Eötvös Loránd Tudományegyetem
Matematika Intézet

Seres István András

A kvantum-információelmélet alapjai

BSc szakdolgozat

Témavezető: dr. Frenkel Péter

ELTE Algebra és Számelmélet Tanszék

2014. Budapest

Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani témavezetőmnek, Frenkel Péternek, akinek segítségével betekintést nyertem a kvantum-információlmélet alapjaiba. Köszönöm rendszeres útmutatásait, tanácsait, megértést segítő példáit. Köszönöm, hogy mindig bátran fordulhattam hozzá a szakdolgozatot érintő legapróbb kérdésekkel is. Az ő irányítása és segítsége nélkül ez a szakdolgozat aligha jöhetett volna létre.

Tartalomjegyzék

Ábrák jegyzéke

Jelölések

Néhány jelölés és konvenció általánosan elfogadott mind a klasszikus, mind a kvantum-információelméletben. Ezeket és a szakdolgozatban használt néhány kvantummechanikai jelölést (Dirac-jelölések) tekintünk át az alábbiakban.

log: a továbbiakban mindig kettes alapú logaritmust értünk rajta

ln: természetes alapú logaritmus

z^*: a z komplex szám konjugáltja

A^T: mátrix transzponáltja

A^*: mátrix adjungáltja

$|\psi\rangle$: oszlopvektor, gyakran *ket*-nek nevezzük

$\langle\psi|$: vektor duálisa, sorvektor, gyakran *bra*-nak nevezzük

$\langle\psi|\varphi\rangle$: a $|\psi\rangle$ és a $|\varphi\rangle$ vektorok skalárszorzata

$|\psi\rangle \otimes |\varphi\rangle$: a $|\psi\rangle$ és a $|\varphi\rangle$ vektorok tenzorszorzata, amit gyakran csak $|\psi\rangle|\varphi\rangle$-vel jelölünk

$\langle\psi|A|\varphi\rangle$: a $|\psi\rangle$ és az $A|\varphi\rangle$ vektorok skalárszorzata

1. fejezet

Bevezetés

A kvantum-információelmélet igencsak fiatal ága a matematikának. Ellentétben a klasszikus információelmélettel, amelynek pontos kezdete jól behatárolható — Shannon 1948-as A Mathematical Theory of Communication c. cikkének megjelenéséhez kötik—, addig a kvantum-információelméletnek nem határozható meg ilyen pontosan a kezdete. Megalapozói (Neumann János, Elliott H. Lieb, Mary Beth Ruskai és mások) az 50-es illetve 60-as évektől kezdve publikálták főbb eredményeiket.

Nem csak a téma újdonsága adja a szakdolgozat aktualitását, érdekességét. Matematikai szempontból számos nemtriviális lineáris algebrai illetve információelméleti kérdést vet fel, míg fizikai szempontból a mai napig megválaszolatlan, hogy építhető-e kvantumszámítógép. Biztató jelek vannak ugyan, de korántsem kielégítőek a megépített modellek.

A következőkben a 2. fejezetben a klasszikus információelmélet fontosabb fogalmai és tételei vannak bemutatva, majd ezeknek a kvantum-információelméleti vonatkozásai vannak ismertetve a 4. fejezetben. A 3. fejezetben a 4. és 5. fejezethez nélkülözhetetlen kvantummechanikai alapok és egyéb fontos, később sokszor használt tételek kerülnek bemutatásra. A szakdolgozat legfontosabb része — a kölcsönös kvantum-információra vonatkozó Holevo-korlát bizonyítása és annak következményei — a 5. fejezetben található.

A szakdolgozatomban legfőképp a [1] és a [2] könyvek felépítését, tárgyalását követtem. Hasznomra volt még a [3] könyv és a [4] interneten elérhető információelméleti jegyzet is.

2. fejezet

Klasszikus információelméleti mennyiségek és tulajdonságaik

Ebben a fejezetben az entrópia, a feltételes illetve kölcsönös információ és ezek fontosabb tulajdonságai kerülnek bemutatásra. A továbbiakban kizárólag véges halmazokon értelmezett diszkrét valószínűségi mértékekkel foglalkozunk.

2.1. Shannon-entrópia

Legyen adott egy X diszkrét valószínűségi változó. Ekkor szeretnénk definiálni az entrópiát, azt a várható információmennyiséget, amit akkor nyerünk, amikor X értékét megtudjuk. Jelölje $H(X)$ ezt a függvényt.

Mielőtt egy valószínűségi változó információmennyiségét definiálnánk, definiáljuk egy esemény információmennyiségét. Egy ilyen függvényt jelöljünk I-vel. Nyilván ennek a függvénynek teljesítenie kell a következő kritériumokat:

(1) $I(E)$ csak az E esemény valószínűségének függvénye, így írhatjuk, hogy $I = I(p)$, ahol p az E esemény valószínűsége $(0 \leq p \leq 1)$.

(2) I nemnegatív, sima függvény.

(3) Ha p és q két független esemény valószínűsége, akkor $I(p) + I(q) = I(pq)$, vagyis két független esemény külön-külön annyi információt szolgáltat, mint együttes bekövetkezésük.

(4) Megállapodás szerint az egységet válasszuk meg az alábbi módon: $I(1/2) = 1$.

