(\vec{a}, \lambda)\hbar/2$ , el resultat de mesurar l'spin de la partícula 1 en la direcció  $\vec{a}$  ( $A$  només pot prendre valors  $+1$  i  $-1$ ) i, per

mitjà de  $B(\vec{b}, \lambda)\hbar/2$ , el resultat de mesurar l'spin de la partícula 2 en la direcció  $\vec{b}$ . Com que l'spin total ha de ser 0 en qualsevol direcció,  $A(\vec{a}, \lambda)\hbar/2 = -B(\vec{a}, \lambda)\hbar/2$ . En una teoria local,  $A$  pot dependre de  $\vec{a}$  i de  $\lambda$ , però no de  $\vec{b}$ , i la mitjana del producte de les mesures simultànies fetes en direccions  $\vec{a}$  i  $\vec{b}$  serà (hem tret el factor  $\hbar^2/4$  per simplicitat)

$$\epsilon(\vec{a}, \vec{b}) = \int p(\lambda) A(\vec{a}, \lambda) B(\vec{b}, \lambda) d\lambda \quad (1.83)$$

notem que  $\epsilon(\vec{a}, \vec{a}) = E(\vec{a}, \vec{a}) = -1$  (la definició de  $E(\vec{a}, \vec{b})$  es va donar a (1.80)), com era d'esperar.

Comparem ara diferents mesures:

$$\begin{aligned} \epsilon(\vec{a}, \vec{b}) - \epsilon(\vec{a}, \vec{c}) &= \int p(\lambda) (A(\vec{a}, \lambda) B(\vec{b}, \lambda) - A(\vec{a}, \lambda) B(\vec{c}, \lambda)) d\lambda \\ &= - \int p(\lambda) A(\vec{a}, \lambda) A(\vec{b}, \lambda) (1 + A(\vec{b}, \lambda) B(\vec{c}, \lambda)) d\lambda \end{aligned} \quad (1.84)$$

ja que  $A(\vec{a}, \lambda)^2 = 1$ . Aleshores,

$$|\epsilon(\vec{a}, \vec{b}) - \epsilon(\vec{a}, \vec{c})| \leq \int p(\lambda) (1 + A(\vec{b}, \lambda) B(\vec{c}, \lambda)) d\lambda = 1 + \epsilon(\vec{b}, \vec{c}) \quad (1.85)$$

ja que  $A$  només pot agafar els valors  $+1$  o  $-1$  i  $p(\lambda)$  és sempre positiu. Això és una desigualtat de Bell i l'ha de complir qualsevol teoria local.

Com a exemple inventat de teoria local podem considerar el següent: les dues partícules que surten en direccions oposades tenen una direcció de spin en el pla perpendicular a l'eix de propagació, que forma un angle respecte a un eix vertical  $z$ , de  $\lambda$  per a la primera partícula i  $\lambda + \pi$  per a la segona. Aleshores, suposarem que el resultat de mesurar l'spin de la primera partícula en la direcció  $\vec{a}$ , continguda també en el pla perpendicular a l'eix de propagació, que indicarem per mitjà d' $A(\vec{a}, \lambda)$ , és  $+1$  o  $-1$ , segons que l'angle entre el seu spin i  $\vec{a}$  sigui menor o major que  $\pi/2$ . Per analogia, es defineix  $B(\vec{b}, \lambda)$  sobre la segona partícula.

Per la mateixa construcció es compleix:  $A(\vec{a}, \lambda) = -B(\vec{a}, \lambda)$ . Si considerem que els estats són generats amb una distribució uniforme en  $\lambda$ , la probabilitat d'obtenir les dues mesures,  $A(\vec{a}, \lambda)$  i  $B(\vec{b}, \lambda)$ , amb signe contrari és de

$$\mathcal{P}(-) = 1 - \frac{\phi}{\pi} \quad (1.86)$$

on  $\phi$  ( $\phi < \pi$ ) és l'angle entre els vectors  $\vec{a}$  i  $\vec{b}$ . Aleshores,

$$\epsilon(\vec{a}, \vec{b}) = 1 - 2\mathcal{P}(-) = -1 + 2\frac{\phi}{\pi} \quad (1.87)$$

Si ara considerem un tercer vector  $\vec{c}$  entre els vectors  $\vec{a}$  i  $\vec{b}$  ( $\phi_{ab} = \phi_{ac} + \phi_{cb}$ ), tindrem

$$|\epsilon(\vec{a}, \vec{b}) - \epsilon(\vec{a}, \vec{c})| = |2\frac{\phi_{ab}}{\pi} - 2\frac{\phi_{ac}}{\pi}| = \frac{2}{\pi}|\phi_{bc}| = 1 + \epsilon(\vec{b}, \vec{c}) \quad (1.88)$$

que compleix, evidentment, la desigualtat.

### Violació de les desigualtats de Bell per part de la MQ

Hem vist que qualsevol teoria local de variables amagades ha de complir en aquests tipus d'experiment

$$|E(\vec{a}, \vec{b}) - E(\vec{a}, \vec{c})| \leq 1 + E(\vec{b}, \vec{c}) \quad (1.89)$$

per a qualsevol conjunt de direccions arbitràries de  $\vec{a}$ ,  $\vec{b}$  i  $\vec{c}$ . No obstant això, si agafem, per exemple,  $2\pi/3$  l'angle entre el vector  $\vec{a}$  i el vector  $\vec{b}$ , i  $\vec{c}$  al mig dels dos, aleshores, segons els càlculs obtinguts en el marc de la MQ,

$$\begin{aligned}|E(\vec{a}, \vec{b}) - E(\vec{a}, \vec{c})| &= |-\cos(2\pi/3) + \cos(\pi/3)| = 1 \\ 1 + E(\vec{b}, \vec{c}) &= 1 - \cos(\pi/3) = 1/2\end{aligned}\tag{1.90}$$

que violen clarament la desigualtat de Bell.

Com que els resultats experimentals concorden amb les prediccions de la MQ i per tant violen clarament les *desigualtats de Bell*, haurem de descartar la possibilitat de descriure les observacions experimentals amb alguna teoria local de *variables amagades*.

#### 1.7.4 La base de Bell de l'espai de Hilbert de dos qbits $C^2 \otimes C^2$

En l'espai de Hilbert de dos qbits  $C^2 \otimes C^2$  podem treballar amb infinites bases, però entre elles tenim la base de Bell que és molt útil per caracteritzar estats quànticament entrellaçats.

Designem per  $|0\rangle$  i  $|1\rangle$  els dos estats ortogonals possibles i A o B són els dos sistemes entrellaçats. Les bases més utilitzades són:

1) La base desacoplada

$$\{|0\rangle_{(A)}|0\rangle_{(B)}, |0\rangle_{(A)}|1\rangle_{(B)}, |1\rangle_{(A)}|0\rangle_{(B)}, |1\rangle_{(A)}|1\rangle_{(B)}\}$$

2) La base acoplada

$$\begin{aligned}|1, 1\rangle &= |0\rangle_{(A)}|0\rangle_{(B)} \\ |1, 0\rangle &= \frac{1}{2}(|0\rangle_{(A)}|1\rangle_{(B)} + |1\rangle_{(A)}|0\rangle_{(B)}) \\ |1, -1\rangle &= |1\rangle_{(A)}|1\rangle_{(B)} \\ |0, 0\rangle &= \frac{1}{2}(|0\rangle_{(A)}|1\rangle_{(B)} - |1\rangle_{(A)}|0\rangle_{(B)})\end{aligned}\tag{1.91}$$

3) La base de Bell:

$$\begin{aligned}|\Psi^-\rangle &= \sqrt{\frac{1}{2}}(|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B) \\ |\Psi^+\rangle &= \sqrt{\frac{1}{2}}(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B) \\ |\varphi^-\rangle &= \sqrt{\frac{1}{2}}(|0\rangle_A|0\rangle_B - |1\rangle_A|1\rangle_B) \\ |\varphi^+\rangle &= \sqrt{\frac{1}{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)\end{aligned}\tag{1.92}$$

Aquest darrers formen una base ortogonal i properament veurem que són estats màximament entrellaçats (observant l'entropia de les seves parts, i evidentment aquest entrellaçament quàntic es manté sota transformacions unitàries del tipus  $U_{AB} = U_A \otimes U_B$  al ser un canvi de base local). També tenen la propietat

$$(1 \otimes \sigma_x)|\Psi^\pm\rangle = |\varphi^\pm\rangle\tag{1.93}$$

#### 1.7.5 Descomposició de Schmidt

Donat un vector (estat pur) de l'espai de Hilbert de dos qbits  $C^2 \otimes C^2$ , molt sovint és complicat veure si és un estat entrellaçat o no.

En l'àlgebra lineal, la descomposició de Schmidt (que porta el nom del seu creador Erhard Schmidt) es refereix a una manera particular d'expressar un vector com a producte tensorial de dos espais. Té nombroses aplicacions en teoria de la informació quàntica, per exemple en la determinació de si un estat és entrellaçat o no.

L'estat més general de  $H = H_A \otimes H_B = C^{d_A} \otimes C^{d_B}$  és

$$|\Psi\rangle_{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} t_{ij} |i\rangle_A |j\rangle_B \quad (1.94)$$

on  $|i\rangle_A (|j\rangle_B)$  és una base ortonormal del primer(segón) espai que anomenen A(B). El que podem fer és aplicar el *single value decomposition* a la matriu  $t_{ij}$  de dimensió  $d_A \times d_B$  i escriure-la com  $t = U\Lambda V$  on U (de  $d_A \times d_A$ ) i V (de  $d_B \times d_B$ ) són matrius unitàries i  $\Lambda$  és una matriu diagonal  $d_A \times d_B$  amb  $d \leq \min(d_A, d_B)$  nombres reals i estrictament positius en la diagonal i la resta tots zeros. Per tant

$$t_{ij} = \sum_{k=1}^d U_{ik} \lambda_k V_{kj} \quad (1.95)$$

i l'estat  $|\Psi\rangle_{AB}$  es pot escriure com

$$|\Psi\rangle_{AB} = \sum_{i=1}^{d_A} \sum_{j=1}^{d_B} \sum_k^d U_{ik} \lambda_k V_{kj} |i\rangle_A |j\rangle_B = \sum_k^d \lambda_k |\Lambda\rangle_k |\Omega\rangle_k \quad (1.96)$$

on  $|\Lambda\rangle_k = U_{ik} |i\rangle_A$  ( $|\Omega\rangle_k = V_{kj} |j\rangle_B$ ) és una base ortonormal de A(B).

Veiem que d'aquesta manera podrem veure fàcilment si l'estat  $|\Psi\rangle_{AB}$  està entrellaçat o no: si  $d = 1$  serà NO entrellaçat (hi ha un únic terme i per tant l'estat en cada subsistema queda perfectament definit) i si  $d > 1$  l'estat és entrellaçat.

Es defineix la matriu densitat d'un estat pur com  $\rho \equiv |\Psi\rangle\langle\Psi|$  i les matrius densitat marginals de cada subespai com  $\rho_A = Tr_B(\rho) \equiv \sum_j \langle j|_B \rho |j\rangle_B$  on  $|j\rangle_B$  és una base ortonormal de B. Com l'operador  $\rho_A$  és independent de la base escollida per B (si fem un canvi de base en B, veiem que les transformacions unitàries que les implementen cancel·len en la definició de  $\rho_A$ ), aleshores a la pràctica el que farem és obtenir les matrius densitat marginals de cada subespai i al diagonalitzar-les hem d'obtenir:

$$\begin{aligned} \rho_A &= \sum_k^d \lambda_k^2 |\Lambda\rangle_k \langle \Lambda|_k \\ \rho_B &= \sum_k^d \lambda_k^2 |\Omega\rangle_k \langle \Omega|_k \end{aligned} \quad (1.97)$$

i el nombre  $d$  de valors  $\lambda_k \neq 0$  determina si és o no entrellaçat (recordem que  $d \leq \min(d_A, d_B)$  i, com són matrius densitat,  $\sum_{k=1}^d \lambda_k^2 = 1$ ).

Mètode:

$$\rho = |\Psi\rangle\langle\Psi| \rightarrow \rho_A = \sum_j \langle j|_B \rho |j\rangle_B \rightarrow \text{diagolalitzar}(\rho_A) \rightarrow \lambda_k^2$$

**exemple.** Determinar si l'estat  $|\Psi\rangle = (1/\sqrt{10})(2|00\rangle + |01\rangle + |10\rangle + 2|11\rangle)$  és un estat entrellaçat o no.

Per això calculem  $\rho_A = Tr_B(\rho)$  i la diagonalitzem per trobar els  $\lambda_k^2$ . Com

$$\rho = \frac{1}{10}(2|00\rangle + |01\rangle + |10\rangle + 2|11\rangle)(2\langle 00| + \langle 01| + \langle 10| + 2\langle 11|) \quad (1.98)$$

tenim

$$\begin{aligned}\rho_A &= Tr_B(\rho) = \langle 0|_B \rho |0\rangle_B + \langle 1|_B \rho |1\rangle_B \\ &= \frac{1}{10}(2|0\rangle + |1\rangle)(2\langle 0| + \langle 1|) + \frac{1}{10}(|0\rangle + 2|1\rangle)(\langle 0| + 2\langle 1|) \\ &= = \frac{1}{10}(5|0\rangle\langle 0| + 4|0\rangle\langle 1| + 4|1\rangle\langle 0| + 5|1\rangle\langle 1|)\end{aligned}\quad (1.99)$$

Per tant, la representació matricial de  $\rho_A$  en la base d'A  $\{|0\rangle, |1\rangle\}$  és

$$\rho_A = \frac{1}{10} \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} \quad (1.100)$$

que ja podem diagonalitzar i trobar el valors propis ( $\lambda_k^2$ ) i els vector propis  $|\Lambda\rangle_k$ .

$$0 = \det(\rho_A - \alpha I) = \begin{vmatrix} 5/10 - \alpha & 4/10 \\ 4/10 & 5/10 - \alpha \end{vmatrix} = (5/10 - \alpha)^2 - (4/10)^2 \Rightarrow \alpha_1 = \frac{1}{10}, \alpha_2 = \frac{9}{10} \quad (1.101)$$

(recordem que els valors propis són precisament  $\lambda_k^2$  i per tant hem obtingut  $\lambda_1 = 1/\sqrt{10}$  i  $\lambda_2 = 3\sqrt{10}$ ) Els vectors propis associats són

$$|\Lambda\rangle_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad i \quad |\Lambda\rangle_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (1.102)$$

Per la simetria que mostra l'estat  $|\Psi\rangle$  és obvi que per  $\rho_B$  obtindrem els mateixos resultats i per tant podem escriure (exceptuant una fase relativa entre els dos termes):

$$|\Psi\rangle = \sum_{k=1}^2 \lambda_k |\Lambda\rangle_k |\Omega\rangle_k = \lambda_1 |\Lambda\rangle_1 |\Omega\rangle_1 + \lambda_2 |\Lambda\rangle_2 |\Omega\rangle_2 \quad (1.103)$$

d'on veiem que al haver-hi dos termes, l'estat  $|\Psi\rangle$  és entrellaçat.

Anem comprovar que el resultat que hem trobar és correcte expressant-lo en la base original

$$\begin{aligned}|\Psi\rangle &= \lambda_1 |\Lambda\rangle_1 |\Omega\rangle_1 + \lambda_2 |\Lambda\rangle_2 |\Omega\rangle_2 \\ &= \left(\frac{1}{\sqrt{10}}\right) \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) + \left(\frac{3}{\sqrt{10}}\right) \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \\ &= \frac{1}{\sqrt{10}} (2|00\rangle + |01\rangle + |10\rangle + 2|11\rangle)\end{aligned}\quad (1.104)$$

### 1.7.6 Matriu densitat

En primer lloc suposem que tenim un experimentador que genera els estats  $|S_z, \pm\rangle$  de forma aleatòria (amb 50% amunt i 50% avall) i que els envia a un observador, que desconeix aquest fet, per a que mesuri el valor esperat del spin en qualsevol direcció arbitrària  $\vec{n}$ . Quins són els valors que obtindrà?

Com que

$$\begin{aligned}\mathcal{P}_{|S_z+\rangle} \left( S_{\vec{n}} : +\frac{\hbar}{2} \right) &= |\langle S_{\vec{n}}, +|S_z+\rangle|^2 = \cos^2(\theta/2) \\ \mathcal{P}_{|S_z+\rangle} \left( S_{\vec{n}} : -\frac{\hbar}{2} \right) &= |\langle S_{\vec{n}}, -|S_z+\rangle|^2 = \sin^2(\theta/2) \\ \mathcal{P}_{|S_z-\rangle} \left( S_{\vec{n}} : +\frac{\hbar}{2} \right) &= |\langle S_{\vec{n}}, +|S_z-\rangle|^2 = \sin^2(\theta/2) \\ \mathcal{P}_{|S_z-\rangle} \left( S_{\vec{n}} : -\frac{\hbar}{2} \right) &= |\langle S_{\vec{n}}, -|S_z-\rangle|^2 = \cos^2(\theta/2)\end{aligned}\quad (1.105)$$

mesurarà un valor esperar 0 en qualsevol direcció doncs  $P(+\hbar/2) = (1/2)\cos^2(\theta/2)+(1/2)\sin^2(\theta/2) = 1/2 = P(-\hbar/2)$ .

En segon lloc suposem que tenim un procés que genera estats entrellaçats de dues partícules de spin 1/2 de la forma

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( |S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} - |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)} \right) \quad (1.106)$$

que, com ja sabem, són estats amb spin total 0 en qualsevol direcció. Aleshores, si hi ha un experimentador que desconeix que les partícules estan entrellaçades i que vol determinar l'estat de spin de la primera partícula, en mesurar-lo es trobarà la meitat amb  $+\hbar/2$  i l'altra amb  $-\hbar/2$  en qualsevol direcció que mesuri i per tant, mesurarà un valor esperat 0 en qualsevol direcció.

En ambdós casos, com que l'observador que fa la mesura troba un valor esperat 0 en qualsevol direcció, aquest es veurà incapaç de descriure l'estat de la partícula com un estat pur. És a dir, si considerem l'estat pur més general possible

$$|\phi\rangle = \frac{1}{\sqrt{|\alpha|^2 + |\beta|^2}} (\alpha|S_z, +\rangle + \beta|S_z, -\rangle) \quad (1.107)$$

no existeixen valors d' $\alpha$  i  $\beta$  que compleixin les equacions

$$\begin{aligned} \mathcal{P}_{|\phi\rangle} \left( S_{\vec{n}} : +\frac{\hbar}{2} \right) &= |\langle S_{\vec{n}}, +|\phi\rangle|^2 = \frac{1}{2} \\ \mathcal{P}_{|\phi\rangle} \left( S_{\vec{n}} : -\frac{\hbar}{2} \right) &= |\langle S_{\vec{n}}, -|\phi\rangle|^2 = \frac{1}{2} \end{aligned} \quad (1.108)$$

per a qualsevol valor del vector  $\vec{n}$ .

Aleshores, per descriure aquest dos tipus de sistemes s'utilitza l'anomenada matriu densitat. És a dir, utilitzarem la matriu densitat per fer una descripció parcial d'un sistema o per sistemes que es generen amb probabilitats clàssiques com en el primer exemple donat. Que ambdós casos la descripció es pugui fer per matrius densitats és degut a que estan relacionats. Per exemple, tornant a l'estat entrallaçat de l'equació 1.106, la mesura de qualsevol observable  $O$  sobre una de les parts, ens porta a un valor esperat de

$$\begin{aligned} \langle \psi | O \otimes I^{(2)} | \psi \rangle &= \frac{1}{2} \langle S_z, + |^{(1)} O | S_z, + \rangle^{(1)} + \frac{1}{2} \langle S_z, - |^{(1)} O | S_z, - \rangle^{(1)} \\ &= \mathcal{P}_{|\psi\rangle} \left( S_z^{(1)} : +\frac{\hbar}{2} \right) \langle S_z, + |^{(1)} O | S_z, + \rangle^{(1)} + \mathcal{P}_{|\psi\rangle} \left( S_z^{(1)} : -\frac{\hbar}{2} \right) \langle S_z, - |^{(1)} O | S_z, - \rangle^{(1)} \end{aligned} \quad (1.109)$$

que és el mateix resultat que trobaríem si tenim un sistema físic que barreja estats purs  $|S_z, \pm\rangle$  amb probabilitats  $p_{\pm} = \mathcal{P}_{|\phi\rangle}(S_z^{(1)} : \pm\hbar/2) = 1/2$ .

En les següents seccions definirem la matriu densitat  $\rho$  per estats purs, barreja i subsistemes d'estats-purs entrellaçats. En tots els casos veurem que es compleix:

$$\langle O \rangle = \text{Tr}(\rho O) \quad (1.110)$$

$$P_{\rho}(A : \lambda) = \text{Tr}(\rho \Pi_{\lambda}) \quad (1.111)$$

### Estats purs

Donat un estat pur  $|\psi\rangle$ , definirem la seva matriu densitat associada com

$$\rho = |\psi\rangle\langle\psi| \quad (1.112)$$

aleshores es compleix 1.110, doncs

$$\langle O \rangle = \langle \psi | O | \psi \rangle = \text{Tr}(|\psi\rangle\langle\psi|O) = \text{Tr}(\rho O) \quad (1.113)$$

on podem pensar que al fer la traça, resultat que sabem és independent de la base, hem agafat com a base una que té com un dels seus elements precisament  $|\psi\rangle$ .

Una altre forma de veure-ho és la següent. Si en una b.o. tenim  $|\psi\rangle = \sum_i a_i |i\rangle$  aleshores  $\langle O \rangle = \langle \psi | O | \psi \rangle = \sum_{i,j} a_i^* a_j \langle i | O | j \rangle$ . Com en el cas d'un estat pur  $\rho = |\psi\rangle\langle\psi| = \sum_{i,j} a_i a_j^* |i\rangle\langle j|$  aleshores comprovem que  $\text{Tr}(\rho O) = \sum_k \langle k | \rho O | k \rangle = \sum_{k,j} a_k a_j^* \langle j | O | k \rangle = \langle O \rangle$ .

També es compleix 1.111 doncs

$$P_{|\psi\rangle}(A : \lambda_i) = \| \Pi_i |\psi\rangle \|^2 = \langle \psi | \Pi_i | \psi \rangle = \text{Tr}(|\psi\rangle\langle\psi|\Pi_i) = \text{Tr}(\rho\Pi_i) \quad (1.114)$$

Com exemple considerem l'estat d'una partícula d'spin 1/2:  $|\psi\rangle = \alpha|+\rangle + \beta|-\rangle$ . Si mesurem el seu spin en direcció  $z$  tenim:

$$\begin{aligned} P_{|\psi\rangle}(S_z : +\hbar/2) &= \langle \psi | \Pi_+ | \psi \rangle = |\langle + | \psi \rangle|^2 = |\alpha|^2 \\ P_{|\psi\rangle}(S_z : -\hbar/2) &= |\beta|^2 \\ \langle S_z \rangle_{|\psi\rangle} &= (+\hbar/2)|\alpha|^2 + (-\hbar/2)|\beta|^2 \end{aligned} \quad (1.115)$$

Anem ara a comprovar aquests resultats amb el formalisme que acabem d'introduir de la matriu densitat. Sabem que:

$$\begin{aligned} \rho &= |\psi\rangle\langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix} \\ \Pi_+ &= |+\rangle\langle+| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \Pi_- = |-\rangle\langle-| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \quad (1.116)$$

i per tant

$$\begin{aligned} P_{|\psi\rangle}(S_z : +\frac{\hbar}{2}) &= \text{Tr}(\rho\Pi_+) = \text{Tr}\left[\begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right] = \text{Tr}\begin{pmatrix} |\alpha|^2 & 0 \\ 0 & 0 \end{pmatrix} = |\alpha|^2 \\ \langle S_z \rangle_{|\psi\rangle} &= \text{Tr}(\rho S_z) = \text{Tr}\left[\begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix} \begin{pmatrix} +\hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix}\right] = \frac{\hbar}{2} \text{Tr}\begin{pmatrix} |\alpha|^2 & 0 \\ 0 & -|\beta|^2 \end{pmatrix} \\ &= (+\hbar/2)|\alpha|^2 + (-\hbar/2)|\beta|^2 \end{aligned} \quad (1.117)$$

### Estats barreja

Suposem el cas d'un sistema físic que es prepara barrejant estats purs  $|\alpha\rangle$  amb probabilitats relatives  $p_\alpha$  (aquests estats estan normalitzats, però no són necessàriament ortogonals). En aquest cas definim la matriu densitat com

$$\rho \equiv \sum_\alpha |\alpha\rangle p_\alpha \langle\alpha|, \quad (1.118)$$

(notem que si l'estat és pur, això és un cas particular on totes les  $p_\alpha = 0$  exceptuant una que val evidentment 1).

Aquí també es compleix 1.110, doncs

$$\langle O \rangle = \sum_\alpha p_\alpha \langle \alpha | O | \alpha \rangle = \sum_{\alpha,n} p_\alpha \langle \alpha | O | n \rangle \langle n | \alpha \rangle = \sum_{\alpha,n} \langle n | \alpha \rangle p_\alpha \langle \alpha | O | n \rangle = \sum_n \langle n | \rho O | n \rangle = \text{Tr}(\rho O) \quad (1.119)$$

on  $|n\rangle$  és una base ortonormal completa.

Tambés es compleix 1.111, doncs

$$\begin{aligned} P_\rho(A : \lambda) &= \sum_{\alpha} p_{\alpha} P_{|\alpha\rangle}(A : \lambda) = \sum_{\alpha} p_{\alpha} \langle \alpha | \Pi_{\lambda} | \alpha \rangle = \sum_{\alpha} p_{\alpha} \text{Tr} [\Pi_{\lambda} | \alpha \rangle \langle \alpha |] \\ &= \text{Tr} \left[ \Pi_{\lambda} \left( \sum_{\alpha} p_{\alpha} | \alpha \rangle \langle \alpha | \right) \right] = \text{Tr} (\Pi_{\lambda} \rho) \end{aligned} \quad (1.120)$$

En el cas de fer una mesura filtrant, si sobreviu, l'estat barreja inicial esdevé també un estat barreja després de la mesura:

$$\rho = \sum_{\alpha} p_{\alpha} | \alpha \rangle \langle \alpha | \Rightarrow \rho^* = \frac{1}{\text{Tr}(\sum_{\lambda} \Pi_{\lambda} \rho \Pi_{\lambda})} \sum_{\lambda} \Pi_{\lambda} \rho \Pi_{\lambda} \quad (1.121)$$

Per simplicitat, anem a demostrar-ho en el cas que no hi ha degeneració, es a dir,  $\Pi_{\lambda} = |\lambda\rangle\langle\lambda|$ . En aquest cas

$$\begin{aligned} \sum_{\lambda} \Pi_{\lambda} \rho \Pi_{\lambda} &= \sum_{\lambda} |\lambda\rangle\langle\lambda| (\sum_{\alpha} p_{\alpha} | \alpha \rangle \langle \alpha |) |\lambda\rangle\langle\lambda| \\ &= \sum_{\lambda} \sum_{\alpha} p_{\alpha} |\langle \alpha | \lambda \rangle|^2 |\lambda\rangle\langle\lambda| = \sum_{\lambda} \text{Tr}(\Pi_{\lambda} \rho) |\lambda\rangle\langle\lambda| \end{aligned} \quad (1.122)$$

quedaria com una barreja dels estats que es filtreuen, com esperaríem. Com la suma de les probabilitats ha de ser 1, això explica el factor de 1.121.