2.1.1. Tétel (Egyedi információ)**.** $I(p) = -\log(p)$.

Bizonyítás. Legyen $\varphi(x) = I(2^{-x})$, ahol $0 \leq x$. A bevezetett új jelölés mellett teljesülnek az alábbiak: $\varphi(x+y) = \varphi(x)+\varphi(y)$ és $\varphi(1) = 1$. Ennek az ún. Cauchy-féle függvényegyenletnek pedig tudjuk, hogy egyedül a $\varphi(x) = x$ az egyetlen folytonos megoldása. □

A 2.1.1. tétel motiválja, hogy a következő módon definiáljuk egy X valószínűségi változó információmennyiségét, másszóval entrópiáját.

2.1.2. Definíció. Egy X diszkrét valószínűségi változó *Shannon entrópiáján az alábbi mennyiséget értjük:*

$$H(X) = H(p_1, ..., p_n) = -\sum_{i=1}^{n} p_i \log p_i$$

.

Megállapodás szerint legyen $0 \log 0 = 0$. Ezt alátámasztja az az intuició is, hogy egy 0 valószínűségű esemény nem hordoz információt illetve az a tény is, hogy az $x \log x$ függvény a 0-ban folytonos és ott a határértéke 0.

Egy diszkrét valószínűségi változó Shannon entrópiájára gondolhatunk úgy is, mint a valószínűségi változó által meghatározott elemi események egyedi információinak várható értékére.

2.2. A relatív entrópia

A következő mennyiséggel valamilyen értelemben valószínűségi változók távolságát tudjuk mérni.

2.2.1. Definíció. Legyenek $p(x)$ és $q(x)$ valószínűségeloszlások ugyanazon $x \in X$ indexhalmazon értelmezettek. Ekkor $p(x)$ *relatív entrópiája* $q(x)$-re nézve a következő mennyiség:

$$H(p(x)||q(x)) = \sum_x p(x) \log \frac{p(x)}{q(x)} = -H(X) - \sum_x p(x) \log q(x).$$

Megállapodás szerint legyen $-0 \log 0 = 0$ illetve $-p(x)\log 0 = +\infty$ ha $p(x) > 0$. A következő tétel megmutatja miért is használható a relatív entrópia valószínűségi változók távolságának mérésére.

2.2.2. Tétel (A relatív entrópia nemnegativitása)**.** A relatív entrópia nemnegatív, azaz $H(p(x)||q(x)) \geq 0$, és egyenlőség akkor és csak akkor teljesül, ha $p(x) = q(x)$ minden x esetén.

Bizonyítás. Alkalmazzuk az alábbi, elemi analízisből ismert egyenlőtlenséget: $\log x \ln 2 = \ln x \leq x - 1$, ahol $x > 0$, és egyenlőség akkor és csak akkor teljesül, ha $x = 1$.

Ezt átrendezve kapjuk: $-\log x \geq \frac{1-x}{\ln 2}$. Ezzel becsüljük a relatív entrópiát alulról:

$$H(p(x)||q(x)) = -\sum_x p(x) \log \frac{q(x)}{p(x)} \tag{2.1}$$

$$\geq \frac{1}{\ln 2} \sum_x p(x) \left(1 - \frac{q(x)}{p(x)}\right) \tag{2.2}$$

$$= \frac{1}{\ln 2} \sum_x (p(x) - q(x)) \tag{2.3}$$

$$= \frac{1}{\ln 2}(1 - 1) = 0, \tag{2.4}$$

és ez az, amit bizonyítani akartunk. Egyenlőség pedig akkor és csak akkor áll fenn, ha (2.2)-nél egyenlőség van, azaz $\frac{q(x)}{p(x)} = 1$ minden x esetén, vagyis ha a két valószínűségi változó eloszlása megegyezik. □

A relatív entrópia bizonyos értelemben a valószínűségi változók között egy távolságot definiál, de nem metrika. Könnyen látható, hogy már a szimmetrikusság sem teljesül. Az információelméletben be szokták vezetni a relatív entrópia szimmetrizált változatát, az információs divergenciát, de még az sem teljesíti a háromszög-egyenlőtlenséget. A továbbiakban látni fogjuk, hogy milyen hasznos fogalom a relatív entrópia. A relatív entrópia első alkalmazásaként felső korlátot adunk az entrópiára.

2.2.3. Tétel (Felső korlát az entrópiára)**.** Legyen X egy valószínűségi változó d darab kimenettel. Ekkor $H(X) \leq \log d$, egyenlőség akkor és csak akkor teljesül, ha X egyenletes eloszlású a d kimeneten.

Bizonyítás. Legyen $p(x)$ tetszőleges valószínűségi eloszlás, míg $q(x)$ legyen egyenletes eloszlású a d kimeneten.

$$H(p(x)||q(x)) = H(p(x)||\frac{1}{d}) = -H(X) - \sum_x p(x) \log \frac{1}{d} = \log d - H(X). \tag{5.1}$$

A relatív entrópia nemnegativitása miatt $\log d - H(X) \geq 0$. Ebből átrendezéssel adódik az állítás, illetve egyenlőség akkor és csak akkor teljesül, ha $p(x)$ is egyenletes eloszlású. □

2.3. Kölcsönös entrópia, feltételes entrópia és kölcsönös információ

Legyen X és Y két valószínűségi változó. Arra vagyunk kíváncsiak, hogy milyen kapcsolat van X és Y információmennyisége között. Ennek megválaszolása céljából vezetjük be a következő fogalmakat.