Com exemple considerem l'estat barreja d'una partícula d'spin 1/2 amb 50% up i 50% down. Ja hem vist en la introducció d'aquesta secció que si mesurem el seu spin en direcció  $z$  tenim:  $P_{\rho}(S_z : +\hbar/2) = 1/2$  i  $\langle S_z \rangle = 0$ .

La matriu densitat per aquest estat barreja en la base  $\{|S_z, \pm\rangle\}$  és:

$$\rho = \frac{1}{2} |+\rangle\langle+| + \frac{1}{2} |-\rangle\langle-| = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \ 0) + \frac{1}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \ 1) = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.123)$$

L'expressió matricial de la matriu densitat evidentment depèn de la base escollida. Ara bé, per aquest estat barreja hem vist en la introducció d'aquesta secció que si mesurem el seu spin en qualsevol direcció  $\hat{n}$  tenim:  $P_{\rho}(S_{\hat{n}} : +\hbar/2) = P_{\rho}(S_{\hat{n}} : -\hbar/2) = 1/2$  i per tant  $\langle S_{\hat{n}} \rangle = 0$ . Aleshores l'expressió de la matriu densitat és sempre  $\rho = \frac{1}{2} I$  en qualsevol base  $\{|S_{\hat{n}}, \pm\rangle\}$  (de fet  $\rho = \frac{1}{2} |S_z, +\rangle\langle S_z, +| + \frac{1}{2} |S_z, -\rangle\langle S_z, -| = \frac{1}{2} |S_{\hat{n}}, +\rangle\langle S_{\hat{n}}, +| + \frac{1}{2} |S_{\hat{n}}, -\rangle\langle S_{\hat{n}}, -|$  per a qualsevol  $\hat{n}$ ).

Ara podem reproduir els resultats ja coneguts utilitzant el formalisme de la matriu densitat:

$$\begin{aligned} P_{\rho}(S_z : +\frac{\hbar}{2}) &= \text{Tr}(\rho \Pi_+) = \text{Tr} \left[ \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right] = \text{Tr} \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} \\ \langle S_z \rangle_{\rho} &= \text{Tr}(\rho S_z) = \text{Tr} \left[ \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \begin{pmatrix} +\hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} \right] = \frac{\hbar}{2} \text{Tr} \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} = 0 \end{aligned} \quad (1.124)$$

Al fer una mesura filtrant, si sobreviu, l'estat barreja inicial esdevé, en aquest cas, un estat pur. Per exemple, si filtrem up tenim

$$\rho \Rightarrow \rho^* = N \Pi_+ \rho \Pi_+ = N \Pi_+ \left( \frac{1}{2} |+\rangle\langle+| + \frac{1}{2} |-\rangle\langle-| \right) \Pi_+ = N \frac{1}{2} |+\rangle\langle+| = |+\rangle\langle+| \quad (1.125)$$

### Matriu densitat d'un subsistema.

La motivació principal per considerar matrius densitat és la consideració de sistemes i els seus subsistemes.

En la introducció d'aquesta secció hem considerat el cas d'un observador que rep estats  $|S_z \pm\rangle$  generats pel seu company que els barreja (genera) amb la mateixa probabilitat. Ara sabem que aquell observador pot descriure les seves observacions amb la matriu densitat

$$\rho = \frac{1}{2}|S_z, +\rangle\langle S_z, +| + \frac{1}{2}|S_z, -\rangle\langle S_z, -| \quad (1.126)$$

En la mateixa introducció hem considerat també un segon exemple, el d'un observador que fa mesures sobre una part de l'estat pur i entrelaçat  $\frac{1}{\sqrt{2}}(|+\rangle|-\rangle - |-\rangle|+\rangle)$  i hem vist que els resultats experimentals que observa són els mateixos que en el cas anterior. Per tant, també pot descriure les observacions que fa sobre una part del sistema total (un subsistema) utilitzant, en aquest cas, la mateixa matriu densitat. La pregunta que ens poden fer és si aquest formulació és generalitzable, és a dir, si les observacions sobre subsistemes sempre es poden descriure a través de matrius densitat. La resposta és si i primer donarem la recepta general per obtenir la matriu densitat associada a un subsistema a partir de l'estat total del sistema i després ho demostrarem.

Com ja hem dit, la motivació principal per considerar matrius de densitat és la consideració de sistemes i els seus subsistemes. Suposem que tenim dos sistemes quàntics, descrits pels espais de Hilbert  $H_1$  i  $H_2$ . El sistema compost és llavors el producte tensorial  $H_1 \otimes H_2$ . Generalment, encara que el sistema total estigui en un estat pur, els diversos subsistemes que el componen hauran de ser descrits per estats barreja (exemple anterior). Per tant, l'ús de matrius de densitat és inevitable. Si denotem per  $\rho_{12}$  la matriu de densitat del sistema compost de dos sistemes (si és un estat pur seria evidentment  $\rho_{12} = |\psi\rangle\langle\psi|$ ), aleshores veurem que l'estat del subsistema  $H_1(H_2)$  es pot descriure mitjançant l'anomenat operador de densitat reduïda  $\rho_1(\rho_2)$ , obtingut prenent la "traça parcial" de  $\rho_{12}$  utilitzant una base ortonormal  $|j_2\rangle\langle j_1\rangle$  de  $H_2(H_1)$

$$\begin{aligned} \rho_1 &= \sum_{j_2} \langle j_2 | \rho | j_2 \rangle = \text{Tr}_2 \rho_{12} \\ \rho_2 &= \sum_{j_1} \langle j_1 | \rho | j_1 \rangle = \text{Tr}_1 \rho_{12} \end{aligned} \quad (1.127)$$

Abans de fer la demostració anem a comprovar que dona el resultat esperat en l'exemple que hem considerant fins ara, on l'estat conjunt és un estat pur

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left( |S_z, +\rangle^{(1)} \otimes |S_z, -\rangle^{(2)} - |S_z, -\rangle^{(1)} \otimes |S_z, +\rangle^{(2)} \right) = \frac{1}{\sqrt{2}} (|+\rangle|-\rangle - |-\rangle|+\rangle) \quad (1.128)$$

la matriu densitat global serà

$$\begin{aligned} \rho_{12} &= |\psi\rangle\langle\psi| \\ &= \frac{1}{2} |+\rangle|-\rangle\langle+| \langle-| + \frac{1}{2} |-\rangle|+\rangle\langle-| \langle+| - \frac{1}{2} |+\rangle|-\rangle\langle-| \langle+| - \frac{1}{2} |-\rangle|+\rangle\langle+| \langle-| \end{aligned} \quad (1.129)$$

Si ara com a base de  $H_2$  agafem  $\{|S_z, \pm\rangle^{(2)}\}$ , tindrem

$$\begin{aligned} \rho_1 &= \text{Tr}_2 \rho_{12} = \sum_{j_2} \langle j_2 | \rho_{12} | j_2 \rangle = \sum_{j_2} \langle j_2 | (|\psi\rangle\langle\psi|) | j_2 \rangle \\ &= \langle S_z, + |^{(2)} | (|\psi\rangle\langle\psi|) | S_z, + \rangle^{(2)} + \langle S_z, - |^{(2)} | (|\psi\rangle\langle\psi|) | S_z, - \rangle^{(2)} \\ &= \frac{1}{2} |S_z, -\rangle\langle S_z, -| + \frac{1}{2} |S_z, +\rangle\langle S_z, +| \end{aligned} \quad (1.130)$$

que reproduceix el resultat esperat.

Anem a fer el mateix exemple però en notació matricial. Escollint  $\{|++, |+-\}, |-+\}, |--\}\}$  la base del sistema, la matriu densitat global  $\rho_{12}$  la podem expressar com (notació  $\langle ab|\rho_{12}|cd\rangle = \rho_{abcd}$ )

$$\rho_{12} = \begin{pmatrix} \rho_{++++} & \rho_{+++-} & \rho_{+-+-} & \rho_{+---} \\ \rho_{+-+-} & \rho_{+-+-} & \rho_{+-+-} & \rho_{+-+-} \\ \rho_{-+++} & \rho_{-+++} & \rho_{-+++} & \rho_{-+++} \\ \rho_{---+} & \rho_{---+} & \rho_{---+} & \rho_{---+} \end{pmatrix} = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & \rho_{44} \end{pmatrix} \quad (1.131)$$

i per tant

$$\begin{aligned} \rho_{12} &= \frac{1}{2}|+\rangle|-\rangle\langle+|\langle-| + \frac{1}{2}|-\rangle|+\rangle\langle-|\langle+| - \frac{1}{2}|+\rangle|-\rangle\langle-|\langle+| - \frac{1}{2}|-\rangle|+\rangle\langle+|\langle-| \\ &= \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (1.132)$$

aleshores (els punts indiquen sumar sobre index iguals al ser una traça)

$$\begin{aligned} \rho_1 &= Tr_2(\rho_{12}) = \begin{pmatrix} \rho_{++} & \rho_{+-} \\ \rho_{-+} & \rho_{--} \end{pmatrix} = \begin{pmatrix} \rho_{11} + \rho_{22} & \rho_{13} + \rho_{24} \\ \rho_{31} + \rho_{42} & \rho_{33} + \rho_{44} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \rho_2 &= Tr_1(\rho_{12}) = \begin{pmatrix} \rho_{++} & \rho_{+-} \\ \rho_{-+} & \rho_{--} \end{pmatrix} = \begin{pmatrix} \rho_{11} + \rho_{33} & \rho_{12} + \rho_{34} \\ \rho_{21} + \rho_{43} & \rho_{22} + \rho_{44} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \quad (1.133)$$

En el cas general, si expressem  $\rho_{12}$  en funció de bases ortonormals pels dos espais  $H_1$  i  $H_2$ , l'expressió més general per a una matriu densitat global serà

$$\rho_{12} = \sum_{i_1, j_1, i_2, j_2} a_{i_1, i_2, j_1, j_2} |i_1\rangle\langle i_2| |j_1\rangle\langle j_2| \quad (1.134)$$

i la traça parcial sobre  $H_2$  ens porta a

$$\rho_1 = Tr \rho_{12} = \sum_{k_2} \langle k_2 | \rho_{12} | k_2 \rangle = \sum_{i_1, j_1} \left( \sum_{k_2} a_{i_1, k_2, j_1, k_2} \right) |i_1\rangle\langle j_1| \quad (1.135)$$

Si l'estat de  $H_1 \otimes H_2$  és una matriu de densitat de la forma especial  $\rho_{12} = \rho_1 \otimes \rho_2$ , llavors la traça parcial de  $\rho_{12}$  respecte a  $H_2$  és només  $\rho_1$ . Tanmateix, generalment  $\rho_{12}$  no serà d'aquesta forma.

Anem finalment a comprovar que en una mesura sobre un subsistema també es compleix que  $\langle O_1 \rangle = Tr(\rho_1 O_1)$ . Efectivament, utilitzant el fet que sobre el sistema global sabem que es compleix  $\langle O \rangle = Tr_{12}(\rho_{12} O)$  tenim

$$\begin{aligned} \langle O_1 \rangle &= \langle O_1 \otimes I_2 \rangle = Tr_{12}(\rho_{12}(O_1 \otimes I_2)) \\ &= \sum_{k_1, k_2} \langle k_1 | \langle k_2 | (\rho_{12} O_1 \otimes I_2) | k_1 \rangle | k_2 \rangle = \sum_{k_1, k_2, k'_1, k'_2} \langle k_1 | \langle k_2 | (\rho_{12} |k'_1\rangle\langle k'_2|) | k'_1 \rangle | k'_2 \rangle | O_1 \otimes I_2 \rangle | k_1 \rangle | k_2 \rangle \\ &= \sum_{k_1, k_2, k'_1} \langle k_1 | \langle k_2 | \rho_{12} | k'_1 \rangle | k_2 \rangle | k'_1 \rangle | O_1 | k_1 \rangle = \sum_{k_1, k'_1} \langle k_1 | \rho_1 | k'_1 \rangle | k'_1 \rangle | O_1 | k_1 \rangle \\ &= \sum_{k_1} \langle k_1 | \rho_1 O_1 | k_1 \rangle = Tr_1(\rho_1 O_1) \end{aligned} \quad (1.136)$$

on hem utilitzat la definició, que ara veiem motivada, de  $\rho_1 = \sum_{k_2} \langle k_2 | \rho_{12} | k_2 \rangle = Tr_2(\rho_{12})$  com la traça parcial de  $\rho_{12}$  sobre  $H_2$  i per tant obtenim el resultat esperat.

Acabem de demostrar doncs que

$$\langle O_1 \rangle = \langle O_1 \otimes I_2 \rangle = Tr_{12}(\rho_{12}(O_1 \otimes I_2)) = Tr_1(\rho_1 O_1) \quad (1.137)$$

Anem a tornar a demostrar la relació anterior en el cas particular on  $\rho_{12}$  sigui un estat pur i que l'expresssem en funció de bases ortonormals per els dos espais  $H_1$  i  $H_2$

$$|\psi\rangle = \sum_{i_1, i_2} a_{i_1, i_2} |i_1\rangle |i_2\rangle \quad (1.138)$$

aleshores el valor esperar de  $O_1$  és

$$\begin{aligned} \langle O_1 \rangle &= \langle \psi | O_1 \otimes I_2 | \psi \rangle = \sum_{i_1, i_2, j_1, j_2} a_{i_1, i_2}^* \langle i_1 | \langle i_2 | (O_1 \otimes I_2) | j_1 \rangle | j_2 \rangle a_{j_1, j_2} \\ &= \sum_{i_1, j_1, k_2} a_{i_1, k_2}^* a_{j_1, k_2} \langle i_1 | O_1 | j_1 \rangle \end{aligned} \quad (1.139)$$

com en aquest cas

$$\rho_{12} = |\psi\rangle\langle\psi| = \sum_{i_1, i_2, j_1, j_2} a_{i_1, i_2} a_{j_1, j_2}^* |i_1\rangle\langle i_2| |j_1\rangle\langle j_2| \quad (1.140)$$

i per tant  $\rho_1 = \sum_{i_1, j_1, k_2} a_{i_1, k_2} a_{j_1, k_2}^* |i_1\rangle\langle j_1|$ , aquest resultat és idèntic a fer

$$\langle O_1 \rangle = Tr(\rho_1 O_1) = \sum_{k_1} \langle k_1 | \rho_1 O_1 | k_1 \rangle = \sum_{k_1, j_1, k_2} a_{k_1, k_2} a_{j_1, k_2}^* \langle j_1 | O_1 | k_1 \rangle \quad (1.141)$$

que és el que havíem trobat intercanviant  $k_1, j_1 \rightarrow j_1, i_1$ .

### Puresa d'un estat

Volem trobar una quantitat que pugui determinar si un estat és pur o barreja i en el darrer cas, "mesurar" d'alguna manera la quantitat de "barreja" (no és el mateix barrejar spin up i down al 50% i 50% que fer-ho al 99% i 1%).

Quan val  $\rho^2$ ? Com els seus valors propis han de ser positius ( $\rho$  és un operador autoadjunt i per tant amb valors propis reals) tenim que  $Tr(\rho^2) \geq 0$ . Veiem els seus límits. Per un estat pur  $|\psi\rangle$  tenim que  $\rho^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = \rho$  i per tant  $Tr(\rho^2) = 1$ . En el cas de màxima ambigüïtat, és a dir, màximament barrejat, tindrem que la probabilitat per associada a cada possible estat serà  $1/d$  (assumint que barregem d'estat purs) en aquest cas que trobem que  $Tr(\rho^2) = d(1/d)^2 = 1/d$  establint així un límit inferior. Per tant

$$\frac{1}{d} \leq Tr(\rho^2) \leq 1 \quad (1.142)$$

acabem de trobar una magnitud que ens indica com de barrejat està un estat. Així, a  $Tr(\rho^2)$  l'anomenarem puresa.

### Exemples d'estats purs o barreja amb entrellaçament o no

Suposem que tenim dos subsistemes (1) i (2). Un estat no estarà entrellaçat si la preparació en (1) és "independent" de la preparació en (2). Per tant en aquests casos

$$\rho_{12} = \rho_1 \otimes \rho_2$$

Ara bé, podem admetre correlacions "clàssiques" en la preparació dels estats (quan (1) genera  $\rho_1^{(x)}$ , (2) genera  $\rho_2^{(x)}$ ) i en aquest cas els estats tampoc estan entrellaçats

$$\rho = \sum_i p_i \rho_1^{(i)} \otimes \rho_2^{(i)}$$

Per tant, en general per sistemes composts, anomenarem estats entrellaçats aquells que no es poden escriure de la forma

$$\rho_{12} \neq \sum_i p_i \rho_1^{(i)} \otimes \rho_2^{(i)} \quad (1.143)$$

és a dir, no s'ha fet una preparació clàssica del estat independent en cada subsistema.

Determinar si un estat global pur està o no entrellaçat és fàcil. Sabem per la descomposició d'Schmidt que dos subsistemes que particionen un estat pur estan entrallaçats si i només si els seus estats reduïts són estats barreja. Com exemple considerem l'estat pur de dos qubits (utilitzarem la base habitual)

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|+\rangle^{(1)} \otimes |-\rangle^{(2)} - |-\rangle^{(1)} \otimes |+\rangle^{(2)}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \Rightarrow \rho_{12} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Confirmem que és estat pur veient que  $\rho^2 = \rho$  i per tant la seva puresa val 1. Efectivament

$$\rho_{12}^2 = \frac{1}{4} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \rho_{12} \Rightarrow \text{Tr}(\rho_{12}^2) = 1$$

Per altra banda sabem que les matrius densitat reduïdes d'ambdós subsistemes són

$$\rho_i = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad i = 1, 2$$

que representen estats barreja doncs  $\text{Tr}(\rho_i^2) = 1/2 < 1$  i per tant, com l'estat global és pur, aquest és un estat entrellaçat. Notem que

$$\rho_1 \otimes \rho_2 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \neq \rho_{12}$$

mentre  $\rho_{12}$  representa un estat entrellaçat,  $\rho_1 \otimes \rho_2$  no ho és.

Si el sistema total no és pur, determinar si està entrellaçat és més complicat. Per exemple, suposem un sistema no entrellaçat quànticament però correlat clàssicament: (1) i (2) comparteixen una informació clàssica que determina si generen el seu estat d'una forma una altre (amb probabilitat  $p$ ) (1) genera un estat totalment barrejat (50% up i 50% down) mentre que (2) l'estat pur  $|+\rangle$  i a amb probabilitat  $(1-p)$  (1) genera  $|+\rangle$  mentre que (2) un estat totalment barrejat. En aquest cas la matriu densitat total serà:

$$\begin{aligned} \rho_{12} &= p \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + (1-p) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \frac{p}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{1-p}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (1.144)$$

Confirmem que l'estat total és barreja doncs  $\text{Tr}(\rho_{12}^2) = 1/4(1 + (1-p)^2 + p^2) < 1$ , però per determinar que està entrellaçat, hem d'assegurar-nos que no el podem expressar com  $\rho_{12} \neq \sum_i p_i \rho_1^{(i)} \otimes \rho_2^{(i)}$  (fet que sabem que si es pot fer doncs ho hem generat d'aquest forma, i per tant no és entrellaçat).

Per l'estat que estem estudiant, podem determinar les matrius densitat reduïdes d'ambdós subsistemes fent la traça parcial o bé calculant-les a partir de la forma que sabem s'han generat en cada subsistema. Fent les traces parcials tenim:

$$\begin{aligned}\rho_{12} &= \frac{1}{2}|++\rangle\langle++| + \frac{1-p}{2}|+-\rangle\langle+-| + \frac{p}{2}|-+\rangle\langle-+| \\ \rho_1 &= Tr_2(\rho_{12}) = \frac{1}{2}|+\rangle\langle+| + \frac{p}{2}|-\rangle\langle-| + \frac{1-p}{2}|+\rangle\langle+| = \left(1 - \frac{p}{2}\right)|+\rangle\langle+| + \frac{p}{2}|-\rangle\langle-| \\ \rho_2 &= Tr_1(\rho_{12}) = \frac{1}{2}|+\rangle\langle+| + \frac{1-p}{2}|-\rangle\langle-| + \frac{p}{2}|+\rangle\langle+| = \left(\frac{1+p}{2}\right)|+\rangle\langle+| + \frac{1-p}{2}|-\rangle\langle-|\end{aligned}\quad (1.145)$$

mentre que sabent com estan generats, tenim

$$\begin{aligned}\rho_1 &= p\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (1-p)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1-p/2 & 0 \\ 0 & p/2 \end{pmatrix} \\ \rho_2 &= (1-p)\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + p\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} (1+p)/2 & 0 \\ 0 & (1-p)/2 \end{pmatrix}\end{aligned}\quad (1.146)$$

obtenint evidentment els mateixos resultats.

Veiem també que

$$\rho_1 \otimes \rho_2 = \frac{1}{2}\begin{pmatrix} 2-p & 0 \\ 0 & p \end{pmatrix} \otimes \frac{1}{2}\begin{pmatrix} 1+p & 0 \\ 0 & 1-p \end{pmatrix} = \frac{1}{4}\begin{pmatrix} (2-p)(1+p) & 0 & 0 & 0 \\ 0 & (2-p)(1-p) & 0 & 0 \\ 0 & 0 & p(1+p) & 0 \\ 0 & 0 & 0 & p(1-p) \end{pmatrix} \neq \rho_{12}$$

doncs encara que tan  $\rho_{12}$  com  $\rho_1 \otimes \rho_2$  representen estats barreja, generalment es compleix que

$$\left(\sum_i p_i \rho_i^{(A)}\right) \otimes \left(\sum_i p_i \rho_i^{(B)}\right) \neq \sum_i p_i \rho_i^{(A)} \otimes \rho_i^{(B)}$$

Com a resum, donem cinc exemples de diferents tipus d'estats atenent a les possibles correlacions clàssiques o quàntiques (entrellaçament) entre els seus subsistemes:

1. Estat pur no entrellaçat

$$|\psi\rangle = |++\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \rho = |\psi\rangle\langle\psi| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

2. Estat pur entrellaçat

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \rho = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

3. Estat barreja no correlat (1: 50% up 50% down, 2: 50% up 50% down )

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

4. Estat barreja correlat clàssicament (1 i 2: 50% up-down, 1 i 2: 50% down-up)

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

5. Estat barreja entrellaçat (combinació de 2, pur entrellaçat, amb 3, barreja)

$$\rho = p \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + (1-p) \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

### Evolució temporal de la matriu densitat

Per trobar l'evolució temporal de la matriu densitat podem fer

$$i\hbar \frac{\partial}{\partial t} \rho = i\hbar \frac{\partial}{\partial t} \sum_{\alpha} p_{\alpha} |\alpha\rangle\langle\alpha| = \sum_{\alpha} p_{\alpha} \{H|\alpha\rangle\langle\alpha| - |\alpha\rangle\langle\alpha|H\} = [H, \rho] \quad (1.147)$$

i, per tant, l'equació d'evolució és

$$\frac{\partial \rho}{\partial t} = -\frac{1}{i\hbar} [\rho, H]. \quad (1.148)$$

que és l'equació de Heisenberg per operadors que no depenen explícitament del temps, com és el cas de  $\rho$ . Per tant, hem obtingut que

$$\rho(t) = U(t, t_0)^{\dagger} \rho(t_0) U(t, t_0) \quad (1.149)$$

on  $U$  és el coneut operador d'evolució temporal.

### 1.7.7 L'esfera de Bloch

Qualsevol estat pur d'un qbit serà una combinació del tipus

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \quad (1.150)$$

Com la fase global és irrelevante, la fixem escollint que el coeficient de  $|0\rangle$  sigui real, fet que determina unívocament els angles  $\theta$  i  $\varphi$ . Per tant, qualsevol estat pur d'un qbit es pot representar per un punt en l'esfera de radi unitat (exemple, el punt blau en la figura 1.25), anomenada esfera de Bloch.

La matriu densitat associada al estat pur anterior és

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} \cos^2 \frac{\theta}{2} & \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{-i\varphi} \\ \cos \frac{\theta}{2} \sin \frac{\theta}{2} e^{i\varphi} & \sin^2 \frac{\theta}{2} \end{pmatrix} \quad (1.151)$$

i evidentment es compleix que  $\rho^2 = \rho \Rightarrow Tr(\rho^2) = 1$ , és a dir, la seva puresa val 1.

En general, tot estat d'un qbit, tant pur com barreja, es pot escriure amb la matriu densitat

$$\rho = \frac{1}{2} + \frac{\vec{b} \cdot \vec{\sigma}}{2} = \frac{1}{2} \begin{pmatrix} 1 + \gamma & \alpha - \beta i \\ \alpha + \beta i & 1 - \gamma \end{pmatrix} \quad (1.152)$$

amb  $\vec{b} = (\alpha, \beta, \gamma)$  i  $\vec{b}^2 \leq 1$ . Veiem que  $Tr(\rho) = 1$ . Un punt  $\vec{b}$  on  $\vec{b}^2 = 1$  és un punt de la superfície de l'esfera de Bloch i representa un estat pur on precisament  $\vec{b}$  és la direcció del qbit

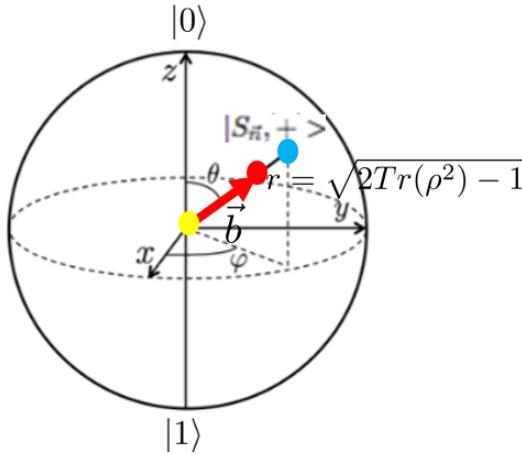


Figure 1.25: Esfera de Bloch

pur (exemple, el punt blau en la figura 1.25). El punt  $\vec{b}$  on  $0 \leq \vec{b}^2 < 1$  és un punt interior de l'esfera de Bloch i representa un estat barreja (exemple, el punt vermell en la figura 1.25), sent l'estat de màxima barreja el corresponent a  $\vec{b} = 0$  doncs, com sabem, és aquell on  $\rho = 1/2I$  (el punt groc en la figura 1.25).