2.3.1. Definíció. X és Y kölcsönös entrópiájának nevezzük a következő mennyiséget:

$$H(X, Y) = -\sum_{x,y} p(x, y) \log p(x, y).$$

X feltételes entrópiáját Y -ra nézve így értelmezzük: $H(X|Y) = H(X, Y) - H(Y)$.

X és Y kölcsönös információján pedig a $H(X : Y) = H(X) + H(Y) - H(X, Y)$ mennyiséget értjük.

Szemléletesen úgy tekinthetünk a kölcsönös entrópiára, mint arra az információmennyiségre, amelyet az (X, Y) pár átlagosan hordoz. A feltételes entrópia azt az átlagos információmennyiséget, bizonytalanságot méri, ami akkor lép fel, amikor ismerjük Y értékét, de X-ét még nem. A kölcsönös információ pedig azt az információmennyiséget méri, hogy az egyik változó entrópiája várhatóan mennyivel csökken, ha a másiknak megtudjuk az értékét.

A definíciókból azonnal jön az alábbi összefüggés: $H(X : Y) = H(X) - H(X|Y)$.

2.3.2. Tétel (A bevezetett mérőszámok alaptulajdonságai)**.** *(1)* $H(X, Y) = H(Y, X)$, $H(X : Y) = H(Y : X)$

(2) $H(Y|X) \geq 0$, *és így* $H(X : Y) \leq H(Y)$, *egyenlőség pontosan akkor, ha* Y függvénye X-nek, $Y = f(X)$.

(3) $H(X) \leq H(X, Y)$, egyenlőség pontosan akkor, ha Y függvénye X-nek.

(4) **Szubadditivitás:** $H(X, Y) \leq H(X) + H(Y)$, egyenlőség akkor és csak akkor, ha X és Y független valószínűségi változók.

(5) $H(Y|X) \leq H(Y)$, tehát $H(Y : X) \geq 0$, egyenlőség akkor és csak akkor, ha X és Y függetlenek.

(6) **Erős szubadditivitás:** $H(X, Y, Z) + H(Y) \leq H(X, Y) + H(Y, Z)$ és egyenlőség pontosan akkor, ha $Z \to Y \to X$ egy Markov-lánc.

(7) A feltételek csökkentik az entrópiát: $H(X|Y, Z) \leq H(X|Y)$.

Bizonyítás. (1) Világos a definíciókból.

(2) Mivel $p(x, y) = p(x)p(y|x)$, ezért

$$H(X, Y) = -\sum_{x,y} p(x, y) \log p(x)p(y|x)$$

$$= -\sum_{x} p(x) \log p(x) - \sum_{x,y} p(x, y) \log p(y|x)$$

$$= H(X) - \sum_{x,y} p(x, y) \log p(y|x)$$

Ezért $H(Y, X) = -\sum_{x,y} p(x, y) \log p(y|x)$. De $-\log p(y|x) \leq 0$, tehát $H(Y, X) \geq 0$. Láthatóan egyenlőség pontosan akkor teljesül, ha Y X-nek függvénye.

(3) Következik az előző pontból.

(4) Ismét alkalmazzuk az analízisből ismert egyenlőtlenséget: $\log x \leq \frac{x-1}{\ln 2}$, ahol $x > 0$, és egyenlőség akkor és csak akkor teljesül, ha $x = 1$. Így kapjuk, hogy:

$$\sum_{x,y} p(x, y) \log \frac{p(x)p(y)}{p(x, y)} \leq \frac{1}{\ln 2} \sum_{x,y} p(x, y)(\frac{p(x)p(y)}{p(x, y)} - 1)$$

$$= \frac{1}{\ln 2} \sum_{x,y} p(x)p(y) - p(x, y) = \frac{1-1}{\ln 2} = 0$$

Ez pedig bizonyítja az állítást. Egyenlőség pedig pontosan akkor áll fenn, ha $p(x, y) = p(x)p(y)$ minden x és y esetén. Vagyis az egyenlőtlenség akkor éles, ha Xés Yfüggetlen.

(5) A szubadditivitásból és a definíciókból következik.

(6) A bizonyítás teljesen hasonlóan megy, mint a szubadditivitásnál, csak itt hármas szummák szerepelnek.

(7) $H(X, Y, Z) - H(Y, Z) \leq H(X, Y) - H(X)$ egyenlőtlenség ekvivalens az állítással, ami éppen az erős szubadditivitás átrendezett alakja. □

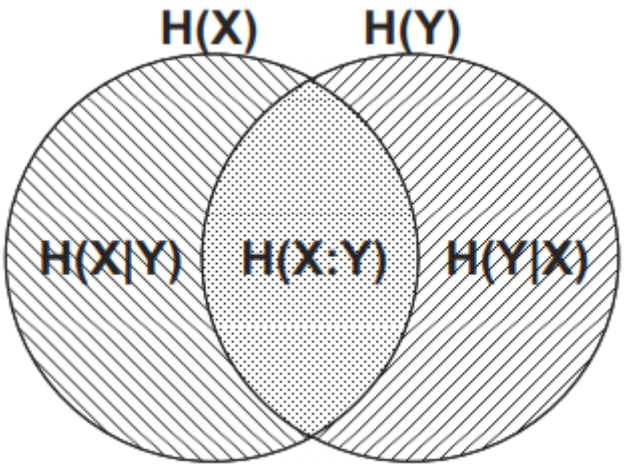
Habár az entrópia szubadditív, a kölcsönös információ nem mindig az, vagyis nem mindig teljesül a következő egyenlőtlenség:

$$H(X, Y : Z) \leq H(X : Z) + H(Y : Z).$$

Valóban, legyenek X és Y független azonos eloszlású valószínűségi változók, melyek 1/2-1/2 valószínűséggel veszik fel a 0-t és az 1-et. Z-t pedig válasszuk eképp: $Z = X \oplus Y$, ahol $\oplus$ a modulo 2 szerinti

összeadást jelöli. Egyszerű számolás mutatja, hogy nem teljesül az egyenlőtlenség, tehát a kölcsönös információ nem mindig szubadditív.

2.1. ábra. Entrópia Venn-diagramm



A különböző klasszikus információelméleti mérőszámok közötti kapcsolatokat, összefüggéseket leginkább egy "entrópia Venn-diagrammon" lehet bemutatni. Ez az ábra inkább az egyes definíciókat szemlélteti jól, mintsem a különböző entrópiafüggvények tulajdonságait mutatja be.

3. fejezet

Kvantummechanikai alapok és egyéb fontos tételek

Mai tudásunk szerint a legpontosabb leírását a világnak a kvantummechanika adja. Ebbe a világba enged bepillantani a következő fejezet, de csak annyira, amennyi szükséges a következő fejezetek (kvantum információelmélet és Holevo-korlát) szempontjából.

3.1. A kvantummechanika posztulátumai

A következő posztulátumok lehetővé teszik a fizikai világ matematikai formalizációját a kvantummechanikai elméleten belül. Az első posztulátum felállítja a kvantummechanika színterét, ahol az egész történet játszódik.

3.1.1. Posztulátum (A kvantummechanika első posztulátuma)**.** Minden fizikai rendszerhez hozzárendelünk egy komplex vektorteret és egy skaláris szorzást ezen a vektortéren, vagyis egy Hilbert-teret. Ezt a Hilbert-teret a rendszerhez tartozó *állapottérnek* nevezzük. Egy fizikai rendszer állapotát az *állapotvektor* írja le, amely egy egységhosszú vektor az állapottérben.

Hogyan változik egy kvantum rendszer $|\psi\rangle$ állapota?

3.1.2. Posztulátum (A kvantummechanika második posztulátuma)**.** Egy *zárt* kvantumrendszer "viselkedése" egy unitér transzformációval írható le. Vagyis, ha a kvantumrendszer a t_1 időpillanatban a $|\psi\rangle$ állapotban van és a t_2 időpillanatban egy $|\psi'\rangle$ állapotba kerül, akkor van olyan U unitér transzformáció mely csak t_1 és t_2-től függ és amelyre teljesül

$$|\psi'\rangle = U|\psi\rangle.$$

Azonnal felmerül a kérdés, hogy mely unitér transzformációkat természetes tekinteni? Könnyen meg lehet mutatni, hogy az összes unitér transzformáció előfordulhat valóságos kvantumállapotok leírásánál.

A valóságban csak a legritkább esetekben találkozunk zárt fizikai rendszerekkel. Hogyan írhatóak le a nyílt kvantumrendszerek viselkedése? Mi történik, ha azon mérést végzünk?

3.1.3. Posztulátum (A kvantummechanika harmadik posztulátuma)**.** A kvantum méréseket egy $\{M_m\}$ halmazzal jelöljük, amelyeket röviden csak *mérésnek* nevezünk. Ezek az operátorok a megmérendő rendszer állapotterén hatnak. Az alsó indexben az m a végeredményre utal. Ha a rendszer a mérés előtt $|\psi\rangle$ állapotban volt akkor az a valószínűség, hogy végeredményül m-et kapunk:

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

alakban írható. A mérés után a rendszer állapotvektora:

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^* M_m|\psi\rangle}}$$

Továbbá a mérések halmaza kielégíti a teljességi relációt, vagyis:

$$\sum_m M_m^* M_m = I.$$

A teljességi reláció lényegében csak annyit fejez ki, hogy a valószínűségek összege 1:

$$1 = \sum_m p(m) = \sum_m \langle\psi|M^* M|\psi\rangle$$

A továbbiakban fontos szerep jut az összetett kvantumrendszereknek. Hogyan néz ki az összetett kvantumrendszerek állapottere? Erre ad választ a következő,

3.1.4. Posztulátum (A kvantummechanika negyedik posztulátuma)**.** Az összetett kvantumrendszer állapottere a komponensek állapottereinek tenzorszorzata. Vagyis ha 1-től n-ig indexeljük kvantum rendszerek egy halmazát és az i-edik a ψ_i állapotban van, akkor az összetett rendszer a $\psi_1 \otimes \psi_2 ... \otimes \psi_n$ állapotban van.