Anem a comprovar que la puresa d'un qbit està acotada entre  $1/2 \leq Tr(\rho^2) \leq 1$  com és d'esperar. Per això calculem

$$\rho^2 = \left( \frac{1 + \vec{b} \cdot \vec{\sigma}}{2} \right)^2 = \frac{1}{4} (1 + 2\vec{b} \cdot \vec{\sigma} + (\vec{b} \cdot \vec{\sigma})^2) = \frac{1}{4} + \frac{1}{2}\vec{b} \cdot \vec{\sigma} + \frac{1}{4}b_i b_j \sigma_i \sigma_j = \frac{1}{4}(1 + \vec{b}^2) + \frac{1}{2}\vec{b} \cdot \vec{\sigma} \quad (1.153)$$

on en el darrer pas hem utilitzat el fet de que  $\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$  i com està multiplicat pel terme simètric  $b_i b_j$  el terme antisimètric no contribueix. Per tant, la puresa d'un qbit,  $Tr(\rho^2) = (1 + \vec{b}^2)/2$ , depèn de la distància entre el punt que determina l'estat del qbit i el centre de l'esfera. Aquesta puresa serà mínima (màxima) quant  $|\vec{b}| = 0(1)$ , amb el valor  $1/2(1)$ , que correspon a l'estat de màxima barreja (un estat pur).

Una combinació d'estat purs  $\vec{b}_i$  (amb  $\vec{b}_i^2 = 1$ ) amb probabilitats  $p_i$ , dona l'estat barreja de punt  $\vec{b} = \sum_i p_i \vec{b}_i$ . Per exemple, considerem l'estat d'un qbit barreja de l'estat pur  $|0\rangle$  (de  $\vec{b} = (0, 0, 1)$ ) amb probabilitat  $p$  i de l'estat pur  $|1\rangle$  (de  $\vec{b} = (0, 0, -1)$ ) amb probabilitat  $1 - p$ . La seva matriu densitat és

$$\rho = p|0\rangle\langle 0| + (1 - p)|1\rangle\langle 1| = \begin{pmatrix} p & 0 \\ 0 & 1 - p \end{pmatrix}$$

i aquesta matriu és precisament la del qbit que es troba en el punt  $\vec{b} = p(0, 0, 1) + (1 - p)(0, 0, -1) = (0, 0, 2p - 1)$  doncs

$$\rho = \frac{1}{2} + \frac{\vec{b} \cdot \vec{\sigma}}{2} = \frac{1}{2} + \frac{(2p - 1)}{2} \sigma_z = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} + \begin{pmatrix} p - 1/2 & 0 \\ 0 & 1/2 - p \end{pmatrix} = \begin{pmatrix} p & 0 \\ 0 & 1 - p \end{pmatrix}$$

Per aquestes combinacions tenim que

$$\rho^2 = p^2|0\rangle\langle 0| + (1 - p)^2|1\rangle\langle 1| = \begin{pmatrix} p^2 & 0 \\ 0 & (1 - p)^2 \end{pmatrix}$$

i per tant la seva puresa val  $Tr(\rho^2) = p^2 + (1 - p)^2 = 2p(p - 1) + 1$ . Com la puresa depèn de la distància al centre, per aquests estats barreja generats amb la probabilitat  $p$  per  $|0\rangle$  i

probabilitat  $1 - p$  per  $|1\rangle$ , tenim la relació

$$|\vec{b}| = \sqrt{4p(p-1) + 1}$$

### 1.7.8 Informació quàntica: entropia de Von Neumann

Suposem un dispositiu (clàssic) que pot trobar-se en dos estats  $(0, 1)$  amb les mateixes probabilitats  $p_0 = p_1 = 1/2$ . Aleshores sabem que la seva entropia (informació clàssica) és  $S = -2 \times (1/2) \log_2(1/2) = 1$  bit.

Suposem ara un dispositiu que pot guardar un qbit, per exemple un spin en qualsevol direcció amb les mateixes probabilitats. Sabem que aquest estat és un estat de màxima barreja que es pot representar per la matriu densitat

$$\rho = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| \quad (1.154)$$

que té probabilitat  $1/2$  de trobar-se en cada un dels estats  $|\pm\rangle$ . Aleshores un podria dir que, com en el cas anterior, l'entropia hauria de ser  $S = -2 \times (1/2) \log_2(1/2) = 1$  bit. Però recordem que també té probabilitat  $1/2$  de trobar-se en  $|S_{\hat{n}}, \pm\rangle$  i això és cert en qualsevol direcció. Per tant, un qbit de informació ve implementat per un dispositiu que pot estar en qualsevol punt de la superfície de l'esfera de Bloch amb la mateixa probabilitat. Estem davant d'un dispositiu que pot "guardar" molt més d'un bit (de fet infinitis bits, un bit en cada direcció espacial, fet que esta relacionat amb el teorema de no teleportació: no podem reduir la informació que compte un qbit en un número finit de bits).

Per copsar aquesta diferència es defineix l'entropia de Von Neumann, o informació quàntica, com

$$S = -Tr(\rho \log_2 \rho) \quad (1.155)$$

i que dona el resultat en qbits, no en bits.

Veiem que dona en l'exemple anterior. Treballant en la base  $|\pm\rangle$  tenim

$$\rho = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.156)$$

i per tant

$$S = -Tr \left( \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \log_2 \frac{1}{2} & 0 \\ 0 & \log_2 \frac{1}{2} \end{pmatrix} \right) = 1 \text{ qbit} \quad (1.157)$$

com cabria esperar.

Veiem que l'expressió que l'entropia Von Neumann es pot calcular en qualsevol base (recordem que la traça és cíclica i per tant podem afegir termes  $CC^{-1}$ , on  $C$  es la matriu de canvi de base, davant i darrera de cada potència de  $\rho$ , fet que deixa el resultat de la traça igual). Per tant, la base més adient per fer el calcul és la que diagonalitza  $\rho$ . Siguin  $\{\lambda_1, \lambda_2, \dots, \lambda_d\}$  tots els autovalors resultants de diagonalitzar la matriu densitat, aleshores

$$S = - \sum_i^d \lambda_i \log_2 \lambda_i \quad (1.158)$$

El valor de l'entropia de Von Neumann per un estat pur, com la seva matriu densitat associada es tal que tots els  $p_i = 0$  exceptuant l'associat al estat pur, és  $S(\rho) = 0\log(0) + 0\log(0) + \dots + 1\log(1) + \dots + 0\log(0) = 0$ . En el cas de la matriu de densitat amb el màxima barreja  $S(\rho) = -\sum_i^d(i/d) \ln(1/d) = \ln d$ .

Veiem que un estat quàntic conté potencialment una gran quantitat d'informació que és inaccessible per a un observador extern. Aquesta informació inaccessible és informació quàntica (o entropia de Von Neumann). La quantitat d'informació que podem extreure d'un estat quàntic mitjançant la mesura és la informació accessible i aquesta és informació clàssica.

### Entropia de Von Neumann per un qbit en estat barreja

Hem vist que un qbit que es trobi en un estat pur té 0 entropia, mentre que un que es trobi en l'estat de màxima barreja té 1 qbit d'entropia (o informació quàntica).

Veiem ara la seva entropia en el cas general de que es trobi en l'estat donat per la matriu densitat de 1.152. Els valors propis d'aquesta matriu densitat són

$$1 \pm r \text{ amb } r = |\vec{b}| = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$$

Per tant la seva entropia de Von Neumann és

$$S = -\frac{1+r}{2} \log_2 \left( \frac{1+r}{2} \right) - \frac{1-r}{2} \log_2 \left( \frac{1-r}{2} \right)$$

Veiem que  $(1+r)/2$  juga el mateix paper que  $p$  en l'entropia d'un bit. Per tant, sabem que varia entre  $1/2$  (per  $r=0$ , estat de màxima barreja, equivalent a un bit amb  $p=1/2$ ) i  $1$  (per  $r=1$ , estat pur, equivalent a un bit amb  $p=1$ ).

Sabent que la seva puresa és  $P = \text{Tr}(\rho^2) = (1+r^2)/2$ , podem relacionar l'entropia de Von Neumann d'un qbit amb la seva puresa:

$$S(P) = -\frac{1+\sqrt{2P-1}}{2} \log_2 \left( \frac{1+\sqrt{2P-1}}{2} \right) - \frac{1-\sqrt{2P-1}}{2} \log_2 \left( \frac{1-\sqrt{2P-1}}{2} \right)$$

### Propietats

Algunes propietats de l'entropia de von Neumann:

- $S(\rho)$  és zero si i només si  $\rho$  representa un estat pur.
- $S(\rho)$  és màxim i és igual a  $\log_2 d$  per a un estat màximament barrejat, sent  $d$  la dimensió de l'espai de Hilbert.
- $S(\rho)$  és invariant per canvis en la base de  $\rho$ , és a dir,  $S(\rho) = S(U\rho U^\dagger)$ , amb  $U$  una transformació unitària.
- $S(\rho)$  és còncau, és a dir, es dóna una col·lecció de nombres positius  $\lambda_i$  que sumen a la unitat tenim

$$S\left(\sum_{i=1}^k \lambda_i \rho_i\right) \geq \sum_{i=1}^k \lambda_i S(\rho_i) \quad (1.159)$$

- $S(\rho)$  és additiu per a sistemes independents. Donades dues matrius de densitat  $\rho_A, \rho_B$  que descriuen sistemes independents A i B, tenim

$$S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B) \quad (1.160)$$

$$\text{doncs } \text{Tr}(f(A \otimes B)) = \text{Tr}(f(A) \otimes f(B)) = \text{Tr}(f(A)) + \text{Tr}(f(B))$$

- $S(\rho)$  és subadditiu per a qualsevol sistema A, i B:

$$0 \leq S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B) \quad (1.161)$$

Tot i que a la teoria de Shannon l'entropia d'un sistema compost mai pot ser inferior a l'entropia de CAP de les seves parts ( $S(\rho_X) \leq S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$  amb  $X = A, B$ ), en teoria quàntica no és així, és a dir, és possible que  $S(\rho_{AB}) = 0$ , mentre que  $S(\rho_A) = S(\rho_B) > 0$ .

Intuïtivament, això es pot entendre de la següent manera: En mecànica quàntica, l'entropia del sistema total pot ser inferior a la suma de l'entropia dels seus components perquè els components es poden entrellaçar. Per exemple, tal com es veu explícitament, l'estat de Bell de dos spin 1/2

$$|\psi\rangle = \frac{1}{\sqrt{2}}|(\uparrow\downarrow) + |\downarrow\uparrow\rangle) \quad (1.162)$$

és un estat pur amb entropia nul·la, però cada spin té l'entropia màxima (1/2) quan es considera individualment la seva matriu de densitat reduïda. L'entropia d'un spin es pot "cancel·lar" correlacionant-se quàticament (entrellaçament) amb l'entropia de l'altre. La desigualtat de l'esquerra es pot interpretar aproximadament com a dir que l'entropia només es pot cancel·lar amb una quantitat d'entropia igual.

Si el sistema A i el sistema B tenen diferents quantitats d'entropia, el més petit només pot cancel·lar parcialment el més gran i s'ha de deixar una mica d'entropia. De la mateixa manera, la part dreta de 1.161 es pot interpretar com a dir que l'entropia d'un sistema compost es maximiza quan els seus components no es correlacionen, en aquest cas l'entropia total és només una suma de les subentropies.

### Conservació de la informació quàntica

Sigui un sistema que inicialment es troba en l'estat  $\rho(t_0)$ . Com l'evolució de la matriu densitat ve donada per l'equació 1.149,  $\rho(t) = U(t, t_0)^\dagger \rho(t_0) U(t, t_0)$ , i com  $U$  és un operador unitari, aleshores degut a la tercera propietat descrita anteriorment tindrem

$$S(\rho(t)) = S(\rho(t_0)) \quad (1.163)$$

es a dir, l'entropia de Von Neumann es manté durant l'evolució temporal.

Per tant, l'entropia zero d'un estat pur es manté a zero durant la seva evolució temporal (l'equivalent en mecànica clàssica es conèixer l'estat inicial amb infinita precisió, i per tant l'entropia o informació promig és zero, i aquesta entropia es manté zero durant la seva evolució doncs la coneixem perfectament donat  $H$ ). Igualment, en un dispositiu que pugui emmagatzemar un qbit, el desconeixement del seu estat el representem per un estat de màxima barreja, i per tant l'entropia es diferent de zero, concretament 1 qbit, i aquesta és manté durant la seva evolució unitària. L'equivalent clàssic és un dispositiu que pugui emmagatzemar un bit, que pot estar en dos diferents estats inicials, cada un coneguts amb infinita precisió i amb probabilitat  $p_i = 1/2$ , i per tant l'entropia és diferent de zero, concretament 1 bit. L'evolució mantindrà aquestes probabilitats i per tant entropia, si no hi ha interacció entre els dos estats possibles i tots dos es mouen sota el mateix potencial (que no es el cas en certes operacions en ordinadors clàssics).

Si considerem l'univers sencer com un sistema tancat, la quantitat total de informació quàntica continua sent la mateixa (obviant evidentment el problema de la mesura i el seu apparent trencament de l'evolució unitària). Al món clàssic, la informació es pot copiar i eliminar perfectament. En el món quàntic, però, la conservació de la informació quàntica ha de significar que la informació no es pot crear ni destruir. Aquest concepte prové de dos teoremes fonamentals de l'evolució unitària de la mecànica quàntica: el teorema de la no clonació (que veuren en la següent subsecció) i el teorema de la no eliminació (procés invers de l'anterior).

També veurem com teleportar un estat quàntic arbitrari, fent-lo desapareixer d'un lloc, on el destruïm fent una mesura sobre ell, i fent-lo aparèixer en un altre lloc. No s'ha de confondre aquest proces amb l'anomenat teorema de no-teleportació que ens diu que no podem convertir la totalitat de la informació quàntica en clàssica. Aquest és un exemple on utilitzem la nostra

interpretació de la mesura com un trencament de l'evolució unitària i utilitzem la comunicació de informació clàssica per mantenir la informació quàntica (l'estat final és exactament l'estat inicial), encara que globalment, admetent que tota evolució és unitaria, la informació mai hauria canviat.

Però el teorema de No-hiding que veurem al final de la secció és la prova definitiva de la conservació de la informació quàntica. El teorema demostra que si l'amplitud de probabilitat desapareix d'un sistema, reapareixerà en un altre sistema. Atès que la funció d'ona conté tota la informació rellevant sobre un sistema físic, aquesta conservació de la funció d'ona equival a la conservació de la informació quàntica.

### 1.7.9 Teorema de no clonació

En estudiar l'experiment de Stern-Gerlach hem introduït el concepte de *preparació* d'un sistema físic en un cert estat. Quan tenim coneixement de l'estat quàntic del sistema, podem fer totes les prediccions probabilístiques que vulguem sobre els resultats de mesures que s'hi realitzin. Però en molts casos no coneixem l'estat. Per exemple, en el mateix experiment de Stern-Gerlach, si hi entra un sol àtom d'argent i en sortir impacta a la zona superior de la pantalla, què podrem dir sobre l'estat quàntic de l'àtom incident? Poca cosa. Ara bé, si tinguéssim tantes còpies com vulguéssim d'un estat incident  $|\phi\rangle$ , llavors el resultat d'un gran nombre d'impacts ens permetria reconstruir amb la precisió que desitgéssim els valors d' $|a_+\rangle$  i  $|a_-\rangle$  de la superposició  $|\phi\rangle = a_+|S_z, +\rangle + a_-|S_z, -\rangle$ . A més, girant l'aparell de Stern-Gerlach obtindriem dades de la fase relativa entre  $a_+$  i  $a_-$ . Hauríem reconstruït l'estat incident amb la precisió desitjada.

El procediment de reconstrucció descrit es basa en el fet de poder fer còpies idèntiques d'un estat que desconeixem. És això possible en MQ? Examinem-ho.

Entendrem com a clonació el procés de fer una copia d'un estat arbitrari  $|\psi\rangle$  fent us de l'evolució unitària de la MQ. Per això ens ajudarem d'un estat "ancilla" o auxiliar:

$$|\psi\rangle|C\rangle \rightarrow U(|\psi\rangle|C\rangle) = |\psi\rangle|\psi\rangle$$

per a qualsevol  $|\psi\rangle$ .

Per exemple, si considerem espais de Hilbert de dos dimensions, vol dir

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \gamma \\ \rho \end{pmatrix} \rightarrow U \left[ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \gamma \\ \rho \end{pmatrix} \right] = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

que en notació matricial podem expressar com

$$\begin{pmatrix} U_{11} & U_{12} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \alpha\gamma \\ \alpha\rho \\ \beta\gamma \\ \beta\rho \end{pmatrix} = \begin{pmatrix} \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{pmatrix}$$

i que ha de complir-se per a qualsevol  $\alpha$  i  $\beta$ . Si com a estat auxiliar utilitzem el  $(1, 0)$ , voldrà dir que estem buscant una matriu unitària tal que

$$\begin{pmatrix} U_{11} & U_{12} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \alpha \\ 0 \\ \beta \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha\alpha \\ \alpha\beta \\ \beta\alpha \\ \beta\beta \end{pmatrix}$$

i això no té solució.

Veiem-ho en un cas general (tots els estats els pensem normalitzats). L'estat sencer de partida és, doncs, el producte tensorial  $|\phi\rangle \otimes |C\rangle$ , i volem que el nostre aparell de clonació realitzi una evolució unitària que ens dugui a un resultat

$$|\phi\rangle \otimes |C\rangle \longrightarrow |\phi\rangle \otimes |\phi\rangle,$$

És clar que voldrem que l'aparell de clonació actuï d'aquesta manera amb un  $|\phi\rangle$  arbitrari, perquè hem de recordar que aquest estat  $|\phi\rangle$  ens és desconegut. Per tant, caldrà que

$$|\psi\rangle \otimes |C\rangle \longrightarrow |\psi\rangle \otimes |\psi\rangle,$$

per un altre estat  $|\psi\rangle$  qualsevol. Notem, però, que, com que l'evolució temporal és unitària, els productes escalarss s'han de conservar:

$$\langle\phi|\psi\rangle = \langle\phi|\psi\rangle\langle\phi|\psi\rangle \quad (1.164)$$

cosa que només és possible si

$$1) \quad \langle\phi|\psi\rangle = 0,$$

o bé

$$2) \quad |\langle\phi|\psi\rangle| = 1.$$

Que 1) garanteix (1.164) és evident. Vegem l'opcio 2. És fàcil demostrar que si el valor absolut del producte escalar de dos estats normalitzats és igual a 1, són el mateix estat excepte d'una possible fase<sup>1</sup>. Per tant, el nostre aparell no pot clonar qualsevol estat. Si és capaç de clonar-ne un, només en podrà clonar d'altres –en tot cas– que li siguin ortogonals –i que siguin ortogonals entre ells. Aquesta és una limitació inesquivable del nostre formalisme. Arribem a la conclusió que *la clonació d'un estat desconegut és impossible en MQ*.

Amb independència del resultat anterior, també podem veure que la linealitat de l'evolució temporal en MQ prohíbeix directament la clonació. Si el nostre aparell és capaç de clonar els estats  $|\phi\rangle$  i  $|\psi\rangle$ , la linealitat de l'evolució temporal implicarà que, per a l'estat de superposició  $|\chi\rangle \equiv \frac{1}{\sqrt{2}}(|\phi\rangle + |\psi\rangle)$ ,

$$|\chi\rangle \otimes |C\rangle \longrightarrow \frac{1}{\sqrt{2}}(|\phi\rangle \otimes |\phi\rangle + |\psi\rangle \otimes |\psi\rangle),$$

que és completament diferent d'allò que l'aparell de clonació se suposava que havia de fer,

$$|\chi\rangle \otimes |C\rangle \longrightarrow |\chi\rangle \otimes |\chi\rangle$$

### Copies aproximades d'estats. Fidelitat

Pel teorema de no-clonació sabem que no podem copiar un qbit en un estat arbitrari però si que podem fer una copia aproximada d'un estat quàntic en el que no obtindrem l'estat perfectament duplicat.

Per definir que vol dir "aproximació" haurem de definir una espècie de distància entre estats tan purs com barreja. Comencem pel cas d'estats purs.

Per comprovar si dos estats purs són el mateix, i per tant només poden diferir en una fase global, el que podem fer és minimitzar  $\min_\alpha |||\Psi_1\rangle - e^{i\alpha}|\Psi_2\rangle||^2$  en funció d'alpha i si dona zero és el mateix estat. Si ho desenvolupem tenim

$$\begin{aligned} \min_\alpha |||\Psi_1\rangle - e^{i\alpha}|\Psi_2\rangle||^2 &= \min_\alpha (\langle\Psi_1| - e^{-i\alpha}\langle\Psi_2|)(|\Psi_1\rangle - e^{i\alpha}|\Psi_2\rangle) \\ &= \min_\alpha [|||\Psi_1\rangle||^2 + |||\Psi_2\rangle||^2 + 2\operatorname{Re}(e^{i\alpha}\langle\Psi_1|\Psi_2\rangle)] \\ &= \min_\alpha (2 - 2\operatorname{Re}(e^{i(\alpha+\theta)}|\langle\Psi_1|\Psi_2\rangle|) = 2(1 - |\langle\Psi_1|\Psi_2\rangle|) \end{aligned} \quad (1.165)$$

on  $\exp(i\theta)$  és la fase de  $\langle\Psi_1|\Psi_2\rangle$  i per tant el mínim es troba quant  $\alpha = -\theta$ .

---

<sup>1</sup>Demostració: escrivim  $|\psi\rangle = \alpha|\phi\rangle + \beta|\xi\rangle$  amb  $|\xi\rangle$  normalitzat i ortogonal a  $|\phi\rangle$ . De fet, podem prendre  $|\xi\rangle$  com  $|\psi\rangle - \langle\phi|\psi\rangle|\phi\rangle$  dividit per la norma. Llavors,  $|\alpha|^2 + |\beta|^2 = 1$ , però com que  $|\langle\phi|\psi\rangle| = 1$ , resulta que també  $|\alpha| = 1$ , d'on inferim que  $\beta = 0$  i, per tant,  $|\psi\rangle = \alpha|\phi\rangle$ .

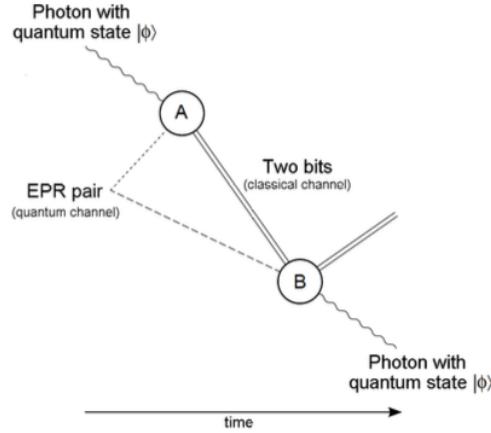


Figure 1.26: Teleportació quàntica de la informació (teleportació d'un estat quàntic).

Si definim la "fidelitat" entre dos estats com  $F \equiv |\langle \Psi_1 | \Psi_2 \rangle|^2$  tenim que el resultat anterior es pot escriure com  $2(1 - \sqrt{F})$ . Quant la fidelitat és 1 els estats són el mateix.

En el cas general definim la fidelitat  $F$  entre un sistema que es troba en l'estat  $\rho$  i un altre que es troba en l'estat  $\rho'$  com

$$F \equiv |Tr\sqrt{\rho^{\frac{1}{2}}\rho'\rho^{\frac{1}{2}}}|^2 \quad (1.166)$$

que en el cas d'estats purs queda reduït a  $F = |\langle \Psi_1 | \Psi_2 \rangle|^2$ . Veiem que la distingibilitat entre dos estats és inversament proporcional al solapament i per tant, inversament proporcional a la fidelitat.

Tot i saber que no podem copiar un qbit en un estat general  $\rho = (1 + \vec{b} \cdot \vec{\sigma})/2$ , , la màquina ideada per Buzek- Hillery és òptima en el sentit que maximitza la fidelitat entre el l'estat que obtenim de la pseudo-clonació i l'original. L'estat barreja duplicat de l'estat original és

$$\rho_{cl} = \frac{1}{4} \left( 1 \otimes 1 + \frac{2}{3} (\vec{b} \cdot \vec{\sigma} \otimes 1 + 1 \otimes \vec{b} \cdot \vec{\sigma}) + \frac{1}{3} \vec{\sigma} \vec{\sigma} \right) \quad (1.167)$$

### 1.7.10 Teleportació quàntica de la informació

Recordem que utilitzant transformacions unitàries no podem ni copiar un estat ni teleportar-lo en el sentit de convertir la informació quàntica que té en informació clàssica. Ara bé, utilitzant el fet de mesurar i una comunicació clàssica, podrem fer una mesura intel·ligent que evidentment destruirà l'estat, però del resultat de la mesura i d'un estat d'altres dues partícules que l'entrellacen amb ell, podem regenerar-lo.

Volem teleportar l'estat desconegut  $|\Psi\rangle_A = a|0\rangle + b|1\rangle$  d'A a B (veure figura 1.26) i per això ens ajudarem de l'estat entrellaçat entre A i B  $|\psi\rangle_{AB} = (1/\sqrt{2})(|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B)$ .

L'estat total de les 3 partícules és

$$\begin{aligned} |\Psi\rangle_A |\psi\rangle_{AB} &= (a|0\rangle_A + b|1\rangle_A) \frac{1}{\sqrt{2}} ((|0\rangle_A|1\rangle_B - |1\rangle_A|0\rangle_B) \\ &= \frac{1}{\sqrt{2}} (a|0\rangle_A|0\rangle_A|1\rangle_B + b|1\rangle_A|0\rangle_A|0\rangle_B - a|0\rangle_A|1\rangle_A|0\rangle_B a - b|1\rangle_A|1\rangle_A|0\rangle_B) \\ &= \frac{1}{\sqrt{2}} \frac{1}{2} \\ &[ (|0\rangle|1\rangle - |1\rangle|0\rangle)_A (-a|0\rangle - b|1\rangle)_B \end{aligned}$$

$$\begin{aligned}
& + (|0\rangle|1\rangle + |1\rangle|0\rangle)_A (-a|0\rangle + b|1\rangle)_B \\
& + (|0\rangle|0\rangle - |1\rangle|1\rangle)_A (b|0\rangle + a|1\rangle)_B \\
& + (|1\rangle|1\rangle + |0\rangle|0\rangle)_A (-b|0\rangle + a|1\rangle)_B ] \\
& = \frac{1}{2} \\
& [ |\Psi^-\rangle_A (-a|0\rangle - b|1\rangle)_B + |\Psi^+\rangle_A (-a|0\rangle + b|1\rangle)_B \\
& + |\varphi^-\rangle_A (b|0\rangle + a|1\rangle)_B + |\varphi^+\rangle_A (-b|0\rangle + a|1\rangle)_B ]
\end{aligned} \tag{1.168}$$

on ho hem expressat amb la base de Bell en A. Aleshores A amb un aparell anomenat Bell analyzer mesura i com a resultat obté un dels 4 estats possibles de la base de Bell. Un cop A ha fet la mesura, sap en quin estat ha quedat l'estat de B i només ha de comunicar-ho a B (com són 4 possible resultats, seran 2 bits clàssics que li haurà d'enviar) i aleshores B haurà de fer la corresponent transformació unitària per portar el seu estat a  $|\Psi\rangle_A = a|0\rangle + b|1\rangle$ . Si el resultat d'A és  $|\Psi^-\rangle$  o  $|\Psi^+\rangle$  o  $|\varphi^-\rangle$  o  $|\varphi^+\rangle$ , A envia a B 00, 01, 10 o 11 respectivament i B haurà de fer les transformacions  $I, -\sigma_x, \sigma_x, \sigma_y$  respectivament.