3.2. Qubitok

Tekintsük a legegyszerűbb esetet, vagyis amikor az állapottér egy 2-dimenziós komplex Hilbert-tér. Ebben az esetben az állapotvektorokat *qubitoknak* nevezzük. A qubitoknak a következő előállítását használjuk:

$$|\psi\rangle = a|0\rangle + b|1\rangle,$$

ahol $|0\rangle$ és $|1\rangle$ egy ortonormált bázisa az állapottérnek és a,b komplex számok. A feltétel, hogy $|\psi\rangle$ egységhosszú vektor azt jelenti, hogy $\langle\psi|\psi\rangle = 1$, vagy ekvivalens megfogalmazásban: $|a|^2 + |b|^2 = 1$. A qubit a legalapvetőbb kvantummechanikai rendszer, akárcsak a bit a klasszikus információelméletben. Lényeges különbség, hogy míg egy bit két állapot (0 és 1) valamelyikében lehet, addig egy qubit ezek *szuperpozíciójában* is lehet, vagyis az $a|0\rangle + b|1\rangle$ állapotban. Ez a szuperpozíció azt jelenti, hogy a qubitról nem tudjuk megmondani teljes biztonsággal, hogy a 0 vagy az 1 állapotban van, hanem *valahol a kettő között.*

Az a és b együtthatókat *valószínűségi amplitúdóknak* is szokták nevezni. Az elnevezésnek az az oka, hogyha meg akarjuk mérni, hogy a qubit melyik állapotban van, akkor $|a|^2$ valószínűséggel azt kapjuk, hogy a 0 állapotban van és $|b|^2$ valószínűséggel azt kapjuk, hogy az 1 állapotban van.

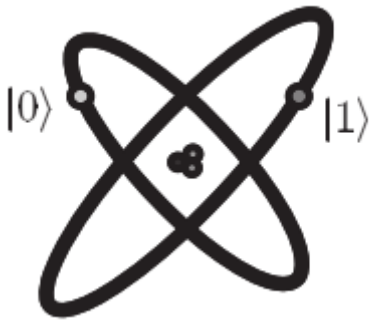
Elsőre a qubit fogalma nem tűnik túl életszerűnek, de a természetben mégis számtalan qubit-realizáció létezik. Csak, hogy említsünk néhányat:

(1) egy foton két polarizációs állapota,

(2) egy elektron két különböző spinje egy mágneses térben,

(3) egy atom körül keringő elektron két állapotban lehet: a 0 az alapállapotnak felel meg, míg az 1 a gerjesztett állapotnak. Ezen állapotok között "ingázhat" az elektron. Számtalan módszerrel (pl.: fénysugárral) gerjeszthetjük az atomot, melynek hatására a 0 állapotból az 1-esbe kerül. Változtatva a sugárzás, gerjesztés mértékét az elektron a két állapot "közé" is kerülhet.

3.1. ábra. Egy qubit realizációja az elektron két állapotának segítségével



Mennyi információt tudunk tárolni egy qubitban? Mint korábban láttuk végtelen sok (kontinuum sok) állapotban lehet egy qubit, de ha mérést végzünk rajta, akkor csak a 0 vagy 1 állapotok valamelyikében lehet a valószínűségi amplitúdóknak megfelelő valószínűségekkel. Tehát egy qubit végtelen sok információt tud tárolni, de a belőle kinyerhető információ ugyanannyi, mint egy klasszikus bit információja.

3.2.1. Többszörös qubitok

Tegyük fel, hogy van két qubitunk. Ha ez két klasszikus bit lenne, akkor 4 állapotunk lenne: 00,01,10,11. Ennek megfelelően a két qubit rendszer egy 4-dimenziós komplex Hilbert-térben él, ahol $|00\rangle$,$|01\rangle$,$|10\rangle$ és $|11\rangle$ egy ortonormált bázis. Ez a két qubit is lehet ezen állapotok szuperpozíciójában, így a két qubit rendszerben az állapotvektor az alábbi alakban írható:

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

Hasonlóan a qubit esethez itt is egy mérés után annak a valószínűsége, hogy az (x = 00,01,10 vagy 11) állapotot mérjük $|a_x|^2$ a valószínűsége. A mérés után a qubit az $|x\rangle$ állapotba kerül. Mivel az

együtthatók itt is valószínűségi amplitúdók, ezért

$$\sum_{x\in\{0,1\}^2} |a_x|^2 = 1.$$

A két qubit esetéhez teljesen hasonlóan értelmezzük az n qubit esetet. Tehát n qubitot 2^n darab valószínűségi amplitúdóval írhatunk le.

3.3. Kvantumállapotok megkülönböztetése

A klasszikus információelméletben mindig meg tudunk különböztetni két állapotot. Semmilyen elvi akadályba nem ütközünk, amikor meg szeretnénk állapítani, hogy egy pénzérme fejre vagy írásra érkezett-e. A kvantummechanikában viszont ez már nincs így.

A problémát a legjobban az alábbi példával lehet szemléltetni: tekintsünk két játékost, Aladárt és Bélát, akik egy nem mindennapi "barkochba"-t játszanak. Kvantumállapotok egy előre meghatározott $|\psi_i\rangle$ $(1 \leq i \leq n)$ halmazából választ egyet Aladár. Bélának az a feladata, hogy meghatározza azt az i indexet, amelyhez az Aladár által választott $|\psi_i\rangle$ kvantumállapot tartozik.