La forma més eficient d'un Bell Analyser és fer una transformació unitària (CNOT) de forma que

$$\begin{aligned}
CNOT|\Psi^\pm\rangle &= |\pm\rangle|0\rangle \\
CNOT|\varphi^\pm\rangle &= |\pm\rangle|1\rangle
\end{aligned} \tag{1.169}$$

on  $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ . Aleshores fent una mesura  $S_x$  sobre el primer estat i  $S_z$  sobre el segon, determinarà unívocament l'estat.

Ho tornarem a veure quant parlem de portes quàntiques on farem una implementació similar a l'anterior encara que no idèntica.

### 1.7.11 El teorema No-hiding

Clàssicament si es perd informació d'un subsistema, pot passar a un altre subsistema o bé pot ocultar-se en la correlació entre els dos subsistemes. Per exemple, si Alice envia un missatge M a Bob, però l'envia encriptat, la informació original del missatge com queda guardada en el missatge un cop encriptat? Es pot demostrar (Shannon, 1949) que el missatge encriptat no conté cap informació sobre el missatge original ni sobre la clau (si fos al contrari Eve podria deduir alguna cosa sobre el missatge o la clau) i tota la informació està en la correlació entre el missatge i la clau.

L'anàleg quàntic del cas anterior seria el poder codificar un estat quàntic en les correlacions entre dos subsistemes, tals que els subsistemes no tenen informació sobre l'estat. Per exemple, si suposem l'estat  $(\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha'|0\rangle + \beta'|1\rangle) = (\alpha\alpha', \alpha\beta', \beta\alpha', \beta\beta')$  (en la base natural), ens estem preguntant si mitjançant transformacions unitàries (matrius unitàries de  $4 \times 4$  sobre aquest vector) es podria eliminar completament la informació de  $\alpha$  i  $\beta$  en tots dos subsistemes (es a dir, les seves matrius densitat reduïdes no tinguin cap informació ni de  $\alpha$  ni de  $\beta$ ) i per tant, aquesta hagi quedat tota en la correlació quàntica (entrellaçament) entre els subsistemes que hi ha en l'estat global transformat. Doncs això és impossible i aquest teorema ens dirà que si s'elimina tota la informació de  $\alpha$  i  $\beta$  del primer subsistema, aquesta passa, en la seva totalitat, al segon subsistema.

La informació quàntica no es pot amagar en correlacions entre un parell de subsistemes. La mecànica quàntica només permet una manera d'ocultar completament un estat quàntic arbitrari d'un dels seus subsistemes: si es perd d'un subsistema, es trasllada a altres subsistemes. La demostració matemàtica (Braunstein i Pati, PRL, 2007) es reproduceix a continuació.

Agafem un estat arbitrari inicial  $\rho_I = |\psi\rangle_I \langle \psi|$  que evoluciona unitàriament juntament amb l'entorn a un estat+entorn de sortida

$$|\psi\rangle_I \otimes |E\rangle \rightarrow |\psi\rangle_{OE} \quad (1.170)$$

Aleshores, si l'estat de sortida es tal que

$$|\psi\rangle_I \rightarrow \sigma_O = Tr_E(|\psi\rangle_{OE} \langle \psi|) \quad (1.171)$$

és independent de l'estat d'entrada, estarem davant d'un procés d'ocultació (qualsevol informació de l'estat inicial  $|\psi\rangle$ , ha desaparegut durant la seva interacció amb l'entorn, arribant a un estat  $\sigma_O$  que ha de ser independent de l'estat inicial, doncs no conté cap informació sobre el mateix). Si això passa, es pot demostrar que tota la informació sobre l'estat inicial que ha "desaparegut" es codifica completament en l'entorn i no hi ha cap informació emmagatzemada en les correlacions entre els dos subsistemes.

Fem la descomposició de l'estat global final

$$|\psi\rangle_I \otimes |E\rangle \rightarrow |\psi\rangle_{OE} = \sum_{k=1}^K \sqrt{p_k} |k\rangle |A_k(\psi)\rangle_E \quad (1.172)$$

on recordem que  $\{|k\rangle\}$  són autovectors de  $\sigma_O$  (amb autovalors  $p_k$  diferents de 0) i  $\{|A_k\rangle\}$  és un conjunt de vectors ortonormals de l'entorn.

Degut a la linealitat de la MQ, s'ha de complir que

$$|A_k(\alpha|\psi\rangle + \beta|\psi_\perp\rangle)\rangle = \alpha|A_k(\psi)\rangle + \beta|A_k(\psi_\perp)\rangle \quad (1.173)$$

on  $|\psi_\perp\rangle$  és ortogonal a  $|\psi\rangle$  (val per qualsevol, però l'agafem ortogonal per conveniència). Com els  $\{|A_k\rangle\}$  són ortonormals tenim

$$\begin{aligned} 0 &= \langle A_l(\alpha|\psi\rangle + \beta|\psi_\perp\rangle) | A_k(\alpha|\psi\rangle + \beta|\psi_\perp\rangle) \rangle = (\alpha^* \langle A_l(\psi) | + \beta^* \langle A_l(\psi_\perp) |)(\alpha|A_k(\psi)\rangle + \beta|A_k(\psi_\perp)\rangle) \\ &= \alpha^* \beta \langle A_l(\psi) | A_k(\psi_\perp) \rangle + \beta^* \alpha \langle A_l(\psi_\perp) | A_k(\psi) \rangle \end{aligned} \quad (1.174)$$

i com s'ha de complir per a qualsevol valor de  $\alpha$  i  $\beta$ , TOTS els termes creuats han de ser zero.

Donada una base pels estats d'entrada  $\{|\psi_j\rangle, j = 1, \dots, d\}$ , el resultat anterior ens permet definir un conjunt d'estats ortonormal en l'entorn  $|A_{kj}\rangle \equiv |A_k(\psi_j)\rangle$  i que som lliures d'escriure com

$$|A_{kj}\rangle = |q_k\rangle \otimes |\psi_j\rangle$$

on  $\{|q_k\rangle\}$  és un conjunt de K estats.

Per linealitat tenim:

$$|A_k(\psi)\rangle = |A_k(\sum_j c_j |\psi_j\rangle)\rangle = \sum_j c_j |A_k(|\psi_j\rangle)\rangle = \sum_j c_j |q_k\rangle |\psi_j\rangle = |q_k\rangle |\psi\rangle \quad (1.175)$$

i per tant

$$|\psi\rangle_I \otimes |E\rangle \rightarrow |\psi\rangle_{OE} = \sum_{k=1}^K \sqrt{p_k} |k\rangle |A_k(\psi)\rangle_E = \sum_{k=1}^K \sqrt{p_k} |k\rangle (|q_k\rangle |\psi\rangle)_E \quad (1.176)$$

d'on veiem que  $|\psi\rangle$  ha quedat TOTALMENT codificat en els estats de l'entorn. La informació quàntica de l'estat inicial ha passat TOTA a l'entorn i no ha quedat res en l'entrellaçament amb l'entorn.

## 1.8 Comunicació quàntica: Codificació quàntica

L'anomenada compressió de Schumacher és l'equivalent al primer teorema de Shannon de compressió de missatges clàssics.

Recordem el primer teorema de Shannon. Suposem que Alice utilitza bits per transmetre un cert alfabet  $\chi$  on els símbols  $x \in \chi$  apareixen amb probabilitats  $p(x)$ , aleshores podrem trobar una codificació òptima tal que longitud promig (en bits) d'aquest alfabet sigui la entropia de Shannon

$$\bar{L} = S(p(x))$$

i que correspon al cas en que cada símbols  $x$  es codifica en  $-\log_2 p(x)$  bits.

Per exemple, imaginem que Alice té un alfabet amb  $2^n$  símbols i cada un d'ells surt amb la mateixa probabilitat  $p = 1/2^n$ . Com que l'entropia de Shannon en aquest cas és  $S = -\sum_i p_i \log_2 p_i = -2^n(1/2^n \log_2(1/2^n)) = n$ , el codi òptim és aquell que en promig codifica els símbols en  $n$  bits (de fet, cada símbol ha de codificar-se en  $-\log_2 2^{-n} = n$ ). Un plantejament similar al problema es considerar que Alice ja utilitza  $n$  bits per enviar missatges ( $2^n$  possibles) i si  $p(0) = p(1) = 1/2$  ja podem dir que la codificació és òptima.

L'equivalent quàntic és el teorema de Schumacher. Suposem que Alice utilitza qbits per transmetre informació, aleshores podrem trobar una codificació òptima tal que longitud promig (en qbits) sigui la entropia de Von Neumann

$$\bar{L} = S(\rho)$$

Per exemple, considerem que Alice utilitza  $n$  qbits per enviar informació. Si cada qbit és de màxima barreja ( $\rho = \frac{1}{2}I^{(2 \times 2)}$ ) ja podem dir que la codificació és òptima doncs

$$\bar{L} = S(\rho_T) = S(\rho_1 \otimes \rho_2 \dots, \rho_n) = S\left(\frac{1}{2^n}I^{(2^n \times 2^n)}\right) = n$$

## 1.9 Comunicació quàntica: Criptografia quàntica

La critografia quàntica es basa en el fet de que Alice i Bob poden detectar que algú està escoltant el seu canal. En el cas que observin que ningú està escoltant el seu canal, aleshores estan segurs que comparteixen una llista de bits que només ells coneixen. Distingirem dos protocols en funció si es basen o no en la utilització d'estats entrellaçats.

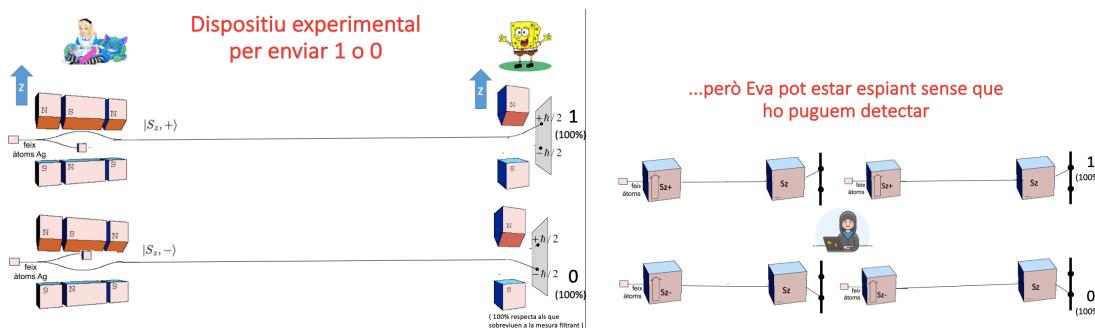


Figure 1.27: Alice fa mesures filtrants en la direcció  $z$  per enviar a Bob els bits desitjats

Figure 1.28: Eve pots espiar el canal sense poder ser detectada

Una possibilitat de comunicació és la indicada en la figura 1.27 on Alice fa mesures filtrants en la direcció z per enviar a Bob els bits desitjats. El problema d'aquest procediment es que Eve pot interceptar els missatges tal i com es mostra en la figura 1.28 (cubs amb fletxa són mesures filtrants en la direcció indicada pel text i els altres cubs són simplement mesures en la direcció indicada també en el text): Eve mesura en la mateixa direcció que utilitza Alice per generar-los i un cop obtingut el resultat fa una mesura filtrant que envia el bit mesurat cap a Bob. Alice i Bob poden comparar la llista de bits i no s'adonaran compte de que Eve els ha interceptat. Com fer possible el detectar si Eve esta escoltant?

### 1.9.1 Protocol BB84

Aquest protocol va ser el primer que es va idear el 1984. Es basa en que Alice codifica el bits en dos bases diferents, per exemple  $\{|0\rangle = |S_z, +\rangle, |1\rangle = |S_z, -\rangle\}$  o bé  $\{|0\rangle = |S_x, +\rangle, |1\rangle = |S_x, -\rangle\}$ . Aleshores tan Alice com Bob decideixen en cada moment de forma aleatòria en quina direcció, si z o x, posen els seus aparells (les 8 possibilitats es mostren en la figura 1.29). Un cop fetes les mesures tots dos fan públiques les seves direccions i on hagin coincidit tenen bits que només ells comparteixen. Alice només li ha de comunicar a Bob les posicions que ha de utilitzar per rebre el missatge.

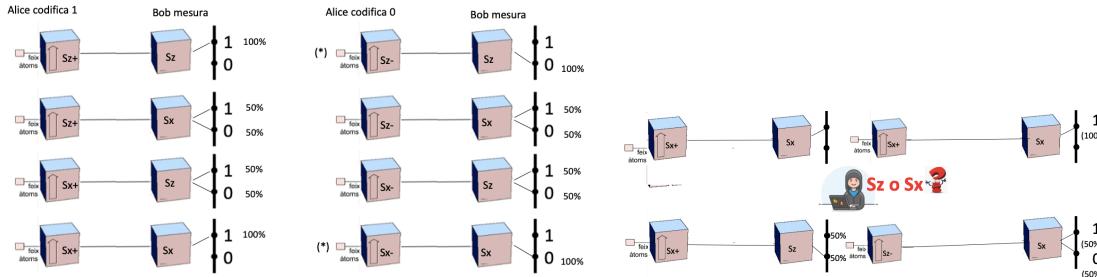


Figure 1.29: les 8 possibilitats de mesura entre Alice i Bob

Figure 1.30: Eve per espiar el canal ha de decidir en quina direcció mesura.

Ara Eve, per espiar el canal, ha de decidir en quina direcció mesura. Si Bob mesura en la mateixa direcció d'Alice, però Eve ha mesurat en l'altre direcció, aleshores Bob obtindrà el bit contrari a Alice amb una probabilitat del 50%.

Un exemple el tenim en la figura 1.31. En blau tenim la seqüència de bits enviats per Alice utilitzant el filtre en groc. Aleshores Eve intercepta l'estat i el mesura en la direcció donada en gris. Atenent el resultat, genera un nou estat que Bob mesura utilitzant un observable donat en verd obtenint un cert resultat donat en gris clar. Degut a la presència d'Eve el resultat de Bob no sempre concorda amb els resultats esperats per Alice quant ambdós mesuren en la mateixa direcció (resultats en vermell).

Si ara Bob publica la seqüència dels seus observables, Alice coneix exactament els bits que hauria d'haver mesurat Bob en les posicions 1,3,4,5,8,10,11,... Si ara Bob publica un subconjunt dels seus resultats, Alice pot comprovar si coincideixen amb els resultats esperats. Si no coincideixen (un 25% diferents) aleshores sabem que algú (EVE!!!) està interceptant els missatges.

### 1.9.2 Ekert 91

Aquest protocol utilitza estats entrellaçats tal i com es mostra en la figura 1.32. Requereix d'una entitat central, com per exemple un satèl·lit, que ha de ser un trusted node (no pot ser EVA!!!!), que enviarà parells de partícules entrellaçades en singlets: una per Alice i una per Bob. Si ambdós decideixen fer les mesures en direcció z, quant Alice observa 1 (0) sap, amb probabilitat 1, que Bob ha observat el contrari, es a dir 0 (1) i per tant Alice coneix exactament

1	0	Sx-	Sz	Sx	1 o 0	0
2	0	Sx-	Sz	Sz	1 o 0	0
3	0	Sx-	Sx	Sx	0	0
4	0	Sx-	Sx	Sx	1 o 0	0
5	0	Sz-	Sx	Sz	1 o 0	0
6	1	Sx+	Sx	Sz	1 o 0	0
7	0	Sx-	Sx	Sz	1 o 0	0
8	0	Sx-	Sz	Sx	1 o 0	0
9	0	Sx-	Sz	Sz	1 o 0	0
10	1	Sx+	Sx	Sx	1	1
11	1	Sz+	Sz	Sz	1	1
12	0	Sz-	Sz	Sx	1 o 0	0
13	0	Sx-	Sz	Sx	1 o 0	0
14	1	Sx+	Sz	Sx	1 o 0	0
15	1	Sx+	Sz	Sz	1 o 0	0
16	1	Sz+	Sz	Sx	1 o 0	0
17	1	Sz+	Sx	Sx	1 o 0	0
18	1	Sx+	Sz	Sx	1 o 0	1
19	1	Sz+	Sx	Sx	1 o 0	0
20	0	Sx-	Sx	Sz	1 o 0	0
21	1	Sx+	Sx	Sz	1 o 0	0
22	1	Sx+	Sz	Sx	1 o 0	1
23	1	Sx+	Sz	Sz	1 o 0	0
24	0	Sz-	Sx	Sz	1 o 0	0
25	1	Sx+	Sz	Sx	1 o 0	1
26	1	Sz+	Sz	Sx	1 o 0	0
27	1	Sx+	Sx	Sz	1 o 0	0
28	1	Sx+	Sx	Sz	1 o 0	0
29	0	Sx-	Sz	Sz	1 o 0	0
30	0	Sx-	Sx	Sz	1 o 0	0
31	1	Sx+	Sx	Sz	1 o 0	0
32	0	Sx-	Sz	Sx	1 o 0	0
33	0	Sz-	Sz	Sz	0	0
34	0	Sx-	Sx	Sz	1 o 0	0
35	1	Sx+	Sx	Sz	1 o 0	0
36	0	Sz-	Sz	Sz	0	0
37	0	Sz-	Sz	Sx	1 o 0	0
38	0	Sx-	Sx	Sz	1 o 0	0
39	1	Sz+	Sx	Sx	1 o 0	0
40	0	Sx-	Sx	Sz	1 o 0	0

Figure 1.31: Exemple de resultats del protocol BB84 (vermell sense Eve espiant, en gris clar amb Eve espiant)

la seqüència de bits que ha obtingut Bob. Per transmetre un missatge només haurà d'especificar les posicions dels bits en la seqüència que només ells dos comparteixen.

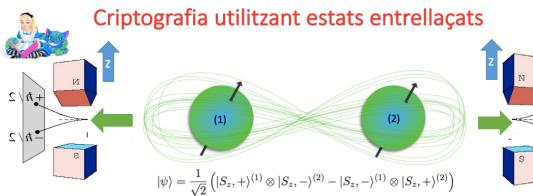


Figure 1.32: mesura d'un estat entrellaçat per Alice i Bob

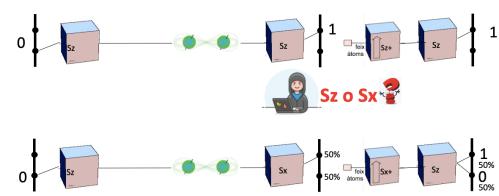


Figure 1.33: Eve per espiar el canal ha de decidir en quina direcció mesura.

Per esbrinar si Eve esta escoltant, farem com en el cas anterior: Alice i Bob decideixen en cada moment de forma aleatòria en quina direcció, si z o x, posen els seus aparells. Un cop fetes les mesures tots dos poden fer públiques les seves direccions i on hagin coincidit saben els bits (en aquest cas oposats) que només ells comparteixen, això si, si Eve no els ha interceptat. Si Eve intercepta el missatge haurà de decidir en quina direcció fa la mesura (figura 1.33). Si Alice i Bob han decidit fer la mesura en la mateixa direcció, però Eve l'ha fet en l'altra direcció, aleshores un 50% del cops el resultat entre Alice i Bob no serà el que ells esperarien.

Un exemple el tenim en la figura 1.31. En groc i verd tenim la seqüència de les direccions de mesura escollides per Alice i Bob. Només en el cas que coincideixin podem esperar resultats opositius com els donats en les columnes en vermell. Ara bé, Eve pot interceptar l'estat, decidint mesurar-lo en la direcció donada en gris i atenent el resultat, genera un nou estat cap a Bob. Degut a la presència d'Eve els resultats realment observats per Alice de Bob (columnes en gris) no sempre concordaran amb els resultats esperats per Alice i Bob quant ambdós mesuren en la mateixa direcció i no hi ha Eve pel mig (columnes en vermell).

Si ara Alice i Bob publiquen la seqüència dels seus observables, saben que en les posicions 2,5,7,8,11,... els seus bits haurien de ser complementaris. Si ara Alice i Bob publiquen un subconjunt dels seus resultats, poden comprovar si coincideixen amb els resultats esperats. Si no coincideixen (un 25% diferents) aleshores sabem que algú (EVE!!!) està interceptant els missatges.

1	Sz	Sz	Sx	1	0	1	0			
2	Sx	Sx	Sx	0	1	0	1			
3	Sz	Sz	Sx	1	0	0	1			
4	Sz	Sz	Sx	1	0	1	0			
5	Sz	Sz	Sz	1	0	1	0			
6	Sz	Sx	Sx	1	0	0	1			
7	Sx	Sz	Sx	1	0	0	1			
8	Sz	Sz	Sz	0	1	0	1			
9	Sz	Sz	Sz	1	0	0	1			
10	Sx	Sx	Sz	1	0	0	1			
11	Sz	Sz	Sz	1	0	1	0			
12	Sz	Sz	Sz	1	0	1	0			
13	Sz	Sz	Sx	1	0	0	1			
14	Sz	Sx	Sx	1	0	1	0			
15	Sx	Sx	Sx	1	0	1	0			
16	Sx	Sx	Sz	1	0	0	1			
17	Sz	Sx	Sz	1	0	1	0			
18	Sz	Sz	Sx	1	0	0	1			
19	Sz	Sx	Sx	1	0	0	1			
20	Sx	Sx	Sz	1	0	1	0			
21	Sx	Sz	Sx	1	0	0	1	0		
22	Sx	Sz	Sz	1	0	0	1			
23	Sz	Sz	Sz	1	0	1	0			
24	Sx	Sz	Sz	1	0	0	1			
25	Sx	Sx	Sx	1	0	1	0			
26	Sz	Sz	Sz	0	1	0	1			
27	Sz	Sz	Sz	1	0	1	0			
28	Sz	Sx	Sx	1	0	0	1			
29	Sz	Sx	Sx	1	0	0	1			
30	Sz	Sx	Sz	1	0	0	1	0		
31	Sz	Sz	Sx	1	0	0	1			
32	Sx	Sz	Sx	1	0	0	1			
33	Sx	Sx	Sx	1	0	1	0			
34	Sx	Sx	Sx	0	1	0	1			
35	Sx	Sz	Sx	1	0	0	1			
36	Sz	Sx	Sz	1	0	0	1			
37	Sx	Sx	Sz	1	0	0	1			
38	Sx	Sz	Sz	1	0	0	1			
39	Sz	Sx	Sz	1	0	0	1	0		
40	Sz	Sx	Sz	1	0	0	1	0		

Figure 1.34: Exemple de resultats del protocol Ekert 91 (vermell sense Eve espiant, en gris clar amb Eve espiant)

## 1.10 Comunicació quàntica: Canals quàntics

A la teoria de la informació quàntica, un canal quàntic és un canal de comunicació que pot transmetre informació quàntica, així com informació clàssica. Un exemple d'informació quàntica és la transmissió de l'estat d'un qbit. Un exemple d'informació clàssica és un document de text transmès per Internet (bits).

Quant parlem de canals hem de parlar de soroll. Ja varem veure que en canals clàssics el soroll limitava la capacitat del canal i que per apropar-nos al límit de la seva capacitat podríem utilitzar mètodes per corregir errors. El problema de corregir errors implica que haurem de mesurar i això sabem quines conseqüències té en el mon quàntic. Per aquest motiu es pensava que no es podria fer cap tipus de correcció en canals quàntics però hi ha certs mètodes que utilitzen el fet de mesurar per disminuir la probabilitat d'error de la transmissió. Discutirem qualitativament el principi d'aquest mètodes i donarem un exemple quan parlem de circuits quàntics. Abans però, parlarem de com podem caracteritzar el soroll que ens podem trobar al enviar un qbit, indicant el seu efecte en l'esfera de Bloch.

### 1.10.1 Modelització del soroll d'un canal

Formalment, els canals quàntics són mapes de conservació de traces completament positives (CP) entre espais dels operadors.

$$\rho \rightarrow \Phi(\rho) \quad (1.177)$$

on  $Tr(\Phi(\rho)) = Tr(\rho)$ . La conservació de la traça implica que es garanteix que l'estat segueix sent quàntic, no s'ha produït un col·lapse.

Exemples de canals sobre un qbit:

- un canal despolaritzador és un canal on el qbit té una probabilitat  $p$  de deixar-lo tal i com està i una probabilitat  $1 - p$  de convertir-se amb un estat de màxima barreja.

$$\Phi(\rho) = p\rho + (1 - p)\frac{1}{2} \quad (1.178)$$

veiem que en aquest cas  $Tr(\Phi(\rho)) = p + (1 - p) = 1$  la traça es manté com demanàvem. En la figura 1.35 podem veure els efectes d'aquest canal en l'esfera de Bloch

- un canal d'amortiment és un canal on el qbit té una probabilitat  $p$  de deixar-lo tal i com està i una probabilitat  $1 - p$  de convertir-se en l'estat fonamentat  $|0\rangle$  (per exemple per emissió espontània d'un fotó en un àtom de dos nivells).

$$\Phi(\rho) = p\rho + (1 - p)|0\rangle\langle 0| \quad (1.179)$$

veiem que en aquest cas també  $\text{Tr}(\Phi(\rho)) = p + (1 - p) = 1$  es manté la traça. En la figura 1.36 podem veure els efectes d'aquest canal en l'esfera de Bloch. Els punt vermells connectats volen indicar l'efecte del soroll sobre un estat inicialment pur de l'esfera de Bloch a l'estat barreja corresponent després de passar pel canal.

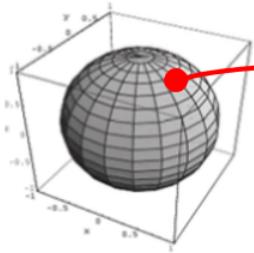


Figure 1.35: canal despolaritzador

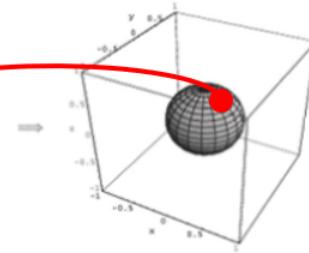


Figure 1.36: canal d'amortiment

### 1.10.2 Correcció d'errors

Un possible mètode de correcció és entrelaçar el qbit, que porta la informació quàntica que volem preservar, amb altres qbites auxiliars dels quals coneixem el seu estat (generalment  $|0\rangle$ ) i fer-ho abans d'enviar-los tots pel canal. Aleshores, a la sortida del canal, podem fer mesures sobre els qbites auxiliars que ens donaran informació sobre el grau de soroll que s'ha produït durant la transmissió i així poder actuar en conseqüència. Actuant d'aquesta forma veurem que podrem reduir la probabilitat de canvi del qbit que porta la informació. En la secció de circuits quàntics donarem un exemple aclaridor, on veurem que serveix per reduir aquesta probabilitat, de  $p$  que dona el canal a  $o(p^2)$ .

## 1.11 Computació quàntica

Un ordinador quàntic agafarà informació codificada en un estat quàntic i després realitzaria "operacions de porta" predeterminades segons les lleis de la mecànica quàntica i produiria un nou estat quàntic, que es pot mesurar per determinar el resultat del càcul.