Először tekintsük azt az esetet, amikor a $|\psi_i\rangle$ állapotok ortogonálisak. Ekkor Béla tud olyan méréseket végezni az állapotvektoron, hogy ki tudja deríteni teljes biztonsággal az i index kilétét. Az eljárás a következő: minden i-re definiáljuk az $M_i = \sqrt{|\psi_i\rangle\langle\psi_i|}$ méréseket, továbbá legyen M_0 a négyzetgyöke az $I - \sum_{i\neq 0} |\psi_i\rangle\langle\psi_i|$ operátornak. Az így definiált operátorok kielégítik a teljességi relációt (vagyis négyzetösszegük az identitás operátor) és ha Aladár a $|\psi_i\rangle$ állapotot készíti el Bélának, akkor $p(i) = \langle\psi_i|M_i|\psi_i\rangle = 1$. Tehát ortogonális állapotvektorok esetén teljes biztonsággal meg tudja Béla mondani, hogy Aladár melyik indexű állapotvektorra gondolt.

Ezzel szemben, ha a $|\psi\rangle$ állapotvektorok nem ortogonálisak, akkor bebizonyítjuk, hogy nem létezik kvantum mérés, mellyel meg tudnánk állapítani teljes biztonsággal az állapotvektor indexét.

3.3.1. Tétel (Nem-ortogonális állapotok megkülönböztethetetlensége)**.** *Tegyük fel, hogy Béla $\{M_j\}$ mérést végez, melynek j az eredménye. Az eredmények alapján egy f függvény segítségével Béla megteszi $i = f(j)$ tippjét a kvantumállapot indexére vonatkozóan. Ekkor a $|\psi_1\rangle$ és $|\psi_2\rangle$ nem-ortogonális állapotok nem megkülönböztethetőek.*

Bizonyítás. Indirekt módon tegyük fel, hogy léteznek olyan mérések, mellyel $|\psi_1\rangle$ és $|\psi_2\rangle$ nem-ortogonális állapotok megkülönböztethetőek. Tehát ha $|\psi_1\rangle$ vagy $|\psi_2\rangle$ állapotokat kapja Béla, akkor 1 annak a valószínűsége, hogy a mérés eredménye j, ahol $f(j) = 1$ vagy $f(j) = 2$. Legyen $E_i = \sum_{j:f(j)=i} M_j^* M_j$, ahol a feltevések szerint

$$\langle\psi_1|E_1|\psi_1\rangle = 1; \langle\psi_2|E_2|\psi_2\rangle = 1. \tag{3.3.1}$$

Mivel $\sum_i E_i = I$, ezért $\sum_i \langle\psi_1|E_i|\psi_1\rangle = 1$ és $\langle\psi_1|E_1|\psi_1\rangle = 1$ miatt $\langle\psi_1|E_2|\psi_1\rangle = 0$, vagyis $\sqrt{E_2}|\psi_1\rangle = 0$. Az állapotok non-ortogonalitásából adódik a következő felbontás: $|\psi_2\rangle = \alpha|\psi_1\rangle + \beta|\varphi\rangle$, ahol $|\varphi\rangle$ ortogonális $|\psi_1\rangle$-re. Mivel az állapotvektorok egységhosszúak, ezért $|\alpha|^2 + |\beta|^2 = 1$, és $|\beta| < 1$, mert

$|\psi_1\rangle$ és $|\psi_2\rangle$ nem ortogonálisak. Vagyis $\sqrt{E_2}|\psi_1\rangle = \beta\sqrt{E_2}|\varphi\rangle$, ami ellentmond 3.3.1-es egyenletnek, hiszen

$$\langle\psi_2|E_2|\psi_2\rangle = |\beta|^2\langle\varphi|E_2|\varphi\rangle \leq |\beta|^2 < 1,$$

ahol az utolsó előtti egyenlőtlenség abból következik, hogy

$$\langle\varphi|E_2|\varphi\rangle \leq \sum_i \langle\varphi|E_i|\varphi\rangle = \langle\varphi|\varphi\rangle = 1.$$

□

3.4. Kvantum mérések és operációk

Az általános formájú 3.1.3. posztulátum fontos speciális esetei a projektív mérések és POVM mérések. Most ezeket vesszük górcső alá.

3.4.1. Definíció. Egy M Hermitikus operátort, amely az állapottéren hat, *projektív mérésnek* nevezünk. M spektrális felbontása:

$$M = \sum_m mP_m,$$

ahol P_m egy projekció az M-nek az m sajátértékhez tartozó sajátalterére. A mérés lehetséges kimenetei megfelelnek az M operátor m sajátértékeinek. Ha egy $|\psi\rangle$ állapotot mérünk meg, akkor annak a valószínűsége, hogy a végeredmény m lesz, a következő alakban adható meg:

$$p(m) = \langle\psi|P_m|\psi\rangle.$$

Ha a mérés kimente m, akkor rögtön a mérés után a rendszer a

$$\frac{P_m|\psi\rangle}{\sqrt{p(m)}}$$

állapotban van.