La diferència clau amb l'ordinador clàssic rau en la capacitat d'un estat quàntic de representar molts estats possibles "clàssics" alhora. Mentre que un bit clàssic pot ser un "0" o un "1", un bit quàntic, o "qbit", és qualsevol combinació o "superposició" possible de "0" i "1", amb nombres complexos com a coeficients de la superposició. En la figura 1.3 podem veure que un "bit" d'informació en un ordinador clàssic es pot representar per dos punts ("0" o "1") mentre que un "qbit" es representa en un ordinador quàntic com qualsevol punt de la superfície d'un 3D esfera: l' "Esfera de Bloch".

La gran potència d'un ordinador quàntic rau en que pot processar tots els estats possibles en paral·lel. No només cada qbit pot estar en un estat de superposició, sinó que el sistema en general pot estar en una superposició de totes les combinacions d'estats diferents de tots els qbites. Aleshores el nombre d'estats possibles que poden estar presents a la superposició és enorme: si tenim  $N$  qbites, hi ha  $2^N$  estats possibles a la superposició mentre que un ordinador

clàssic de N bits, només pot estar en una sola configuració. Un ordinador quàntic amb només 30 qubits tindria 1.073.741.824 estats possibles simultàniament i un ordinador quàntic amb 300 qubits tindria aproximadament el mateix nombre d'estats possibles que el nombre total d'àtoms de l'univers coneugut.

Haurem de tenir en compte tres fets fonamentals a l'hora de pensar en ordinador quàntics :

- L'evolució sempre unitària dels estats sota l'acció de qualsevol Hamiltonià (exceptuant si decidim fer una mesura, evidentment). Això implica dues restriccions molt importants.

- Primer, totes les "gates" quàntiques són forçosament reversibles. La unitarietat de la nostra dinàmica implica reversibilitat doncs totes les transformacions unitàries tenen una inversa ( recordem que el contrari no és cert doncs tenim matrius invertibles que no són unitàries com per exemple  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  ).

NOTA:

Encara que la unitarietat de la nostra dinàmica impliqui reversibilitat (podem passar la pel·lícula dels fets sense ambigüïtat), com és el cas de la MQ, això no implica la NO violació del operador d'inversió temporal T doncs les nostres lleis poden ser invertibles però no ser invariants sota T (la pel·lícula cap endarrere no té per que ser una solució possible de la nostra dinàmica). Per exemple,  $1 = dx/dt$ , que té com a solució  $x(t) = t + C$ , és invertible, es a dir, podem passar la pel·lícula cap endarrere sense ambigüïtat ( $x^T(t) = -t + C'$ ), però no és invariant sota T doncs aquesta no és solució de la nostra dinàmica  $dx^T/dt = -1 \neq 1$ . Una forma alternativa de veure-ho és que la llei de la dinàmica sota T canvia de signe la part dreta i per tant no és invariant.

En MQ, T ve implementat per un operador antiunitari amb  $T^2 = \pm 1$  i  $T^\dagger T = 1$ .

Designem l'estat transformat sota T com  $|\phi^T\rangle \equiv T|\phi\rangle$ , per exemple  $|\vec{x}^T\rangle = T|\vec{x}\rangle = |\vec{x}\rangle$ ,  $T|\vec{p}\rangle = |-\vec{p}\rangle$ ,  $T|j, m\rangle = (-1)^{j-m}|j, -m\rangle$  (canviar la direcció del moment i spin de forma conseqüent).

Direm que el nostre problema presenta invariància sota inversió temporal si  $\langle f|U|i\rangle = \langle i^T|U|f^T\rangle$ , és a dir, la dinàmica donada per U que és la que connecta  $|i\rangle$  anant a  $|f\rangle$ , ha de ser la mateixa que connecti la pel·lícula cap endarrere, és a dir, de  $|f^T\rangle$  anant a  $|i^T\rangle$ .

Com  $\langle i^T|U|f^T\rangle = \langle Ti|U|Tf\rangle = \langle i|T^\dagger UT|f\rangle^* = \langle f|T^\dagger UT|i\rangle$  deduïm que el nostre problema presenta invariància sota inversió temporal només si  $U = T^\dagger U^\dagger T$ . (en el cas de que H no depengui del temps i  $[H, T] = 0$  la relació queda satisfeta doncs  $T^\dagger U^\dagger T = T^\dagger e^{+iHt/\hbar} T = T^\dagger T e^{-iHt/\hbar} = e^{-iHt/\hbar} = U$ ).

Veiem doncs que, encara que l'operador d'evolució temporal sempre és unitari i per tant invertible, només en el cas en que  $U = T^\dagger U^\dagger T$  estem davant d'un problema que és invariant sota inversió temporal T. Només en aquests casos podem assegurar que qualsevol solució de la nostra dinàmica "visualitzada" cap endarrere en el temps, també és solució de la nostra dinàmica. Si H depèn explícitament del temps, i per tant  $U(H(t))$  també, hem de tenir en compte la subtilesa de que mentre en la solució cap endavant aquest temps explícit apareix com  $H(t)$ , en la cap endarrere haurem de posar  $H(-t)$ .

- Segon, no podem fer copies d'un estat quàntic segons el teorema de la no clonació.

- podemaprofitar l'existència d'estats entrellaçats per processar la informació. La mesura d'una part col·lapse instantàniament l'estat de l'altra part, fet que podrem utilitzar per processar informació (però que recordem no permet enviar informació instantàniament).

- la mesura trenca l'evolució unitària amb un resultat probabilístic. Sempre haurem de fer una mesura per obtenir/observar el resultat. Haurem de conviure amb aquesta aleatorietat i, si pot ser, aprofitar-la.

Aquests 3 fets anteriors produeix que hi hagin dues grans dificultats amb els ordinadors quàntics: construir-ne un i determinar com programar aquest sistema .

Construir un ordinador quàntic és un repte, ja que implica manipular un gran nombre d'objectes microscòpics segons les lleis de la mecànica quàntica. La principal dificultat és que l'objecte microscòpic sobre el qual emmagatzemem cada qbit ha d'estar aïllat del seu entorn tant com sigui possible per evitar la descoherència i que la informació passi a l'entorn (teorema de no-hiding). Aquest transvasament d'informació quàntica entre l'ordinador quàntic i el seu entorn, també es pot veure com el soroll que introduceix l'entorn i aleatoritza l'estat del qbit. L'exemple d'ordinador quàntic que tractarem amb més detall serà el dels Ions atrapats: en aquests sistemes, les cadenes d'ions s'emmagatzemen en trampes electromagnètiques i la informació es codifica en estats electrònics de llarga vida dels àtoms

La programació d'un ordinador quàntic es fa més difícil per les lleis de la mesura en mecànica quàntica: quan mesurem el sistema no obtenim tots els resultats possibles, sinó que la mesura col·lapsa l'estat en un resultat dels possibles. Per això, no és fàcil dissenyar algoritmes que facin ús de la potència intrínseca d'un ordinador quàntic. També són difícils d'escriure, perquè són matemàticament més complexes i molt menys intuïtius que els algorismes per a un ordinador clàssic.

En comparació amb la programació d'ordinadors convencionals, la situació actual en la programació quàntica està en la fase de com utilitzar i combinar portes lògiques quàntiques (veurem quines propietats han de tenir i quines són) per solucionar algun problema plantejat. Els passos són bàsicament:

1. Inicialització: Crear un o més qbites i registres clàssics per mesurar-los.
2. Disseny circuit: Crear un circuit que agrupi els qbites en una unitat d'execució lògica.
3. Running: Aplicar portes quàntiques als qbites per aconseguir el resultat desitjat.
4. Resultat: Mesurar els qbites al registre clàssic per obtenir un resultat final.

Un exemple de circuit el tenim en la Figura 1.37 on , es preparen 2 qbites en estat fonamental  $|0\rangle$ . La porta H porta el primer qbit a l'estat  $(|0\rangle + |1\rangle)/\sqrt{2}$  deixant l'estat dels dos qbites a  $(|00\rangle + |10\rangle)/\sqrt{2}$ . A continuació, la porta CNOT gira el segon qbit si el primer està excitat, fent que l'estat dels dos qbites passi a ser  $(|00\rangle + |11\rangle)/\sqrt{2}$ . Finalment fem les mesures sobre els dos qbites i haurem d'obtenir "00" un 50%, "11" un 50% i en cap cas "01" ni "10". Aquest "circuit/programa" el podem simular evidentment fent càlculs en paper (utilitzant matrius unitàries i daus per generar números aleatòria i simular les mesures), en QISKit ( Python) o bé executar-lo en un ordinador quàntic.

Com exemple de simulació en QISKit, a continuació podem trobar el codi en QISKit corresponent a la porta Pauli X (bit flip de la figura 1.38). Primer creem un únic qbit, un registre clàssic per mesurar el qbit i, a continuació, apliquem la porta Pauli X al qbit i finalment, mesurem el seu valor.

```
# ----- Program Circuit X -----
import sys
import qiskit
import logging
from qiskit import QuantumProgram
# Main sub
```

```

def main():
    # create a program qp = QuantumProgram()
    # create 1 qbit
    quantum_r = qp.create_quantum_register("qr", 1)
    # create 1 classical register
    classical_r = qp.create_classical_register("cr", 1)
    # create a circuit
    qp.create_circuit("Circuit", [quantum_r], [classical_r])
    # get the circuit by name
    circuit = qp.get_circuit('Circuit')
    # Pauli X gate to qbit 1 in the Quantum Register "qr"
    circuit.x(quantum_r[0])
    # measure gate from qbit 0 to classical bit 0
    circuit.measure(quantum_r[0], classical_r[0])
    -----
    # backend simulator
    backend =' local_qasm_simulator'
    # Group of circuits to execute
    circuits = ['Circuit']
    # Compile your program
    qobj = qp.compile(circuits,backend)
    # run in simulator
    result = qp.run(qobj,timeout = 240)
    # Show result counts
    print(str(result.get_counts('Circuit')))
```

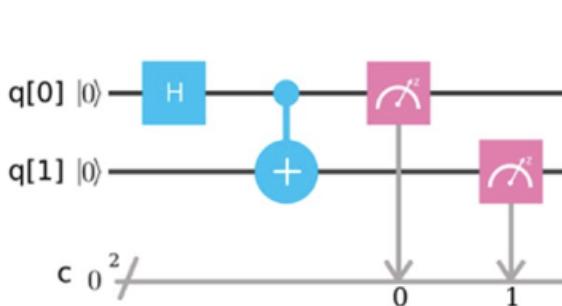


Figure 1.37: Exemple de disseny de circuit quàntic. Es pot simular en python o executar en un ordinador quàntic

La pregunta fonamental que ens hem de fer és: quines són les "portes" que poden existir a nivell quàntic?.

### 1.11.1 Portes quàntiques

Les dues regles fonamentals que tenim i que limiten les possibles portes quàntiques que podem pensar es poden implementar experimentalment són:

1. Evolució unitària i lliniaritat. Per tant vindran representades per matrius unitàries
2. El teorema de no-clonació: no podem fer copies d'estats quàntics arbitraris

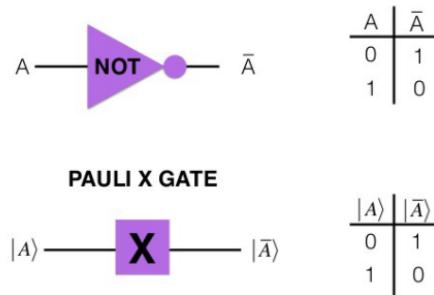


Figure 1.38: Algunes portes quàntiques sobre un qbit

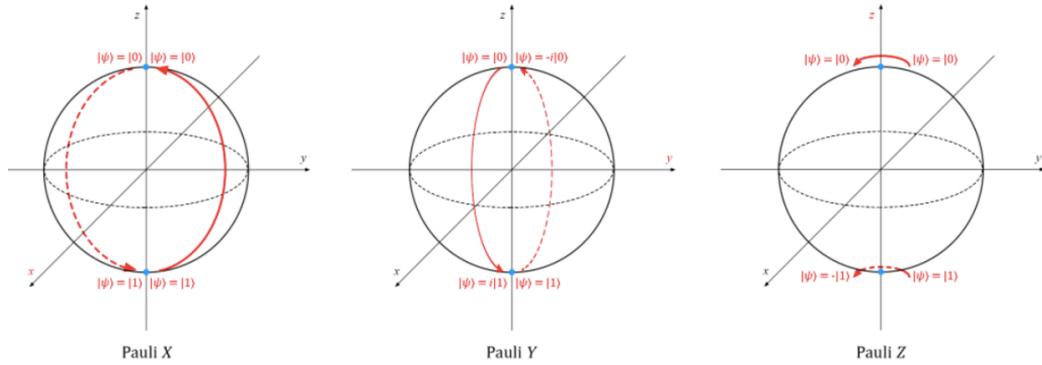


Figure 1.39: Portes Pauli

Primer parlarem de les portes quàntiques sobre un qbit i després sobre dos qbits.

### Portes quàntiques sobre un sol qbit

L'única porta lògica clàssica NO trivial sobre un bit és en la que s'intercanvien els estats  $0 \leftrightarrow 1$  (porta *NOT* de la figura 1.19). Es pot definir una porta quàntica anàloga per a qbits? Imagineu que teníem algun procés que portés l'estat  $|0\rangle$  a l'estat  $|1\rangle$  i viceversa. Com que la porta quàntica actua linealment (la MQ és linial) tindrem que

$$\alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|1\rangle + \beta|0\rangle \quad (1.180)$$

i per tant en la base  $\{|0\rangle, |1\rangle\}$  l'actuació d'aquesta porta ve representada per la matriu unitària:

$$X \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (1.181)$$

anomenada generalment com a porta Pauli-X i que és una rotació en l'esfera de Bloch i per tant totalment realitzable experimentalment (per spin  $1/2$  ho podríem fer amb un camp magnètic en la direcció adequada durant el temps pertinent).

En la figura 1.39 podem veure com actua la porta Pauli-X sobre qualsevol estat: és una rotació d'angle  $\pi$  al voltant de l'eix  $x$ . Notem que aquesta rotació no porta de l'estat  $|S_{\hat{n}}, +\rangle = \cos \frac{\theta}{2}|0\rangle + e^{i\varphi} \sin \frac{\theta}{2}|1\rangle$  a l'estat  $|S_{\hat{n}}, -\rangle = -e^{-i\varphi} \sin \frac{\theta}{2}|0\rangle + \cos \frac{\theta}{2}|1\rangle$  doncs

$$X|S_{\hat{n}}, +\rangle = X \left( \cos \frac{\theta}{2}|0\rangle + e^{i\varphi} \sin \frac{\theta}{2}|1\rangle \right) = e^{i\varphi} \sin \frac{\theta}{2}|0\rangle + \cos \frac{\theta}{2}|1\rangle$$

al ser  $X$  una rotació de  $\pi$  al voltant de  $x$  ens portarà de  $\varphi$  a  $-\varphi$  i no de  $\varphi$  a  $\varphi + \pi$ .

Evidentment tenim infinites portes quàntiques que podem actuar sobre un qbit (qualsevol rotació o combinació), però les més utilitzades són les rotacions donades en les figures 1.39 i 1.40. En la figura 1.41 es donen les seves representacions matricials en la base  $\{|0\rangle, |1\rangle\}$  i també el conveni de símbols que es utilitzen per designar aquestes portes quan formin part de circuits quàntics.

La porta Hadamard (H) és una de les portes quàntiques més útils. Porta l'estat  $|0\rangle$  a l'estat  $(|0\rangle + |1\rangle)/\sqrt{2}$  ortonormal tant a  $|0\rangle$  com a  $|1\rangle$ . En la figura 1.40 veiem que queda implementada com una rotació d'angle  $\pi$  al voltant de l'eix  $x$  seguida d'una rotació d'angle  $-\pi/2$  al voltant de l'eix  $y$ .

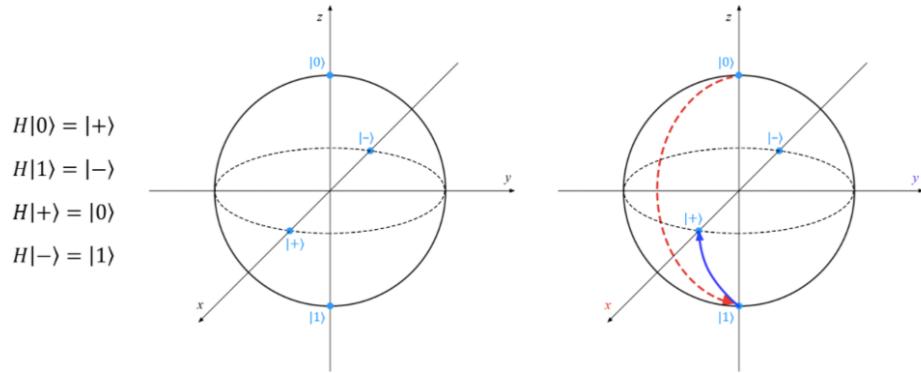


Figure 1.40: Porta Hadamard (H)

Hadamard	$\boxed{H}$	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
Pauli- $X$	$\boxed{X}$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli- $Y$	$\boxed{Y}$	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$
Pauli- $Z$	$\boxed{Z}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Phase	$\boxed{S}$	$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$
$\pi/8$	$\boxed{T}$	$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$

Figure 1.41: Portes quàntiques sobre un qbit

### Portes quàntiques sobre dos qbits

Pregunta: és possible implementar portes clàssiques de dos bits, com la porta NAND (universal) o la porta XOR (veure figura 1.19) en portes quàntiques?

Recordem que les portes quàntiques han de ser unitàries i per tant, reversibles. Ni la NAND ni la XOR són reversibles: donada la sortida X no és possible determinar quines eren les entrades A i B; hi ha una pèrdua irrecuperable d'informació associada a l'acció irreversible de la porta. Per tant, no podrem fer una implementació directament. Comprendre com fer la lògica clàssica en sentit reversible és un pas crucial per entendre com aprofitar el potencial de mecànica quàntica per a càcul.

Hem dit que per raons de irreversibilitat no podem implementar la porta XOR, però si podrem implementar el que anomenarem XOR-generalitzada o C-NOT. Fixem-nos que la porta XOR de la figura 1.19 actua com NOT del bit B sempre i quant el bit A estigui a 1. Per tant, podem considerar la següent taula de transformació de dos bits a dos bits:

$$A = 0, B = 0 \rightarrow A = 0, X = 0, 01 \rightarrow 01, 10 \rightarrow 11, 11 \rightarrow 10$$

i que és totalment reversible (donada la sortida sabem quina ha estat l'entrada). Veiem ara el motiu pel qual que es diu XOR-generalitzada al ser la porta clàssica XOR però conservant el bit d'entrada A, o també controlada-NOT per les raons donades.

Per aquest motiu, la porta lògica quàntica prototípica és la porta NOT-controlada o C-NOT. Aquesta porta té dos qbits d'entrada, conegeuts com a qbit de control i qbit objectiu, respectivament. El qbit objectiu només canvia si el qbit de control val  $|1\rangle$ :

$$|00\rangle \rightarrow |00\rangle, |01\rangle \rightarrow |01\rangle, |10\rangle \rightarrow |11\rangle, |11\rangle \rightarrow |10\rangle \quad (1.182)$$

que es pot resumir en  $|A, B\rangle \rightarrow |A, B \oplus A\rangle$ , on  $\oplus$  és l'addició mòdul dos. La representació com a circuit de la porta C-NOT es mostra en la part esquerra de la figura 1.42 ; la línia superior representa el qbit de control, mentre que la línia inferior representa el qbit objectiu.

L'operador associat a la porta CNOT en la base  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  ve donat per la matriu:

$$U_{CN} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (1.183)$$

que al ser una matriu unitària la podrem implementar experimentalment. Veiem que la seva implementació no es podrà fer utilitzant simples rotacions sobre qbits individuals doncs introduceix entrellaçament entre ells (veurem exemples de com introduir entrellaçament quant parlem de trampes de ions). Sobre qualsevol estat  $|\Psi\rangle = a_1|00\rangle + a_2|01\rangle + a_3|10\rangle + a_4|11\rangle = \sum_i a_i|i\rangle$  el resultat serà  $U_{CN}|\Psi\rangle = U_{ij} a_j|i\rangle$ . Evidentment, si l'estat d'entrada ja està entrellaçat, l'actuació serà

$$\sum_i \alpha_i |A_i, B_i\rangle \rightarrow \sum_i \alpha_i |A_i, B_i \oplus A_i\rangle \quad (1.184)$$

amb  $A_i, B_i$  0 o 1.

La porta C-NOT (NOT-controlada) és una de les infinites portes controlades que podem definir sobre dos qbits. Queda clar que podem definir la porta U-controlada, on U és un operador sobre un qbit, com aquella porta que opera sobre dos qbits, on el primer qbit controla l'actuació de la porta U sobre el segon qbit: si el primer qbit està a  $|1\rangle$  actua U sobre el segon, en cas contrari no es fa res. La representació com a circuit de la porta U-controlada o C-U es mostra en la part dreta de la figura 1.42 ; la línia superior representa el qbit de control, mentre que la línia inferior representa el qbit objectiu. En aquest cas general, l'operador sobre els dos qbits és

$$H_{AB}(U) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes H_U \quad (1.185)$$

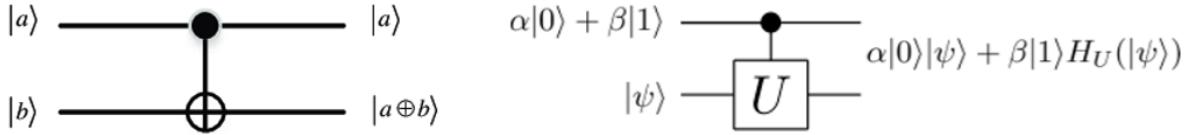


Figure 1.42: porta C-NOT i porta en general controlada C-U. Pel cas C-U, es dona un estat general no entrellaçat d'entrada i la seva sortida, pel cas C-NOT es dona el resultat pel cas particular de la base computacional (on  $|a\rangle$  i  $|b\rangle$  són  $|0\rangle$  o  $|1\rangle$ ), per obtenir el resultat en general només haurem aplicar linealitat.

Si en la base  $\{|0\rangle, |1\rangle\}$  l'operador  $H_U = \begin{pmatrix} x_{00} & x_{01} \\ x_{10} & x_{11} \end{pmatrix}$  aleshores l'operador  $H_{AB}(U)$  en la base  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  ve donar per

$$H_{AB}(U) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x_{00} & x_{01} \\ 0 & 0 & x_{10} & x_{11} \end{pmatrix} \quad (1.186)$$

Notem que encara que l'estat inicial dels 2 qbits no estigui entrellaçat,  $|\varphi\rangle|\psi\rangle$ , l'estat final si que ho pot estar

$$\begin{aligned} H_{AB}(|0\rangle|\psi\rangle) &= |0\rangle|\psi\rangle \\ H_{AB}(|1\rangle|\psi\rangle) &= |1\rangle H_U(|\psi\rangle) \\ H_{AB}[(\alpha|0\rangle + \beta|1\rangle)|\psi\rangle] &= \alpha|0\rangle|\psi\rangle + \beta|1\rangle H_U(|\psi\rangle) \end{aligned} \quad (1.187)$$

i en el cas més general, on l'estat inicial ja estigui entrellaçat, tenim

$$H_{AB}[(\alpha|0\rangle|\psi_0\rangle + \beta|1\rangle|\psi_1\rangle)] = \alpha|0\rangle|\psi_0\rangle + \beta|1\rangle H_U(|\psi_1\rangle) \quad (1.188)$$

### Portes quàntiques sobre n qbits

Portes quàntiques que actuen sobre n qbits són generalment portes controlades: tenim  $(n - 1)$  qbits que només en cas on TOTS estiguin en l'estat  $|1\rangle$ , s'aplicarà la transformació U al l'enèssim qbit. Aleshores l'operador associat serà

$$H_n(U) = \left( I^{\otimes(n-1)} - (|1\rangle\langle 1|)^{\otimes(n-1)} \right) \otimes I + (|1\rangle\langle 1|)^{\otimes(n-1)} \otimes H_U \quad (1.189)$$

Un exemple és la porta Toffoli que és com la porta C-NOT però controlada per dos qbits. La representació com a circuit d'aquesta es mostra en la figura 1.43 on les dues línies superiors representen els qbits de control, mentre que la línia inferior representa el qbit objectiu. La seva representació matricial en la base habitual serà

$$H_{Toffoli} = (I \otimes I - |1\rangle\langle 1| \otimes |1\rangle\langle 1|) \otimes I + (|1\rangle\langle 1| \otimes |1\rangle\langle 1|) \otimes X = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (1.190)$$

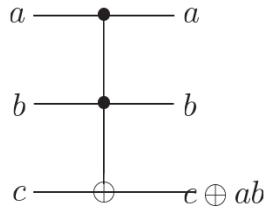


Figure 1.43: Porta Toffoli sobre tres qbit (NOT-controlada). Es dona el resultat pel cas particular de la base computacional (on  $|a\rangle$ ,  $|b\rangle$  i  $|c\rangle$  són  $|0\rangle$  o  $|1\rangle$ ), per obtenir el resultat en general només haurem aplicar linealitat.

on hem utilitzat el fet que el complementari del projector  $|1\rangle\langle 1| \otimes |1\rangle\langle 1|$  és  $I \otimes I - |1\rangle\langle 1| \otimes |1\rangle\langle 1|$ . Com  $I = |0\rangle\langle 0| + |1\rangle\langle 1|$ , també el podem escriure com

$$\begin{aligned} I \otimes I - |1\rangle\langle 1| \otimes |1\rangle\langle 1| &= (|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) - |1\rangle\langle 1| |1\rangle\langle 1| \\ &= |0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes |1\rangle\langle 1| + |1\rangle\langle 1| \otimes |0\rangle\langle 0| \\ &= |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10| \end{aligned} \quad (1.191)$$

i en cap cas és igual a  $|00\rangle\langle 00|$ , error bastant habitual.

### 1.11.2 Circuits quàntics

Acabem de veure algunes de les portes habituals que poden ser utilitzades en circuits quàntics. En general, un circuit quàntic estarà format per un conjunt de portes quàntiques que actuen seqüencialment sobre varis qbiuts. Per conveni, qualsevol circuit s'ha de llegir d'esquerra a dreta i cada línia representa un cable quàntic (aquest fil no correspon necessàriament a un fil físic; en canvi, pot correspondre al pas del temps, o potser a una partícula física com un fotó que es mou d'una ubicació a una altra a través de l'espai). És convencional suposar que l'estat d'entrada al circuit és un estat de base computacional, generalment l'estat que consta de tots  $|0\rangle$ s.

Hi ha algunes característiques permeses en els circuits clàssics que no soLEN estan presents en els circuits quàntics. En primer lloc, no permetem "buckles", és a dir, retroalimentació d'una part del circuit quàntic a una altra; diem que el circuit és acíclic.

En segon lloc, els circuits clàssics permeten que els cables s'uneixin, una operació coneguda com a CROSSOVER en un cable únic que conté el bit de les entrades. Òbviament, aquesta operació no és reversible i, per tant, no és unitària, de manera que no permetem en els nostres circuits quàntics. En tercer lloc, l'operació inversa, mitjançant la qual es produeixen diverses còpies d'un bit, tampoc no es permet als circuits quàntics (teorema de no-clonació).