A definíció alapján könnyen kiszámolhatjuk a mérés várható értékét:

$$\begin{aligned} E(M) &= \sum_m mp(m) \\ &= \sum_m m\langle\psi|P_m|\psi\rangle \\ &= \langle\psi|\left(\sum_m mP_m\right)|\psi\rangle \\ &= \langle\psi|M|\psi\rangle. \end{aligned}$$

Ez egy igen hasznos összefüggés, ami sok számolást leegyszerűsít. Az M várható értékét gyakran így is jelöljük: $\langle M\rangle = \langle\psi|M|\psi\rangle$. Ezzel az új jelöléssel az M mérés szórását az alábbi alakban is írhatjuk:

$$D(M) = \sqrt{\langle M^2\rangle - \langle M\rangle^2}.$$

Eddig úgy adtunk meg méréseket, hogy definiáltuk, hogy melyik kimenetel milyen valószínűséggel fordul elő, majd definiáltuk, hogy a rendszer a mérést követően milyen állapotba kerül. Gyakran nem fontos, hogy a mérés után milyen állapotba kerül a rendszer, mivel a rendszernek csak az aktuális állapota érdekel minket. Ebből a szempontból a méréseknél általánosabb fogalom kerül most bevezetésre:

3.4.2. Definíció. Tegyük fel, hogy az M_m operátorok által leírt mérések egy $|\psi\rangle$ állapotban lévő rendszeren hatnak. Ekkor annak a valószínűsége, hogy a mérés végeredménye m: $p(m) = \langle\psi|M_m M_m^*|\psi\rangle$. Legyen

$$E_m = M_m M_m^*.$$

Ekkor a 3.1.3. posztulátum és lineáris algebrai megfontolások szerint E_m egy pozitív operátor, melyre $\sum_m E_m = I$ és $p(m) = \langle\psi|E_m|\psi\rangle$. Tehát az E_m operátorok meg tudják határozni a kimeneti valószínűségeket. Az E_m operátorokat a méréshez asszociált *POVM* (Positive Operator-Valued Masure) elemeknek nevezzük. A $\{E_m\}$ teljes halmazt pedig POVM mérésnek hívjuk.

Most példát adunk POVM mérésekre. Tekintsük projektív mérések egy P_m halmazát, ahol P_m egy projekció, vagyis: $P_m P_{m'} = \delta_{mm'} P_m$ és $\sum_m P_m = I$. Ebben az esetben az összes POVM elem megegyezik az eredeti mérések operátorával, vagyis $E_m = P_m^* P_m = P_m$.

A következő formalizmus a kvantum zaj leírását teszi lehetővé.

3.4.3. Definíció. Legyen ρ egy sűrűségoperátor a $V_1 \otimes V_2 \otimes V_3$ téren (V_1, V_2, V_3 vektorterek). Jelölje I_1 a V_1 tér identitását, míg U_{23} a $V_1 \otimes V_2$ tér egy unitér transzformációját. Ekkor a kvantum zaj hatására a ρ állapotú kvantumrendszer a

$$\rho' = \Gamma(\rho) = (I_1 \otimes U_{23}^{-1})\rho(I_1 \otimes U_{23})$$

állapotba kerül. A Γ leképezést *kvantum operáció*nak nevezzük.

4. fejezet

A kvantum-információelmélet alapjai

A klasszikus információelméleti bevezető után a megismert fogalmakat kvantumrendszerekre szeretnénk általánosítani. Néhány eredmény analogonja továbbra is érvényben marad, ám lesz egy-két eset, amikor az intuiciónknak ellentmondó eredményeket kapunk. A fejezetben felépítjük a szükséges technikai hátteret, ami a szakdolgozat legfontosabb tételének bizonyításához szükséges.

4.1. Neumann-entrópia

A Neumann-entrópia a klasszikus esetből ismert Shannon entrópiát általánosítja. Míg egy klasszikus rendszert egy diszkrét valószínűségi változó írt le, addig egy kvantumrendszer állapotát egy komplex Hilbert-téren értelmezett sűrűségoperátor írja le. Definíció szerint sűrűségoperátornak egy 1-nyomú, pozitív szemidefinit mátrixot nevezünk. Így 2.1.2. definíció alapján adódik a következő fogalom:

4.1.1. Definíció. Egy ρ kvantumrendszer *Neumann-entrópiáját* így értelmezzük:

$$S(\rho) = -tr(\rho \log \rho).$$

Mivel a nyomfüggvény hasonlóság invariáns, így a diagonalizálás után a λ_x sajátértékek segítségével az alábbi egyszerűbb alakban is írhatjuk a Neumann-entrópiát:

$$S(\rho) = -\sum_x \lambda_x \log \lambda_x,$$

ahol $0 \log 0 = 0$ hasonló megfontolások miatt, mint a Shannon-entrópiánál. A számolásoknál többnyire az utóbbi formulát fogjuk használni.

A továbbiakban amikor entrópiára utalunk, akkor a szövegkörnyezetből mindig egyértelmű lesz, hogy a Shannon- vagy Neumann-entrópiáról van éppen szó.

4.2. Kvantum relatív entrópia

Ismét hasznos lesz bevezetnünk a relatív entrópia fogalmát.