En la figura 1.44 es mostren alguns dels símbols addicionals als ja vistos, que s'utilitzen habitualment en el disseny de circuits, com podem ser els associats a les mesures que en algun moment hem de fer sobre els qbits o com indicar l'utilització de bits (línia doble).

#### Circuit swap

Com a primer exemple de circuit quàntic començarem per un de molt senzil, l'anomenat circuit swap, combinació de tres portes C-NOT (figura 1.45).

El circuit de la figura 1.45 realitza una tasca senzilla però útil: intercanvia els estats dels dos qbits. Per veure que aquest circuit realitza l'operació d'intercanvi, tingueu en compte que la seqüència de 3 portes CNOT té la següent seqüència d'efectes sobre un estat computacional

measurement		Projection onto $ 0\rangle$ and $ 1\rangle$
qubit		wire carrying a single qubit (time goes left to right)
classical bit		wire carrying a single classical bit
$n$ qubits	$/^n$	wire carrying $n$ qubits

Figure 1.44: Altres elements dels circuits quàntics



Figure 1.45: circuit swap format per tres portes CNOT i el símbol utilitzat com equivalència

$|a, b\rangle$  ( a o b són 0 o 1)

$$\begin{aligned}
 |a, b\rangle &\rightarrow |a, a \oplus b\rangle \\
 &\rightarrow |a \oplus (a \oplus b), a \oplus b\rangle = |b, a \oplus b\rangle \\
 &\rightarrow |b, b \oplus (a \oplus b)\rangle = |b, a\rangle
 \end{aligned} \tag{1.192}$$

per tant, l'actuació sobre un estat d'entrada no entrellaçat serà efectivament fer el swap dels qubits

$$(\alpha|0\rangle + \beta|1\rangle) \otimes (\rho|0\rangle + \gamma|1\rangle) \rightarrow (\rho|0\rangle + \gamma|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle) \tag{1.193}$$

mentre que per un de general serà

$$\sum_{a,b} c_{a,b} |a, b\rangle \rightarrow \sum_{a,b} c_{a,b} |b, a\rangle \tag{1.194}$$

### Circuit quàntic per crear estats de la base Bell

Sabem que una porta Hadamard ens porta de  $|0\rangle$  a  $(|0\rangle + |1\rangle)/\sqrt{2}$ . Analitzem doncs el circuit de la figura 1.46 quant a l'entrada tenim el  $|00\rangle$ .

Després de la porta H tindrem l'estat  $(|00\rangle + |10\rangle)/\sqrt{2}$  i si ara apliquem la porta C-NOT sabem que només canvia l'estat del segon qbit quant el primer està a 1 i per tant ens queda l'estat  $(|00\rangle + |11\rangle)/\sqrt{2}$ . En resum

$$|00\rangle \rightarrow \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \rightarrow \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Però què passa amb altres estats d'entrada (o si volem, l'entrada segueix sent  $|00\rangle$  però a aplicant una porta Pauli-X per passar de  $|0\rangle$  a  $|1\rangle$  quan convingui, abans d'entrar al circuit de la figura 1.46). Aleshores és fàcil comprovar que tenim els resultats donats en la figura 1.46.

In	Out
$ 00\rangle$	$( 00\rangle +  11\rangle)/\sqrt{2} =  \beta_{00}\rangle$
$ 01\rangle$	$( 01\rangle +  10\rangle)/\sqrt{2} \equiv  \beta_{01}\rangle$
$ 10\rangle$	$( 00\rangle -  11\rangle)/\sqrt{2} \equiv  \beta_{10}\rangle$
$ 11\rangle$	$( 01\rangle -  10\rangle)/\sqrt{2} \equiv  \beta_{11}\rangle$

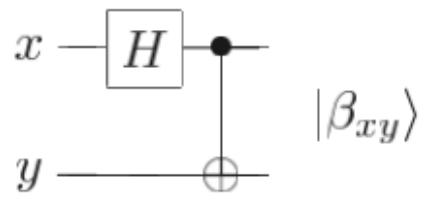


Figure 1.46: Circuit quàntic per crear estats de la base Bell

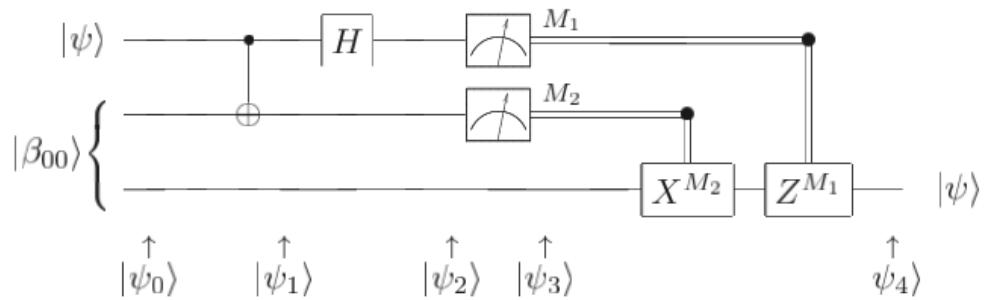


Figure 1.47: Circuit quàntic per teleportar un estat arbitrari d'Alice a Bob

### Circuit per teleportar

El circuit de la figura 1.47 serveix per teleportar un estat arbitrari  $|\psi\rangle_A = \alpha|0\rangle + \beta|1\rangle$  d'Alice a Bob. Les dues línies superiors del circuit representen el sistema d'Alice, mentre que la línia inferior és el sistema de Bob. Recordem que els metres representen una mesura i les línies dobles que en surten porten bits clàssics en lloc de qbits (línies simples).

Per teleportar  $|\psi\rangle$  ens ajudem de l'estat de Bell  $|\beta_{00}\rangle$  que Alice i Bob comparteixen i que sabem que podem crear amb el circuit de la figura 1.46.

Inicialment l'estat és

$$|\psi\rangle_0 = |\psi\rangle_A |\beta_{00}\rangle_{AB} = (\alpha|0\rangle + \beta|1\rangle) \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}} [\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|00\rangle + |11\rangle)] \quad (1.195)$$

on fem servir la convenció que els dos primers qbits (a l'esquerra) pertanyen a Alice i el tercer qbit a Bob. Com hem explicat anteriorment, el segon qbit d'Alice i el qbit de Bob formen un estat EPR.

Primerament, Alice passa els seus dos qbits per una porta CNOT, aconseguint

$$|\psi\rangle_1 = \frac{1}{\sqrt{2}} [\alpha|0\rangle(|00\rangle + |11\rangle) + \beta|1\rangle(|10\rangle + |01\rangle)] \quad (1.196)$$

i després fa passar el seu primer qbit per una porta Hadamard que ens porta a

$$|\psi\rangle_2 = \frac{1}{2} [\alpha(|0\rangle + |1\rangle)(|00\rangle + |11\rangle) + \beta(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)] \quad (1.197)$$

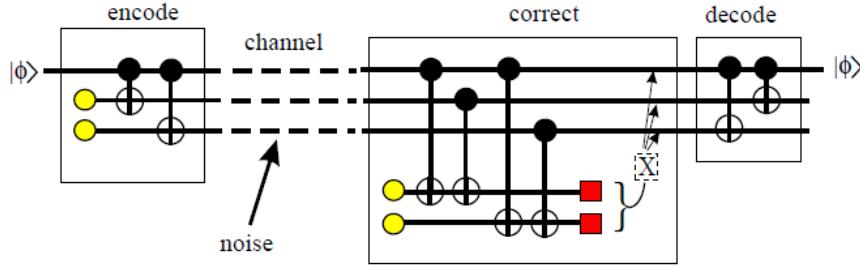


Figure 1.48: Mètode per corregir el soroll d'un canal que amb probabilitat  $p$  canvia  $|0\rangle \leftrightarrow |1\rangle$

estat que podem reescriure de la forma convenient:

$$|\psi\rangle_2 = \frac{1}{2} [|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle) + |10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)] \quad (1.198)$$

Ara Alice mesura els seus dos qbis (les mesures sempre s'entenen que es fan en la base computacional  $\{|0\rangle, |1\rangle\}$  si no s'especifica una altre cosa). En funció del resultat de la mesura dels dos qbits d'Alice, l'estat de Bob col·lapsarà en algun d'aquest 4 estats

$$\begin{aligned} 00 &\rightarrow |\psi\rangle_3 = \alpha|0\rangle + \beta|1\rangle \\ 01 &\rightarrow |\psi\rangle_3 = \alpha|1\rangle + \beta|0\rangle \\ 10 &\rightarrow |\psi\rangle_3 = \alpha|0\rangle - \beta|1\rangle \\ 11 &\rightarrow |\psi\rangle_3 = \alpha|1\rangle - \beta|0\rangle \end{aligned} \quad (1.199)$$

Aleshores Alice només ha d'enviar els dos bits  $M_1, M_2$  resultants de la seva mesura a Bob i Bob pot "arreglar" el seu estat per recuperant  $|\psi\rangle$ , aplicant la porta quàntica adequada. Per exemple, en el cas que la mesura produueixi 00, Bob no necessita fer res. Si la mesura és 01, Bob pot arreglar el seu estat aplicant la porta X. Si la mesura és 10, Bob pot arreglar el seu estat aplicant la porta Z. Si la mesura és 11, Bob pot solucionar el seu estat aplicant primer una porta X i després una porta Z. En resum, Bob ha d'aplicar la transformació unitària  $Z^{M_1}X^{M_2}$  per recuperar  $|\psi\rangle$ .

### Circuit quàntic per la correcció d'errors de transmissió.

Suposem que volem transmetre un qbit a través d'un canal de comunicació sorollós que actua de la següent manera: amb probabilitat  $1 - p$  surt el qbit sense canvis i amb la probabilitat complementaria  $p$  s'aplica l'operador Pauli-X sobre al qbit ( $\sigma_x$ , es a dir,  $|0\rangle \leftrightarrow |1\rangle$ ). Això és un tipus de soroll molt artificial, però ens servirà perfectament per entendre com funciona la correcció per a tipus de soroll molt més realistes.

Un mètode de correcció per aquest soroll ve donat en la figura 1.48 que com veurem serveix per reduir la probabilitat de canvi del qbit, passarem del  $p$  que dona el canal a  $o(p^2)$ .

Sigui  $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$  l'estat (arbitrari) del qbit que volem transmetre. Ens ajudarem de dos qbits auxiliars en l'estat  $|0\rangle$  (punts grocs de la figura 1.48). Aleshores el circuit "encode" fa

$$(\alpha|0\rangle + \beta|1\rangle)|0\rangle|0\rangle = \alpha|000\rangle + \beta|100\rangle \rightarrow \alpha|000\rangle + \beta|110\rangle \rightarrow \alpha|000\rangle + \beta|111\rangle \quad (1.200)$$

Ara enviem els tres qbits pel canal on sabem que degut al soroll tenim una probabilitat  $p$  de flip  $|0\rangle \leftrightarrow |1\rangle$ . Com a resultat, a la sortida del canal (i entrada del circuit "correct") tenim els possibles estats donats a l'esquerra de la figura 1.49 amb les seves probabilitats.

state	probability	state	probability
$a 000\rangle + b 111\rangle$	$(1-p)^3$	$(a 000\rangle + b 111\rangle) 00\rangle$	$(1-p)^3$
$a 100\rangle + b 011\rangle$	$p(1-p)^2$	$(a 100\rangle + b 011\rangle) 11\rangle$	$p(1-p)^2$
$a 010\rangle + b 101\rangle$	$p(1-p)^2$	$(a 010\rangle + b 101\rangle) 10\rangle$	$p(1-p)^2$
$a 001\rangle + b 110\rangle$	$p(1-p)^2$	$(a 001\rangle + b 110\rangle) 01\rangle$	$p(1-p)^2$
$a 110\rangle + b 001\rangle$	$p^2(1-p)$	$(a 110\rangle + b 001\rangle) 01\rangle$	$p^2(1-p)$
$a 101\rangle + b 010\rangle$	$p^2(1-p)$	$(a 101\rangle + b 010\rangle) 10\rangle$	$p^2(1-p)$
$a 011\rangle + b 100\rangle$	$p^2(1-p)$	$(a 011\rangle + b 100\rangle) 11\rangle$	$p^2(1-p)$
$a 111\rangle + b 000\rangle$	$p^3$	$(a 111\rangle + b 000\rangle) 00\rangle$	$p^3$

Figure 1.49: Estas a l'entrada i sortida del circuit "correct"

Posteriorment en el circuit "correct" utilitzem dos nous qbits auxiliars en l'estat  $|0\rangle$  que mitjançant portes C-NOT canviem el seu estat en funció de l'estat dels tres qbits que hem enviat pel canal. El resultat del circuit "correct" (abans de les dues mesures que es indiquen en vermell) queda resumit a la dreta de la figura 1.49.

Si ara mesurem els dos qbits auxiliars que hem utilitzat en "correct", en funció dels resultats, podrem actuar sobre els tres qbits que hem enviat pel canal de forma que augmenti la probabilitat de que el seu estat sigui  $\alpha|000\rangle + \beta|111\rangle$ . Per això s'ha d'aplicar la porta Pauli-X segons el resultat:

$$\begin{aligned}
 00 &\rightarrow \text{no fer res} \\
 01 &\rightarrow \sigma_x \text{ sobre qbit}_3 \\
 10 &\rightarrow \sigma_x \text{ sobre qbit}_2 \\
 11 &\rightarrow \sigma_x \text{ sobre qbit}_2
 \end{aligned} \tag{1.201}$$

(fet indicat dins del mateix circuit "correct" de la figura 1.48 com X i fletxes apuntant les 3 qbits de d'alt).

Per exemple, suposem que el resultat ha estat 10. Aleshores de la part dreta de 1.49 veiem que l'estat dels 3 qbits que hem enviat per canal és  $\alpha|010\rangle + \beta|101\rangle$  amb probabilitat  $p(1-p)^2$  o  $\alpha|101\rangle + \beta|010\rangle$  amb probabilitat  $p^2(1-p)$ . Com el primer és més probable, decidim corregir aplicant una porta Pauli-X al segon qbit. Aleshores a la sortida del circuit "correct" sabem que si hem mesurat 10 i hem fet aquesta correcció, tenim l'estat  $\alpha|000\rangle + \beta|111\rangle$  amb probabilitat  $p(1-p)^2$  o  $\alpha|111\rangle + \beta|000\rangle$  amb probabilitat  $p^2(1-p)$ .

Finalment, per treure el entrelaçament dels 3 qbits, apliquem el circuit "decode" de la figura 1.48, obtenint l'estat  $(\alpha|0\rangle + \beta|1\rangle)|0\rangle|0\rangle$  o  $(\alpha|1\rangle + \beta|0\rangle)|0\rangle|0\rangle$ , sent el primer molt més probable.

Amb aquest circuit, la probabilitat final d'error és igual a la probabilitat de que almenys dos dels 3 qbits que hem enviat pel canal estiguin corruptes i això passa amb una probabilitat

$$3p^2(1-p) + p^3$$

que és més petita que  $p$  (sempre i quant  $p < 1/2$ ) i per tant hem millorar l'error de transmissió.

### 1.11.3 Algorismes quàntics

Tot això de portes quàntiques pinta bé, però com les implementem i com les programem?. La primera pregunta la respondrem més endavant quan parlem de trampes de ions, que tenen l'avantatge de generar uns qbits "nets" (perden difícilment la coherència) però tenen el desavantatge del nombre baix de qbits que de moment es poden implementar. La segona pregunta és de més difícil resposta. De fet, i com veuren en els dos següents exemples, no hi ha cap

relació "algorítmica" de com, donat un problema, podem solucionar-lo en un ordinador quàntic. L'aproximació actual es més aviat de idea feliç: primer trobar com codificar el problema/solució en un estat quàntic i segon trobar quins passos, que es puguin sempre implementar en transformacions unitàries, i quines mesures hem de fer, de forma que ens ajudin a portar-nos fins la solució. Evidentment això no acota de cap manera la forma de solucionar un problema donat, però entenem que un cop trobada la solució, si que es podria implementar en un ordinador quàntic si totes les transformacions que hem fet són unitàries.

A continuació donarem dos exemples d'algoritmes quàntics que solucionen dos problemes concrets. Aquí entenem per algoritme el donar:

- quin és l'estat quàntic que codifica el problema/solució
- quines transformacions unitàries hem de fer sobre aquest estat
- quines mesures hem de fer, tan a meitat del procés per prendre decisions i/o fer correccions com les finals per obtenir resultats.

### Algorisme de cerca de Grover

El problema és buscar un element en una base de dades no estructurada. Per exemple, suposem que se us proporciona un número de telèfon d'un llistin telefònic ordenat alfabèticament i heu de saber a qui pertany. Haureu de revisar tots els números de telèfon i comprovar els noms dels propietaris registrats en cada cas. La cerca d'un article en una base de dades no classificada amb mida  $N$  costa un temps d'execució de l'ordinador clàssic  $O(N)$ , ja que de mitjana cal comprovar  $N/2$  entrades.

Pot un ordinador quàntic cercar una agulla en un paller de manera molt més eficient que el seu homòleg clàssic? Grover, el 1996, va respondre afirmativament a aquesta pregunta proposant un algorisme de cerca que consultés la base de dades només  $O(\sqrt{N})$  vegades.

Els 3 punts que defineixen l'algoritme de Grover són

- Donada una base de dades ordenada de  $N$  elements, ens contruïm l'estat de  $\log_2(N)$  qubits

$$|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_{x=1}^N |x\rangle = \frac{1}{\sqrt{N}}|00\cdots 000\rangle + \frac{1}{\sqrt{N}}|00\cdots 001\rangle + \frac{1}{\sqrt{N}}|00\cdots 010\rangle + \cdots + \frac{1}{\sqrt{N}}|11\cdots 111\rangle$$

on cada entrada de la base de dades està associada al vector  $|x\rangle$  que indica la seva posició. Inicialment doncs, totes les  $N$  entrades tenen el mateix pes ( $1/\sqrt{N}$ ).

- Si  $|a\rangle$  és l'element que estem buscant assumirem que tenim un operador ORACLE que actua com

$$U_a|x\rangle = \begin{cases} 1 & \text{si } x \neq a \\ -1 & \text{si } x = a \end{cases}$$

(seria com donar un nom del llistin i comprovar que efectivament aquest nom correspon al número que estem buscant). Aleshores podemaprofitar el paralelisme innat de la MQ per actuar simultàniament sobre tots els estats:

$$U_a \left( \frac{1}{\sqrt{N}} \sum_x |x\rangle \right) = \frac{1}{\sqrt{N}}(|0\rangle + |1\rangle + \cdots - |a\rangle + \cdots + |N\rangle)$$

- Procés d'amplificació: donat l'estat anterior on tenim l'estat que hem de trobar amb una fase diferent de la resta, farem de l'ordre de  $O(\sqrt{N})$  transformacions unitàries de forma que l'estat final sigui

$$|\psi_f\rangle \approx |a\rangle$$

de forma que quan fem una mesura de l'estat final en la base computacional obtinguem, amb molt bona aproximació, el valor  $a$ .

Anem a veure a continuació com podem implementar cada un d'aquests 3 passos.

### 1. ESTAT INICIAL

Recordem que la porta Hadamard actuant sobre un qbit passa del  $|0\rangle$  ( $|1\rangle$ ) a l'estat  $(|0\rangle + |1\rangle)/\sqrt{2}$  ( $(|0\rangle - |1\rangle)/\sqrt{2}$ ). Partint de l'estat  $|00\rangle$  de dos qbits i aplicant una porta de Hadamard a cada un d'ells, obtenim

$$H_1 \otimes H_2 |00\rangle = (|0\rangle + |1\rangle)/\sqrt{2} \otimes (|0\rangle + |1\rangle)/\sqrt{2} = \frac{1}{\sqrt{4}} (|00\rangle + |01\rangle + |10\rangle + |11\rangle) \quad (1.202)$$

l'estat superposició uniforme per  $n = 2$ .

Per tant, per obtenir l'estat inicial l'únic que haurem de fer és tenir  $n = \log_2(N)$  qbits, tots en l'estat  $|0\rangle$ , i aplicar  $H^{\otimes n}$ , de forma que

$$H_1 \otimes H_2 \otimes \dots \otimes H_n |00\dots00\rangle = \frac{1}{\sqrt{N}} |00\dots000\rangle + \frac{1}{\sqrt{N}} |00\dots001\rangle + \frac{1}{\sqrt{N}} |00\dots010\rangle + \dots + \frac{1}{\sqrt{N}} |11\dots111\rangle$$

### 2. ORACLE

Una forma difícil de implementar-ho, però no impossible, és suposar que podem construir un operador  $U_a$  que actuant sobre qualsevol combinació dels estats de la base computacional  $|\psi\rangle = \sum_x \alpha_x |x\rangle$  actua com

$$U_a \left( \sum_x \alpha_x |x\rangle \right) = \alpha_0 |0\rangle + \alpha_1 |1\rangle + \dots - \alpha_a |a\rangle + \dots + \alpha_N |N\rangle$$

és a dir, canvia el signe del terme  $|a\rangle$ . Evidentment aquest operador vindrà representat per una matriu unitària de  $N \times N$ , diagonal amb tots els termes igual a 1 exceptuant el terme  $(U_a)_{aa} = -1$ .

Are bé, la implementació més realista és suposar que és fàcil disposar d'una una funció booleana  $f : \{1, \dots, N\} \rightarrow \{0, 1\}$  tal que  $f(a) = 1$ , i la resta val 0. Ho podem pensar com una taula de mida  $N$ , on exactament un element té valor 1 i tots els altres són 0.

Aleshores, el nostre "operador" oracle actua de la següent manera (cf. algorisme de Deutsch-Jozsa) sobre un dels registres de la base de dades  $|x\rangle$  i utilitzant també un qbit auxiliar  $|q\rangle$ :

$$O|x\rangle|q\rangle = |x\rangle|q \oplus f(x)\rangle, \quad (1.203)$$

on suposem que  $f$  es pot calcular clàssicament en temps polinòmic i per aquest motiu és més realista la seva implementació.

Si apliquéssim aquest oracle sobre un estat general mentre que el qbit auxiliar és  $|0\rangle$ , tindríem

$$\sum_x \alpha_x |x\rangle |0\rangle \rightarrow \sum_x \alpha_x |x\rangle |f(x)\rangle \quad (1.204)$$

i si el qbit auxiliar també està en superposició, tenim

$$\begin{aligned} \sum_x \alpha_x |x\rangle \left( \frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) &\rightarrow \sum_x \alpha_x \left( \frac{|x\rangle |f(x)\rangle - |x\rangle |\bar{f}(x)\rangle}{\sqrt{2}} \right) \\ &= \sum_x \alpha_x |x\rangle \left( \frac{|f(x)\rangle - |\bar{f}(x)\rangle}{\sqrt{2}} \right) = \sum_x \alpha_x |x\rangle (-1)^{f(x)} \left( \frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) \end{aligned} \quad (1.205)$$

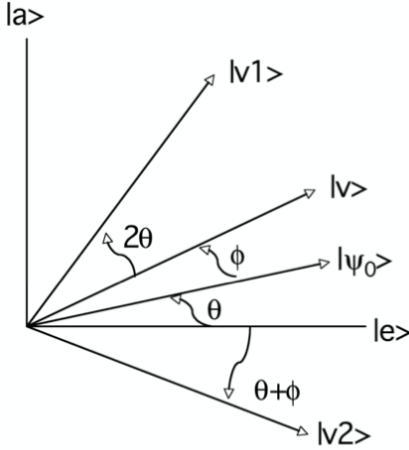


Figure 1.50: Per girar  $|\nu\rangle$  de  $2\theta$  a  $|\nu\rangle_1$ , reflectim al voltant de  $|e\rangle$ , arribant a  $|\nu\rangle_1$ , i després reflectim al voltant de  $|\psi_0\rangle$ .

doncs  $\bar{f}(x)$  significa el complement binari de  $f(x)$ , és a dir, si  $f(x) = 1$ , llavors  $\bar{f}(x) = 0$ . En el darrer pas hem utilitzat el fet que si  $f(x) = 0$  el qbit auxiliar és  $|0\rangle - |1\rangle = (-1)^{f(0)}(|0\rangle - |1\rangle)$ , mentre que si  $f(x) = 1$  el qbit auxiliar és  $|1\rangle - |0\rangle = (-1)^{f(1)}(|0\rangle - |1\rangle)$ .

Per tant, podem resumir l'acció de l'oracle sobre un estat arbitrari com

$$O\left(\sum_x \alpha_x |x\rangle\right) = \sum_x (-1)^{f(x)} \alpha_x |x\rangle \quad (1.206)$$

on hem omès el qbit auxiliar a banda i banda de l'equació perquè es troba en el mateix estat  $(|0\rangle - |1\rangle)/\sqrt{2}$ . Però no oblidem que aquest qbit auxiliar és essencial perquè es realitzi aquesta acció de l'oracle O sobre qualsevol estat. En aquest cas aquest operador vindrà representat per una matriu unitària de  $N \times N$ , diagonal amb valors  $O_{ii} = (-1)^{f(i)}$ , es a dir, tots igual a 1 exceptuant el terme  $O_{aa} = -1$ .

### 3. AMPLIFICACIÓ

Com identificar el terme que ha canviat de signe (el  $|a\rangle$ ) i per tant determinar la solució?. La idea és anar fent petites transformacions/rotacions que ens portin l'estat inicial  $|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle$  a l'estat buscat  $|a\rangle$ .

Per això pensem en el subespai bidimensional que consta dels dos estats:  $|a\rangle$  i la superposició uniforme  $|\psi_0\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle$ . Sigui  $\theta$  l'angle entre  $|\psi_0\rangle$  i  $|e\rangle$ , on  $|e\rangle$  és el vector que és ortogonal a  $|a\rangle$  (en el pla determinat per  $|a\rangle$  i  $|\psi_0\rangle$ , tal com es mostra en la figura 1.50). Aquest angle ha de ser molt petit doncs

$$\langle \psi_0 | a \rangle = \cos(\pi/2 - \theta) = \sin(\theta) = \frac{1}{\sqrt{N}} \sum_x \langle x | a \rangle = \frac{1}{\sqrt{N}} \sum_x \delta_{x,a} = \frac{1}{\sqrt{N}} \quad (1.207)$$

i per tant  $\sin \theta \approx \theta \approx 1/\sqrt{N}$ .

La idea aleshores és començar per un estat  $|\nu\rangle_0 = |\psi_0\rangle$  i anar aplicant-li rotacions d'angle  $2\theta$  els cops que siguin necessàries fins apropar-se al nostre objectiu que és  $|a\rangle$ . Per tant, necessitem  $O(\frac{\pi/2}{2/\sqrt{N}}) = O(\sqrt{N})$  iteracions per que ens acostem molt a  $|a\rangle$ , i després amb una alta probabilitat, una mesura de l'estat a la base computacional (tots els qbits) produirà  $a$ . Evidentment no s'ha d'iterar més enllà d'aquest punt. Les iteracions posteriors giraran de nou el vector  $|\nu\rangle$  lluny de  $|a\rangle$ . Es pot demostrar que per a  $N$  gran, iterarant  $r = \pi\sqrt{N}/4$  vegades, la corresponent probabilitat d'error és  $O(1 - \cos^2 \theta) = O(\sin^2 \theta) = O(N^{-1})$ .

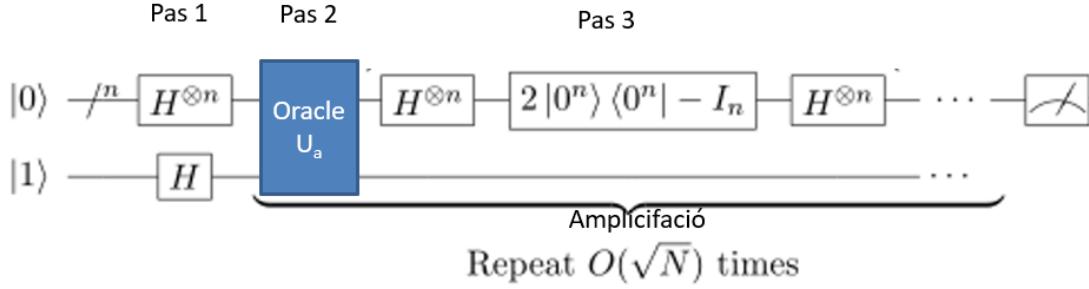


Figure 1.51: Circuit de Grover

Com aconseguim fer aquestes rotacions d'angle  $2\theta$  en cada iteració? Suposem que ja hem fet varies iteracions i estem en l'estat  $|\nu\rangle$  de la figura 1.50. Una manera de girar aquest vector un angle  $2\theta$  és fer dues reflexions: reflectint primer sobre  $|e\rangle$  i després reflectint sobre  $|\psi_0\rangle$ . Aquesta transformació també s'il-lustra a la figura 1.50. La primera reflexió transforma un vector arbitrari  $|\nu\rangle$  a  $|\nu_2\rangle$  i la segona reflexió transforma  $|\nu_2\rangle$  a  $|\nu_1\rangle$ .

Com implementem les dues reflexions?

- La reflexió sobre  $|e\rangle$  és fàcil.  $|e\rangle$  és el vector ortogonal a  $|a\rangle$ , de manera que tot el que hem de fer és capgirar la fase del component de la funció d'ona de la base de dades en la direcció de  $|a\rangle$ , és a dir, enviem qualsevol component  $|a\rangle$  a  $-|a\rangle$  i deixeu tots els altres components tal qual. Per aconseguir-ho, només actuem amb l'oracle:

$$O|\nu\rangle = O \left( \sum_x \alpha_x |x\rangle \right) = \sum_{x \neq a} \alpha_x |x\rangle - \alpha_a |a\rangle = (I - 2|a\rangle\langle a|) |\nu\rangle \quad (1.208)$$

i per tant

$$R_{|e\rangle} = O = I - 2|a\rangle\langle a| \quad (1.209)$$

La notació  $R_{|e\rangle}$  significa una reflexió sobre el vector  $|e\rangle$ .

- Què passa amb la reflexió sobre  $|\psi_0\rangle$ ? Recordem que  $H^{\otimes n}$  passa de l'estat  $|00\dots 00\rangle$  a la superposició uniforme de  $n$  qbits  $|\psi_0\rangle$ , i a l'inrevés (doncs  $H^\dagger = H$ ). Per tant, primer apliquem  $H^{\otimes n}$  per passar de  $|\psi_0\rangle \rightarrow |00\dots 0\rangle$ , a continuació, reflexionem al voltant de  $|00\dots 0\rangle$  i, finalment, apliqueu  $H^{\otimes n}$  per tornar a la base original.

Com una reflexió al voltant del vector  $|00\dots 00\rangle$  vol dir canviat totes les fases de signe menys la del  $|00\dots 00\rangle$ , la rotació que estem buscant és

$$R_{|\psi_0\rangle} = H^{\otimes n} (-I + 2|00\dots 00\rangle\langle 00\dots 00|) H^{\otimes n} \quad (1.210)$$

Per fer un pas d'iteració combinem les dues reflexions, donant l'operador Grover (o operador d'amplificació):

$$G = R_{|\psi_0\rangle} R_{|e\rangle}. \quad (1.211)$$

Aleshores només hem d'aplicar aquest operador de Grover  $r = \pi\sqrt{N}/4$  vegades per passar de  $|\psi_0\rangle$  a aproximadament  $|a\rangle$ , i la corresponent probabilitat d'error és  $O(N^{-1})$ .

En la figura 1.51 es pot veure el circuit que implementa l'algoritme de Grover amb els tres passos que hem comentat i els seus efectes en la figura 1.52.

Com exemple considerem el circuit de la figura 1.53 que implementa l'algorisme de Grover utilitzant 3 qbits, es a dir, cerca sobre una base de  $2^3$  entrades, i on hem utilitzat el fet que

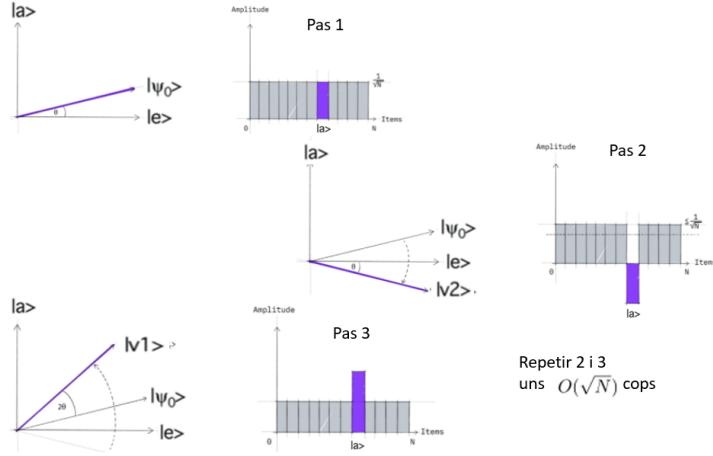


Figure 1.52: Els 3 passos de l'algoritme de Grover

$(X \otimes X \otimes X)Z_{controlada}(X \otimes X \otimes X) = (-1)(2|0\rangle\langle 0|^3 - I^{\otimes 3})$ , es a dir, és l'operador que implementa la segona reflexió exceptuant una fase global  $(-1)$  irrellevant.

Els operadors que apareixen en aquest circuit són:

$$H \otimes H \otimes H = \begin{pmatrix} \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} & -\frac{1}{2\sqrt{2}} \end{pmatrix}$$

$$X \otimes X \otimes X = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$Z_{controlada} = (P_1^{(1)} \otimes P_1^{(2)}) \otimes Z^{(3)} + (I^{(1)} \otimes I^{(2)} - P_1^{(1)} \otimes P_1^{(2)}) \otimes I^{(3)} =$$

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

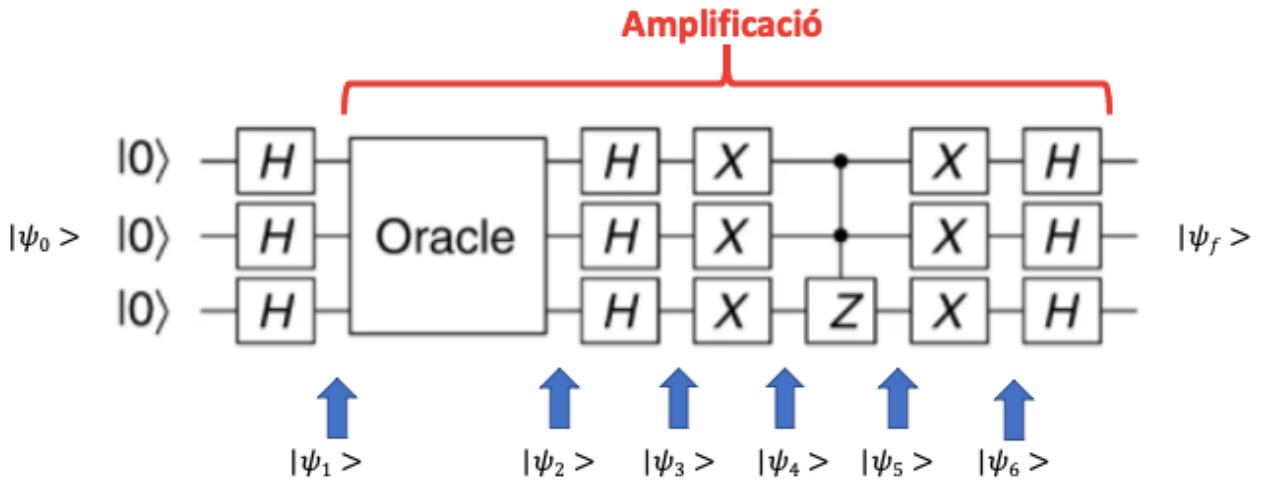


Figure 1.53: Circuit de Grover amb 3 qbits

i per tant, si assumim que l'Oracle "selecciona" l'element  $|4\rangle = |100\rangle$ , els estats que es van produint al llarg del circuit són

$$\langle\psi_0| = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\langle\psi_1| = \left( \frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \right)$$

$$\langle\psi_2| = \left( \frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \quad -\frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \quad \frac{1}{2\sqrt{2}} \right)$$

$$\langle\psi_3| = \left( \frac{3}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \right)$$

$$\langle\psi_4| = \left( \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad \frac{3}{4} \right)$$

$$\langle\psi_5| = \left( \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad -\frac{3}{4} \right)$$

$$\langle\psi_6| = \left( -\frac{3}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad -\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{4} \right)$$

$$\langle\psi_f| = \left( -\frac{1}{4\sqrt{2}} \quad -\frac{1}{4\sqrt{2}} \quad -\frac{1}{4\sqrt{2}} \quad -\frac{1}{4\sqrt{2}} \quad -\frac{5}{4\sqrt{2}} \quad -\frac{1}{4\sqrt{2}} \quad -\frac{1}{4\sqrt{2}} \quad -\frac{1}{4\sqrt{2}} \right)$$

Aleshores, les probabilitats d'obtenir cada un dels elements de la base habitual  $\{|n\rangle\}$ , ( $n = 0, \dots, 7$ ) al fer una mesura sobre l'estat final  $|\psi_f\rangle$  són

$$probabilitats = \left\{ \frac{1}{32}, \frac{1}{32}, \frac{1}{32}, \frac{1}{32}, \frac{25}{32}, \frac{1}{32}, \frac{1}{32}, \frac{1}{32} \right\}$$

el valor esperat de  $n$  i la seva dispersió.

$$\langle n \rangle = 3.9 + / - 1.2$$

Si al circuit de la figura 1.53 li afegim un grup addicional de transformacions que hem designat en la mateixa figura com "amplificació", l'estat final i el valor esperat de  $n$  són:

$$\langle \psi_{f2} | = \left( -\frac{1}{8\sqrt{2}} \quad -\frac{1}{8\sqrt{2}} \quad -\frac{1}{8\sqrt{2}} \quad -\frac{1}{8\sqrt{2}} \quad \frac{11}{8\sqrt{2}} \quad -\frac{1}{8\sqrt{2}} \quad -\frac{1}{8\sqrt{2}} \quad -\frac{1}{8\sqrt{2}} \right)$$

$$probabilitats_2 = \left\{ \frac{1}{128} \quad \frac{1}{128} \quad \frac{1}{128} \quad \frac{1}{128} \quad \frac{121}{128} \quad \frac{1}{128} \quad \frac{1}{128} \quad \frac{1}{128} \right\}$$

$$\langle n \rangle = 4. + / - 0.59$$

Si afegim una tercera amplificació, l'estat final i el valor esperat de  $n$ . són:

$$\langle \psi_{f3} | = \left( \frac{7}{16\sqrt{2}} \quad \frac{7}{16\sqrt{2}} \quad \frac{7}{16\sqrt{2}} \quad \frac{7}{16\sqrt{2}} \quad -\frac{13}{16\sqrt{2}} \quad \frac{7}{16\sqrt{2}} \quad \frac{7}{16\sqrt{2}} \quad \frac{7}{16\sqrt{2}} \right)$$

$$probabilitats_3 = \left\{ \frac{49}{512} \quad \frac{49}{512} \quad \frac{49}{512} \quad \frac{49}{512} \quad \frac{169}{512} \quad \frac{49}{512} \quad \frac{49}{512} \quad \frac{49}{512} \right\}$$

$$\langle n \rangle = 3.6 + / - 2.$$

Veiem que amb aquesta tercera iteració ens allunya dels resultats esperats, La raó és que ja ens hem passat de la solució buscada. Com  $\cos(\pi/2 - \theta) = \langle 4 | \psi_1 \rangle = 1/\sqrt{8}$ ,  $\theta \approx 20^\circ$  i per tant, girem uns  $40^\circ$  en cada amplificació. Inicialment estem a  $70^\circ$  de  $|4\rangle$ .

- La primera amplificació ens porta a  $30^\circ$  de  $|4\rangle$ .
- La segona amplificació, encara que "sobrepassa"  $|4\rangle$ , ens deixa més apropi, a només a  $10^\circ$  de  $|4\rangle$ .
- La tercera amplificació ens deixa molt lluny, a  $50^\circ$  de  $|4\rangle$ .

### Algorisme de factorització de Shor

El primer algorisme per a una computadora quàntica va ser proposat per Peter Shor el 1994 amb l'objectiu de factoritzar. Mentre que l'algorisme clàssic més ràpid de factorització va com  $t \sim \exp(N)$ , l'algorisme de Shor és polinòmic  $t \sim N^p$

Per fer una comparació amb els ordinadors clàssics, si suposem que un superordinador ràpid trigaria aproximadament un any a factoritzar un nombre de 150 díigits, el mateix ordinador requeriria aproximadament l'edat actual de l'univers per factoritzar un nombre de 400 díigits. L'acceptació que aquests són els ordres de temps necessaris per factoritzar un número en els ordinadors clàssics, constitueix la base dels sistemes de xifratge actuals. Si es pogués construir un ordinador quàntic que factoritzés un nombre de 150 díigits en un mes, llavors el mateix ordinador quàntic podria factoritzar un nombre de 400 díigits en aproximadament un any.

L'algorisme de Shor fa ús del fet que els ordinadors quàntics serien bons per trobar el període d'una funció periòdica, que es pot relacionar per la teoria de nombres amb el problema de factoritzar un nombre. Com qualsevol algorisme quàntic, és probabilístic amb probabilitat  $P = 1 - \epsilon$  de trobar la solució. Per tant, la probabilitat que Shor sigui exitós si l'executem k cops, serà  $P = 1 - \epsilon^k$ .

L' algorisme de Shor es descompon en dues grans parts:

- Reducció del problema de descompondre en factors al problema de trobar un període, com veurem a continuació
- Trobar un algoritme quàntic per solucionar el problema de trobar el període.

Necessitarem uns quants passos per poder solucionar el problema que ens hem proposat. La formulació clàssica de problema seria:

1. Escollim un nombre pseudo-aleatori  $a < N$  i calculem el màxim comú divisor de  $a$  i  $N$ , és a dir  $mcd(a, N)$ . Si  $mcd(a, N) \neq 1$ , a ja serà un factor no trivial de  $N$  i així que ja hem acabat.
2. En el cas que  $mcd(a, N) = 1$  haurem de trobar el període  $r$  de  $f(x) = a^x \text{mod}(N)$ , es a dir, hem de trobar el nombre més petit  $r$  pel qual  $f(x + r) = f(x)$ . Per exemple: si  $N = 15$  i  $a = 7$  i obtindrem el següent:

$$\begin{aligned}
 7^0 \text{mod}15 &= 1 \\
 7^1 \text{mod}15 &= 7 \\
 7^2 \text{mod}15 &= 4 \\
 7^3 \text{mod}15 &= 13 \\
 7^4 \text{mod}15 &= 1
 \end{aligned} \tag{1.212}$$

i per tant, el període és  $r = 4$ .

3. Si trobem una  $r$  imparella haurem de tornar al pas número 1.
4. Si  $a^{r/2} = 1 \text{ mod}(N)$  hem de tornar al pas 1.
5. Per últim, els factors de  $N$  són el  $mcd(a^{r/2} \pm 1, N)$  i ja hem acabat.

El problema queda reduït a solucionar el pas de trobar el període utilitzant un ordinador Quàntic, doncs la resta de passos es poden fer de forma eficient en un ordinador clàssic.

Anem ara a donar l'algorisme de Shor per trobar aquest període i com exemple ho farem per  $N = 15$

1. Triar un enter  $q$  de manera que  $N < q < 2^N$  (pel nostre exemple escollim  $q = 256$ ).
2. Triar un nombre enter aleatori  $a$  tal que  $mcd(a, N) = 1$  (pel nostre exemple seleccionem  $a = 7$ ).
3. Crear dos registres quàntics:

- Registre d'entrada: ha de contenir prou qbits per representar nombres tan grans com  $q - 1$  (pel nostre exemple fins a 255, de manera que necessitem 8 qbits)
- Registre de sortida: ha de contenir prou qbits per representar números tan grans com  $N - 1$  (pel nostre exemple fins a 14, de manera que necessitem 4 qbits)

4. Inicialització (suposem que tots els qbits estan a l'estat  $|0\rangle$ ): utilitzar portes Hadamard per portar el registre d'entrada a una superposició igual de ponderada de tots els enters de 0 a  $q - 1$  (0 a 255 en el nostre exemple). L'estat inicial del sistema en l'exemple considerat serà

$$|\psi\rangle = \left( \frac{1}{\sqrt{256}} \sum_{m=0}^{255} |m\rangle \right) \otimes |0000\rangle$$

5. Aplicar la transformació unitària  $a^m \text{mod}N$  a cada número  $m$  del registre d'entrada, emmagatzemant el resultat de cada càlcul al registre de sortida. En l'exemple considerat (recordem que hem escollit  $a = 7$ ) tenim

Input Register	$(7^m \text{Mod}15)$	Output Register
$ 0\rangle$	$7^0 \text{Mod}15$	1
$ 1\rangle$	$7^1 \text{Mod}15$	7

$ 2\rangle$	$7^2 \text{Mod}15$	4
$ 3\rangle$	$7^3 \text{Mod}15$	13
$ 4\rangle$	$7^4 \text{Mod}15$	1
$ 5\rangle$	$7^5 \text{Mod}15$	7
$ 6\rangle$	$7^6 \text{Mod}15$	4
$ 7\rangle$	$7^7 \text{Mod}15$	13

i per tant l'estat ara és:

$$|\psi\rangle = \frac{1}{\sqrt{256}} \sum_{m=0}^{255} (|m\rangle \otimes |7^m \text{Mod}15\rangle)$$

6. Fer una mesura al registre de sortida, fet que col·lapsarà la superposició del registre d'entrada. En el nostre cas, el registre de sortida pot col·lapsar a  $|1\rangle, |4\rangle, |7\rangle$  o  $|13\rangle$ . Si suposem, per exemple, que ha donat  $|1\rangle$ , aleshores el registre d'entrada es col·lapsarà a:

$$\frac{1}{\sqrt{64}}(|0\rangle + |4\rangle + |8\rangle + |12\rangle, \dots) = \frac{1}{\sqrt{64}} \sum_{m \in M} |m\rangle$$

on  $M = \{0, 4, 8, 12, \dots, 252\}$  (el nostre registre d'entrada es troba ara en una superposició igual de 64 valors  $(0, 4, 8, \dots, 252)$ ).

7. Aplicar la transformada de Fourier quàntica al registre d'entrada. La transformada de Fourier té l'efecte de prendre un estat  $|m\rangle$  i transformar-lo en un estat donat per:

$$|m\rangle \rightarrow \frac{1}{\sqrt{q}} \sum_{c=0}^q \exp(2\pi imc/q) |c\rangle \quad (1.213)$$

i per tant, en el nostre exemple, l'estat passa a ser (recordem que hem escollit  $q = 256$ ):

$$\frac{1}{\sqrt{64}} \sum_{m \in M} \frac{1}{\sqrt{256}} \sum_{c=0}^{256} \exp(2\pi imc/256) |c\rangle = \frac{1}{\sqrt{64}} \frac{1}{\sqrt{256}} \sum_{c=0}^{256} \left( \sum_{m \in M} \exp(2\pi imc/256) \right) |c\rangle \quad (1.214)$$

Aquest nou estat tindrà amplituds màximes per certs  $|c\rangle$ , quan totes les fases  $\exp(2\pi imc)$  donin 1. En el nostre exemple  $m \in \{0, 4, 8, 12, \dots, 252\}$  i recordem que  $0 \leq c \leq q = 256$ , aleshores resulta que aquests  $c$  són els múltiples enters de  $256/4 = 64$ , es a dir  $|0\rangle, |64\rangle, |128\rangle, |192\rangle$ . Per tant, les amplituds de probabilitat dels estats anteriors són ara superiors a la resta d'estats del nostre registre d'entrada.

8. Mesurar en registre d'entrada. En el nostre exemple col·lapsarà amb alta probabilitat a un d'aquests múltiples de 64  $\{0, 64, 128, 192\}$ , anomenem aquest valor  $p$ .
9. Amb el nostre coneixement de  $q$  i  $p$ , hi ha mètodes per calcular el període (un mètode és l'expansió de fracció contínua de la proporció entre  $q$  i  $p$ ). Això es pot fer en un ordinador clàssic i el resultat serà  $r = 4$ . Un cop determinat el període, els factors de  $N$  es poden determinar prenent el màxim comú divisor de  $N$  respecte a  $a^{r/2} + 1$  i  $a^{r/2} - 1$ . La idea aquí és que aquest càcul es farà també en un ordinador clàssic. En el nostre exemple on hem escollit  $a = 7$  i hem trobat  $r = 4$  tindrem

$$\begin{aligned} Gcd(7^{4/2} + 1, 15) &= 5 \\ Gcd(7^{4/2} - 1, 15) &= 3 \end{aligned} \quad (1.215)$$

i per tant, hem acabat de factoritzar  $N = 15$

Pot passar que aquest procediment ens doni el període equivocat. La probabilitat de que això passi depèn de la vostra elecció de  $q$ . Com més gran sigui la  $q$ , més gran serà la probabilitat de trobar el període correcte. Si es produueix algun d'aquests casos, tornem al principi i escollim una nova  $a$  o  $q$ .

### 1.11.4 Implementació física de computadores quàntiques

Ja hem discutit el cas ideal, però no factible, de crear computadors quàntics utilitzant SGs per crear qbits en la direcció desitjada, camps magnètics uniformes per "girar-los" i l'acoblament spin-spin per portar-los a estats entrellaçats. El que farem aquí és explicar una de les realitzacions actuals de computadors quàntics implementades utilitzant trampes de ions. En les trampes de ions manipulem els qbits mitjançant lasers de la freqüència i durada adequada. Anem a veure quines han de ser aquestes freqüències i temps estudiant els sistema de dos nivells sotmés a una pertorbació externa.

#### Pertorbació sinusoidal

Suposem un sistema amb un  $H_0$  que té dos nivells  $|a\rangle$  i  $|b\rangle$ , de manera que  $H_0 = E_a|a\rangle\langle a| + E_b|b\rangle\langle b|$ . Suposem també que el sotmetem a una pertorbació  $H'(t)$  harmònica que connecta els dos nivells esmentats:

$$\lambda H'(t) = \lambda e^{-i\omega t}|b\rangle\langle a| + \lambda e^{i\omega t}|a\rangle\langle b| \quad (1.216)$$

amb  $\omega > 0$  per definició. Un exemple pot ser una partícula de spin 1/2 dins un camp magnètic en direcció  $z$ , de manera que  $H_0 = \frac{e\hbar}{2m}B_z\sigma_z$ , amb els dos nivells no perturbats  $|S_z, +\rangle$  i  $|S_z, -\rangle$ . Aleshores es sotmet a una pertorbació donada per un camp magnètic giratori en el pla  $x, y$ . Aleshores  $\lambda H'(t) = \frac{e\hbar}{2m}B_{\text{pert}}(\cos(\omega t)\sigma_x + \sin(\omega t)\sigma_y)$ .

Suposarem  $E_b > E_a$ , de manera que  $\omega_{ba} > 0$ . Definim, com sempre,  $H'_{ba} = \langle b|H'|a\rangle = e^{-i\omega t}$ . L'estat en qualsevol instant de temps el podem escriure com

$$|\psi(t)\rangle = c_a(t)e^{-iE_a t/\hbar}|a\rangle + c_b(t)e^{-iE_b t/\hbar}|b\rangle \quad (1.217)$$

i l'equació de Schrödinger la podem escriure com

$$\begin{aligned} \begin{pmatrix} \dot{c}_a(t) \\ \dot{c}_b(t) \end{pmatrix} &= \frac{-i}{\hbar} \begin{pmatrix} H'_{aa} & H'_{ab} e^{i\omega_{ab}t} \\ H'_{ba} e^{i\omega_{ba}t} & H'_{bb} \end{pmatrix} \begin{pmatrix} c_a(t) \\ c_b(t) \end{pmatrix} \\ &= \frac{-i}{\hbar} \begin{pmatrix} 0 & \lambda e^{i(\omega - \omega_{ba})t} \\ \lambda e^{-i(\omega - \omega_{ba})t} & 0 \end{pmatrix} \begin{pmatrix} c_a(t) \\ c_b(t) \end{pmatrix} \end{aligned} \quad (1.218)$$

o

$$\dot{c}_a(t) = -i(\lambda/\hbar) e^{i(\omega - \omega_{ba})t} c_b(t) \quad (1.219)$$

$$\dot{c}_b(t) = -i(\lambda/\hbar) e^{-i(\omega - \omega_{ba})t} c_a(t) \quad (1.220)$$

El canvi a unes noves variables

$$x_a(t) \equiv e^{-i\frac{(\omega - \omega_{ba})t}{2}} c_a(t), \quad x_b(t) \equiv e^{i\frac{(\omega - \omega_{ba})t}{2}} c_b(t),$$

elimina la dependència temporal en els coeficients de les equacions diferencials anteriors, que llavors esdevenen de resolució molt fàcil. Suposem que, en l'instant inicial  $t = 0$ , el sistema es troba en l'estat fonamental  $|a\rangle$ ,

$$|c_a(0)| = 1, \quad c_b(0) = 0; \quad (1.221)$$

la solució és

$$\begin{aligned} c_b(t) &= -i \frac{(\lambda/\hbar)}{\omega_r} \exp(-i\xi t) \sin(\omega_r t) \\ c_a(t) &= -i \frac{1}{\omega_r} \exp(i\xi t) [\xi \sin(\omega_r t) + i\omega_r \cos(\omega_r t)] \end{aligned} \quad (1.222)$$

on

$$\xi = \frac{\omega - \omega_{ba}}{2}, \quad \omega_r \equiv \frac{1}{2} \sqrt{(\omega_{ba} - \omega)^2 + \lambda^2 4/\hbar^2} \quad (1.223)$$

Per tant,

$$\mathcal{P}_{a \rightarrow b}(t) = \lambda^2 \frac{4}{\hbar^2} \frac{\sin^2(\omega_r t)}{(2\omega_r)^2} \quad (1.224)$$

Aquesta solució, exacta, es coneix com a *fórmula de Rabi*.

Aquesta probabilitat de transició oscil·la sinusoidalment en funció del temps (figura 1.54) entre 0, per a temps  $t_n = n\pi/\omega_r$ , i el màxim  $\lambda^2 4/\hbar^2 (2\omega_r)^2 \leq 1$ . Aquest màxim només és igual a la unitat si  $\omega = \omega_{ba}$ , per a temps  $t_n = (2n+1)\pi\hbar/2\lambda$ .

D'altra banda, la probabilitat de transició en funció de  $\omega$  (figura 1.55) té un pic en  $\omega = \omega_{ba}$ , d'alçada  $\sin^2(\lambda t/\hbar)$  i d'amplada

$$4 \sqrt{\frac{\pi^2}{t^2} - \frac{\lambda^2}{\hbar^2}} \quad (1.225)$$

En les figures 1.54 i 1.55, es mostra l'evolució del sistema. És clar que, si inicialment estem en l'estat  $b$ , com que  $|H'_{ba}| = |H'_{ab}|$ , la probabilitat de transició a l'estat  $a$  serà la mateixa que hem trobat abans:

$$\mathcal{P}_{b \rightarrow a}(t) = \mathcal{P}_{a \rightarrow b}(t) \quad (1.226)$$

Per exemple, si il·luminem un àtom amb una certa llum monocromàtica de freqüència adequada  $\omega$ , podem provocar un salt d'un electró a una capa superior amb una certa probabilitat, que és exactament la mateixa probabilitat que l'electró salti a la capa inferior estant en aquella capa superior. (a més a més hem vist que aquesta probabilitat és exactament 1 quan hi ha resonància ( $\omega = \omega_{ba}$ ) i actuem durant un  $\Delta t = \pi\hbar/2\lambda$ ). El primer fenomen l'anomenarem *absorció estimulada*, ja que la radiació externa estimula el salt de l'electró cap a una capa superior absorbint un fotó. El segon fenomen l'anomenarem *emissió estimulada*, ja que la radiació externa estimula, i de fet amb la mateixa probabilitat que en el cas de l'absorció estimulada, el salt de l'electró de la capa superior a la inferior emetent un fotó.

### Trampes de ions

En la figura 1.56 es mostra una trampa d'ions amb una cadena d'ions atrapats. Les quatre pales estan a alta tensió (pales veïnes amb potencial oposat) oscil·lant a radiofreqüència, fet que proporciona un confinament tipus trampes Paul en les direccions radials. Els electrodes de punta dels extrems estan en alta tensió positiva i atrapen els ions axialment. Un làser s'adreça individualment als ions i en manipula el seu estat quàntic. La fluorescència de ressonància dels ions es detecta en una càmera CCD. A la part inferior de la figura es mostra la imatge CCD d'una cadena de vuit ions freds excitats amb làser (ions  $Ca^+$ ). La distància entre els ions extrems és d'uns  $70\mu m$ .

El qbit atòmic (o qbit d'informació doncs és on codificarem la informació quàntica) en el cas d'un ió requereix dos nivells estables que poden ser un estat fonamental i un estat excitat metastable connectat per una transició òptica prohibida ("qbit òptic"), o bé dos subnivells hiperfins de l'estat fonamental d'un ió amb spín nuclear diferent de zero ("qbit hiperfine"). Tots dos casos es tracten com a sistemes atòmics de dos nivells  $\{|g\rangle, |e\rangle\}$  o  $\{| \downarrow \rangle, | \uparrow \rangle\}$  i que utilitzarem com la base  $\{|0\rangle, |1\rangle\}$ .

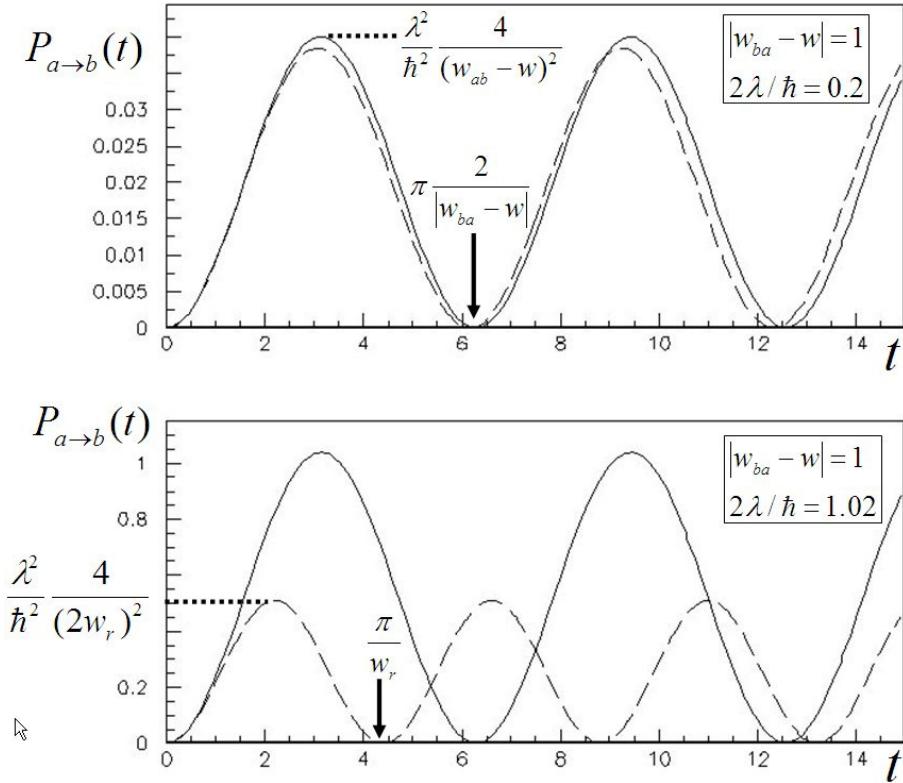


Figure 1.54: Probabilitat de transició en funció del temps. La línia contínua és la solució fins a ordre  $\lambda$  i la discontínua, la solució exacta.

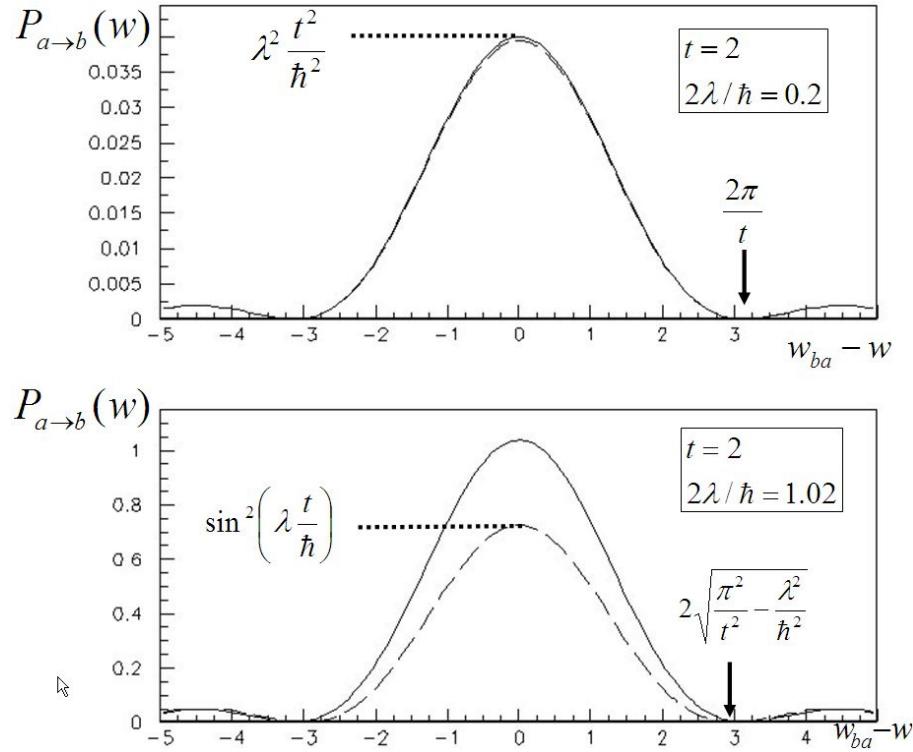


Figure 1.55: Probabilitat de transició en funció de la freqüència de la pertorbació. La línia contínua és la solució fins a ordre  $\lambda$  i la discontínua, la solució exacta.

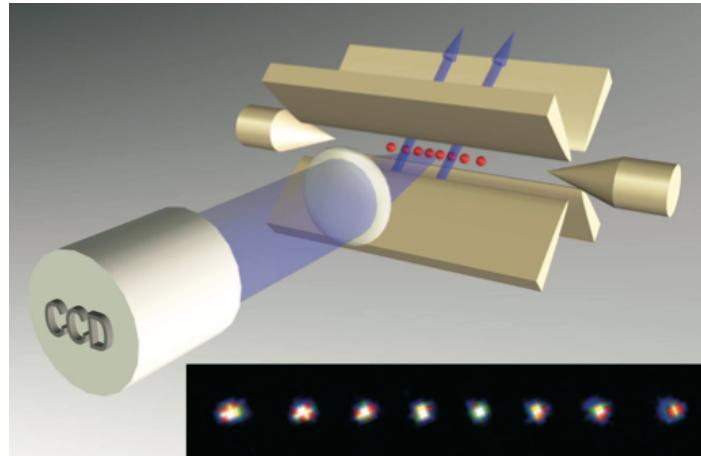


Figure 1.56: Trampa de ions

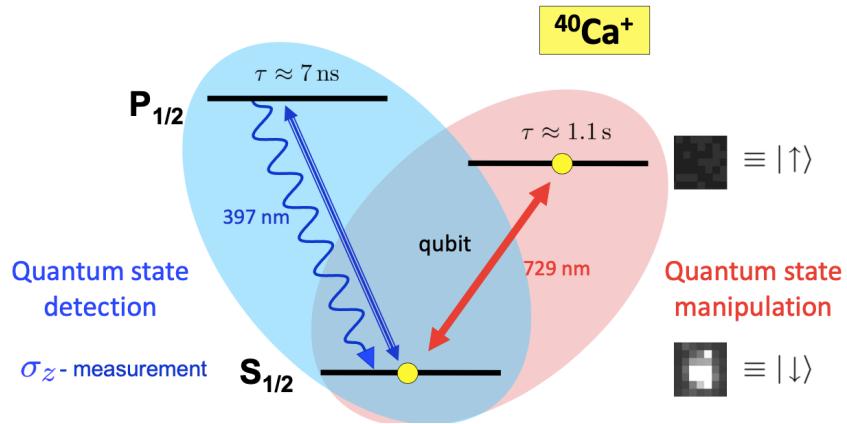


Figure 1.57: qbit atòmic i forma de detecció del seu estat.

Per exemple, en el qbit òptic de  $Ca^+$  (figura 1.57), i que a partir d'ara és el que utilitzarem com el qbit atòmic o d'informació, es seleccionen els subestats Zeeman particulars dels nivells  $S_{1/2}$  i  $D_{5/2}$  per guardar els dos estats del qbit (és on es troben les dues boletes de color groc). L'estat  $S_{1/2}$  es manté estable mentre que l'estat  $D_{5/2}$  té una vida mitja de  $\approx 1.1s$  i per tant els càlculs s'han de fer en temps més curts. Per saber si l'ió està en un estat o l'altre, es pot enviar un làser de la freqüència adequada entre  $S_{1/2}$  i  $P_{1/2}$  (397 nm): si l'estat esta en  $D_{5/2}$  no passarà res, mentre que si esta en  $S_{1/2}$ , immediatament saltarà  $P_{1/2}$  i com aquest té una vida molt curta, immediatament tornarà a  $S_{1/2}$  emeten llum fluorescent. Es a dir, si l'ió es troba en  $S_{1/2}$  veurem llum i en cas contrari no.

Un cop tenim el qbit per exemple en l'estat fonamental de  $S_{1/2}$ , podrem manipular-lo amb un làser de la freqüència adequada (729 nm) per a que passi a l'estat  $D_{5/2}$  o en una combinació dels dos, dependent del temps d'actuació de la perturbació del laser.

Evidentment, el Hamiltonià que descriu aquesta interacció serà

$$H_C = \hbar \frac{\Omega}{2} (|e\rangle\langle g| + h.c.) \quad (1.227)$$

on C es per indicar "carrier" i distingir-lo d'altres processos que veurem més endavant i  $\Omega \propto \exp(i\omega t)$ , amb  $\omega$  la freqüència de 729nm.

Aquest Hamiltonià induceix operacions unitàries sobre el qbit de la forma

$$R_C(\theta, \phi) = \exp \left[ -i \frac{\theta}{2} (e^{i\phi}|e\rangle\langle g| + e^{-i\phi}|g\rangle\langle e|) \right] \quad (1.228)$$

on  $\theta = \omega t$  és l'angle de rotació d'un pols de durada  $t$ , i  $\phi$  és la fase del làser.

Utilitzant la base habitual  $\{|g\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |e\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , podem expressar  $R_C$  com

$$R_C(\theta, \phi) = \begin{pmatrix} \cos(\theta/2) & -ie^{-i\phi} \sin(\theta/2) \\ -ie^{i\phi} \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \quad (1.229)$$

i per tant,  $R_C(\theta, \phi)$  representa un rotació d'angle  $\theta$  al voltant de l'eix en el pla  $X - Y$  determinat per l'angle  $\phi$ .

La fase làser és arbitrària quan s'aplica per primer cop d'una sèrie de polsos en una transició, però s'ha de fer un seguiment de totes les operacions posteriors fins a la mesura de l'estat final. Per exemple si inicialment l'iò es troba en l'estat  $S_{1/2}$  ( $|\psi(0)\rangle = |g\rangle$ ) i apliquem un làser de 729 nm un temps tal que  $\theta = \omega t = \pi/2$  tindrem (escollim la fase arbitrària com  $\phi_0 = \pi/2$  per conveniència)

$$R_C(\pi/2, \phi_0)|\psi(0)\rangle = R_C(\pi/2, \pi/2)|g\rangle \frac{1}{\sqrt{2}} (|g\rangle + |e\rangle) \quad (1.230)$$

o si  $\theta = \omega t = \pi$  el portarem a  $|e\rangle$ .

Acabem de veure que podem manipular l'estat de qualsevol qbit de la trampa d'ions utilitzant un làser de 729nm que apunti al qbit que volem manipular, actuant durant el temps adequat i amb la fase relativa adequada respecta el primer pols. Però com podrem entrellaçar-los?. Per això, haurem d'utilitzar l'existència d'alguna interacció entre ells: el potencial que els atrapa conjuntament.

Per a una cadena de  $N$  ions, hi ha  $3N$  modes de vibració normals,  $2N$  radials i  $N$  axials. Els dos modes axials de freqüència més baixa són el mode centre de massa, on tots els ions oscil·len com un cos rígid i el mode d'estirament, on l'amplitud d'oscil·lació de cada iò és proporcional a la seva distància al centre (figura 1.58). Com s'explicarà a continuació, l'entrellaçament entre diferents ions s'aconsegueix mitjançant una excitació làser coherent de transicions entre els estats quàntics més baixos  $|n = 0\rangle$  i  $|n = 1\rangle$  d'un dels modes de vibració axial, generalment el centre de massa (veure figura 1.59). Es a dir, aquests dos estats, el fonamental i el primer nivell excitat del mode centre de massa no es utilitzaran com a qbits d'informació, sinó com a "bus" per entrellaçar dos ions de la trampa. Per aquesta raó aquests dos estats se'ls anomena qbits de moviment o qbits bus.

Refredant adequadament la trampa d'ions, la col·lectivitat d'ions es porta a l'estat fonamental vibracional del centre de masses. Aleshores, els dos nivells més baixos del mode vibracional, que designarem per  $|0\rangle$  i  $|1\rangle$ , formen el que hem anomenat qbit de moviment o qbit de bus (noteu la desafortunada notació, doncs als veritable qbits que utilitzarem per emmagatzemar l'informació quàntica els hem anomenat com  $\{|g\rangle, |e\rangle\}$ , mentre que aquests que només els utilitzarem com a "bus" els anomenem  $\{|0\rangle, |1\rangle\}$ ). La diferència d'energia dels qbits de bus és més petita que la diferència d'energia dels qbits d'informació. Aleshores, la base per a operacions lògiques quàntiques, o "subespai computacional" (CS) de qualsevol iò de la trampa està format pels quatre estats  $\{|g, 0\rangle, |g, 1\rangle, |e, 0\rangle, |e, 1\rangle\}$  (figura 1.60), on per exemple  $|g, 0\rangle$ , indica que l'iò està en l'estat fonamental  $|g\rangle$  com a qbit de informació, i que la col·lectivitat d'ions està en l'estat fonamental de vibració.

Per a una manipulació de l'estat quàntic de l'iò dins del CS, podem utilitzar làsers de la freqüència adequada, tal com es mostra en la figura 1.61. Observem que les transicions vermella, portadora (C) o blava, al ser de diferents freqüències ( $\omega_- = \omega - \Delta\omega, \omega, \omega_+ = \omega + \Delta\omega$ ) podem escollir quina transició volem aplicar.

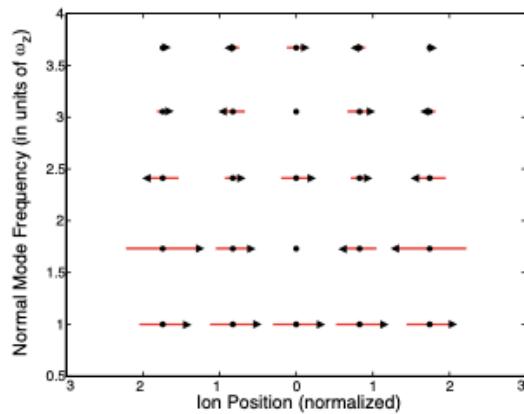


Figure 1.58: Modes normals de vibració axial: freqüències i amplituds per a 5 ions. L'inferior és el mode centre de massa i el immediatament superior el mode d'estirament.

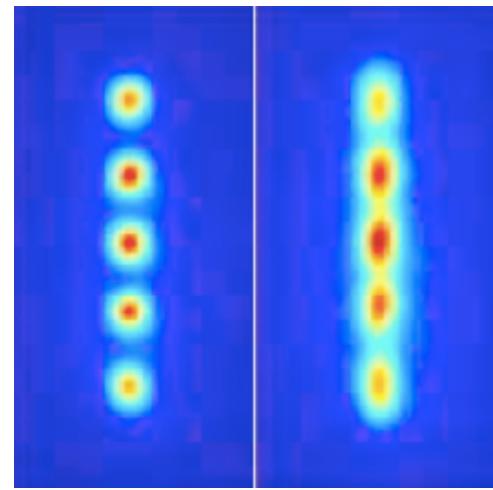


Figure 1.59: Mode centre de masses: nivell fonamental i nivell excitat

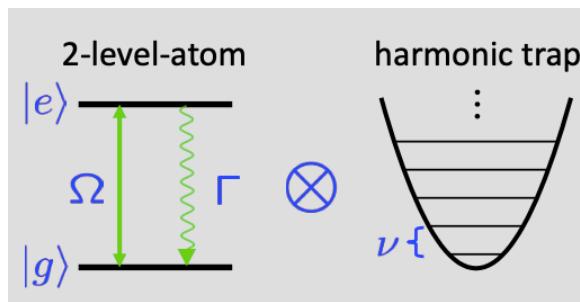


Figure 1.60: El qbit i el potencial harmònic que el manté en la trampa. Ambdós defineixen el "subespai computacional" (CS).

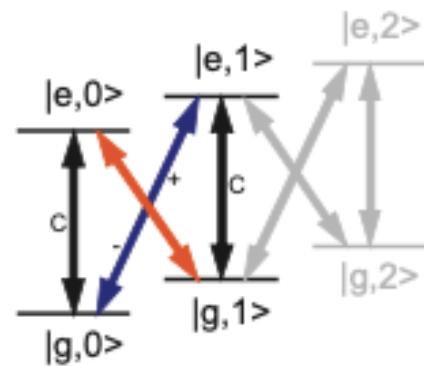


Figure 1.61: nivells enegètics del qbit considerant també els efectes de translació

Els Hamiltonians que descriuen aquestes transicions seran

$$\begin{aligned} H_- &= \hbar \frac{\eta_- \Omega}{2} (|e\rangle\langle g|a + h.c.) \\ H_+ &= \hbar \frac{\eta_+ \Omega}{2} (|e\rangle\langle g|a^\dagger + h.c.) \end{aligned} \quad (1.231)$$

on hem posat  $\eta_\pm$  per indicar la lleugera diferència de la freqüència d'aquestes transicions respecta a la  $\Omega$  de "carrier"

Aquests Hamiltonians indueixen operacions unitàries sobre el qbit de la forma

$$\begin{aligned} R_-(\theta, \phi) &= \exp \left[ -i \frac{\theta}{2} \left( e^{i\phi} |e\rangle\langle g|a + e^{-i\phi} |g\rangle\langle e|a^\dagger \right) \right] \\ R_+(\theta, \phi) &= \exp \left[ -i \frac{\theta}{2} \left( e^{i\phi} |e\rangle\langle g|a^\dagger + e^{-i\phi} |g\rangle\langle e|a \right) \right] \end{aligned} \quad (1.232)$$

on  $\theta = \omega_\pm t$  és l'angle de rotació d'un pols de durada  $t$ , i  $\phi$  és la fase del làser. Per exemple, si inicialment l'iò es troba en l'estat  $S_{1/2}$  ( $|\psi(0)\rangle = |g0\rangle$ ), és a dir, l'iò es troba en l'estat fonamental  $|g\rangle$  i la col·lectivitat en l'estat fonamental de vibració  $|0\rangle$ , i apliquem el làser lateral "blau" un temps tal que  $\theta = \omega_+ t = \pi/2$  tindrem

$$R_+(\pi/2, \phi_0) |\psi(0)\rangle = R_+(\pi/2, \pi/2) |g0\rangle \frac{1}{\sqrt{2}} (|g0\rangle + |e1\rangle) \quad (1.233)$$

i per tant introduceix entrallaçament entre el qbit atòmic i el qbit vibracional del mateix iò (l'estat és una superposició de l'iò en estat fonamental amb tots els ions en l'estat vibracional fonamental i l'iò en estat excitat amb tots els ions en l'estat vibracional excitat). L'excitació de banda lateral d'un iò modifica l'estat de moviment de tota la cadena, proporcionant així l'acobllament iò-iò necessari.

La generalització del subespai computacional a partir del cas d'un iò és directa: el CS de dos ions (que poden ser dos qualsevol en una cadena molt llarga) és l'espai producte. Els estats i les transicions resultants es mostren a la figura 1.62. Els impulsos es defineixen de manera anàloga a  $R_\pm(\theta, \phi)$  però amb superíndex addicionals (1) o (2) que indica l'iò adreçat. Per exemple

$$|gg0\rangle \xrightarrow{R_+^1(\pi/2, \phi_0)} \frac{1}{\sqrt{2}} (|gg0\rangle + |eg1\rangle) \xrightarrow{R_-^2(\pi, \phi_0)} \frac{1}{\sqrt{2}} (|gg0\rangle + |ee0\rangle) \quad (1.234)$$

ens porta d'estat no entrellaçat  $|gg0\rangle$  de dos ions a l'estat entrellaçat  $1/\sqrt{2}(|g,g\rangle + |e,e\rangle)$  deixant el qbit de moviment intacte. Ho hem aconseguit utilitzant dos únics polsos làser, el primer actuant en el primer iò i el segon sobre el segon iò. Hem utilitzat el qbit moviment com a "bus" per entrellaçar els dos qbits atòmics.

## 1.12 Perspectives de la Computació Quàntica

La implementació de la criptografia quàntica utilitzant llum polaritzada ja és un fet. Per altra banda, ja s'estan buscant alternatives clàssiques als protocols com el RSA que no puguin ser desxifrats ni en computadores clàssiques ni quàntiques. Tot apunta que si en algun moment hi han ordinadors quàntics que puguin implementar l'algorisme de Shor per a gran nombres, només servirà per desxifrar missatges del passat.

El problemes principals de la Computació Quàntica actualment són

- L'escalabilitat de la tecnologia

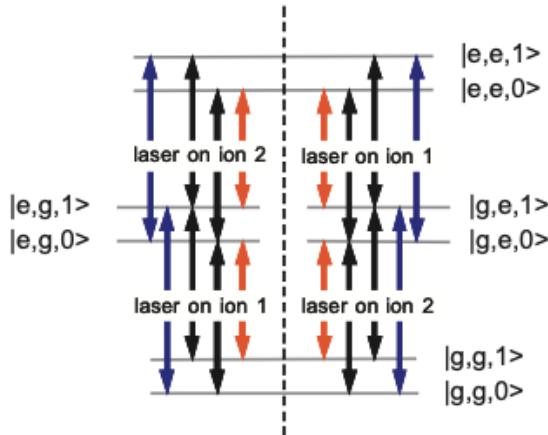


Figure 1.62: Subespai computacional de dos ions i un mode vibracional, amb totes les transicions possibles. Les fletxes curtes, mitjanes i llargues són transicions de banda lateral vermella, portadora i blava, respectivament.

- El soroll. Els qbits actuals són sorollosos (perden la coherència). El tipus més comú d'ordinador quàntic universal utilitza qbits fets a partir d'unions superconductores. Totes les màquines d'aquest tipus fins ara necessiten 1.000 qbits bruts per obtenir un qbit net. Un ordinador quàntic universal útil necessitaria uns 50 qbits nets, que requeririen 50.000 qbits bruts.
- La comunicació/interfície entre la informació clàssica i la quàntica. No sabem carregar dades clàssiques directament a una memòria quàntica com tampoc podem extreure tota la informació quàntica que hi ha en un ordinador quàntic (teorema de no teleportació de la informació quàntica).

Veiem un exemple del problema de carregar dades. Un Petabit són  $10^{15}$  bits. Una base de dades d'un Petabit permet  $2^{10^{15}}$  configuracions diferents i si totes són equiprobables, l'informació clàssica (entropia de Shannon) que podem guardar en aquesta base de dades és evidentment de  $10^{15}$  bits. Com 1 qbit d'informació quàntica és equivalent a infinitis bits de informació clàssica, en un sol qbit podríem guardar qualsevol configuració de la base de dades, per exemple transformant-la en una direcció biunívoca del qbit. Una alternativa seria utilitzar 50 qbits en lloc d'un. En aquest cas, com que la dimensió de l'espai de Hilbert és de  $2^{50} \approx 10^{15}$ , podríem guardar qualsevol configuració de la base de dades en la combinació dels elements de la base de l'espai de Hilbert de coeficients precisament els donats per la seqüència de 0s i 1s de la configuració. El problema d'aquestes solucions és que no sabem com construir un codificador que faci aquesta feina.

L'altre problema és com extreure la informació. Si tenim un ordinador quàntic amb 50 qbits i fem una mesura al final del càcul, només obtindrem 50 bits de informació clàssica i per tant, haurem de repetir moltes vegades el càcul i fer la posterior mesura (que a més a més és aleatòria) fins extreure tota la informació rellevant que ens interessa (per exemple, els pesos d'una enorme xarxa neuronal).

Podem concloure que:

- Hi ha moltes expectatives en la computació quàntica per resoldre problemes.
- Per alguns problemes, no els resoldrà durant molt de temps.

- Per als problemes d'anàlisi de Big Data, que constitueixen la major part de la actual inversió empresarial en informàtica, és difícil veure com ho podria fer.
- .... però jugar amb estats quàntics de molts graus de llibertat, a part de ser divertit, pot portar-nos sorpreses en apropar-nos a la frontera entre el món quàntic i el món clàssic.