4.2.1. Definíció. Tegyük fel, hogy ρ és σ két sűrűségoperátor. Ekkor ρ *relatív entrópiáját* σ-ra nézve így értelmezzük:

$$S(\rho||\sigma) = tr(\rho \log \rho) - tr(\rho \log \sigma).$$

A klasszikus relatív entrópiához hasonlóan a kvantum relatív entrópia is lehet olykor $+\infty$. Megállapodás szerint legyen $+\infty$ a kvantum relatív entrópia, ha σ magja nem merőleges ρ képterére. Különben véges. Valóban, mert ha fennáll, hogy σ magtere merőleges ρ képterére, akkor ezzel ekvivalens, hogy $Ker\sigma \subset Ker\rho$, tehát ekkor elég ρ és σ helyett az $Im\sigma$-ra vett megszorításukat venni. Ezek kvantum relatív entrópiája (ami véges) éppen ρ és σ kvantum relatív entrópiája. Fontos eredmény a következő:

4.2.2. Tétel (Klein-egyenlőtlenség)**.** *A relatív entrópia nemnegatív,*

$$S(\rho||\sigma) \geq 0,$$

egyenlőség akkor és csak akkor, ha $\rho = \sigma$.

Bizonyítás. Legyen $\rho = \sum_i p_i |i\rangle\langle i|$ és $\sigma = \sum_j q_j |j\rangle\langle j|$ a megfelelő sűrűségoperátorok mátrixa a diagonalizálás után. A relatív entrópia definíciója miatt írhatjuk, hogy:

$$S(\rho||\sigma) = \sum_i p_i \log p_i - \sum_i \langle i|\rho \log \sigma|i\rangle.$$

Ebbe az egyenlőségbe behelyettesítve a következőket: $\langle i|\rho = p_i\langle i|$ és

$$\langle i| \log \sigma |i\rangle = \langle i| \Big(\sum_j \log q_j |j\rangle\langle j| \Big) |i\rangle = \sum_j \log(q_j) P_{ij},$$

ahol $P_{ij} = \langle i|j\rangle\langle j|i\rangle \geq 0$ kapjuk, hogy

$$S(\rho||\sigma) = \sum_i p_i \big(\log p_i - \sum_j P_{ij} \log(q_j) \big).$$

P_{ij} duplán sztochasztikus mátrix ($\sum_i P_{ij} = 1, \sum_j P_{ij} = 1$), mivel mind a két rendszer ortonormált bázis. A log függvény szigorú konkavitása miatt $\sum_j P_{ij} \log q_j \leq \log r_i$, ahol $r_i = \sum_j P_{ij} q_j$, egyenlőséggel akkor és csak akkor, ha létezik olyan j, melyre $P_{ij} = 1$. Vagyis

$$S(\rho||\sigma) \geq \sum_i p_i \log \frac{p_i}{r_i},$$

egyenlőséggel akkar és csak akkor, ha minden i-re létezik olyan j, melyre $P_{ij} = 1$, vagyis ha P_{ij} permutációmátrix. Az előbbi egyenlőtlenség jobb oldalán a klasszikus relatív entrópia található, amiről már beláttuk, hogy nemnegatív. Ebből adódik a kvantum relatív entrópia nemnegativitása is.

$$S(\rho||\sigma) \geq 0,$$

egyenlőség akkor és csak akkor, ha $p_i = r_i$ minden i-re és P_{ij} egy permutációmátrix.

Az egyenlőségre vonatkozó feltételt egyszerűbben is megfogalmazhatjuk: a permutációmátrix sorait és oszlopait felcserélve feltehető, hogy P_{ij} az egységmátrix. Ez pedig azt jelenti, hogy ρ és σ ugyanabban a bázisban diagonálisak. A $p_i = r_i$ feltételből pedig következik, hogy ρ és σ megfelelő sajátértékei megegyeznek, ami azt jelenti, hogy egyenlőség akkor és csak akkor teljesül, ha ρ=σ. □

4.3. Schmidt-felbontás és purifikáció

4.3.1. Definíció. A 1-rangú ρ sűrűségoperátort *tiszta állapotnak* nevezzük. Ha ρ^A az A kvantumrendszernek az n-dimenziós H_1 komplex Hilbert-téren értelmezett sűrűségoperátora és hasonlóan ρ^B a B kvantumrendszernek az m-dimenziós H_2 komplex Hilbert-téren értelmezett sűrűségoperátora, akkor az AB összetett kvantumrendszer sűrűségoperátora, ρ^{AB} egy olyan sűrűségoperátor, melyre teljesülnek a következők: ρ^{AB} a $H_1 \otimes H_2$ téren értelmezett, $n \times n$-es diagonális blokkjainak összege ρ^A, továbbá ha ρ^{AB}-t $n \times n$-es blokkokra osztjuk, akkor a k-adik sor j-edik oszlopában található blokkmátrix nyoma legyen a ρ^B mátrix k-adik sorának j-edik eleme. Ekkor az AB összetett rendszer marginálisainak nevezzük az A és B rendszereket.

Könnyen látható, hogy minden 1-rangú ρ sűrűségoperátor előállítható egy egységhosszú vektor önmagával vett tenzorszorzataként és hasonlóan minden egységhosszú vektor önmagával vett tenzorszorzata egy 1-rangú sűrűségoperátort határoz meg (az egységhosszúság azért van feltéve, hogy az operátor nyoma 1 legyen). Ez a tiszta állapotok és állapotvektorok közötti (egységhosszú komplex számmal való szorzástól eltekintve) bijekció gyakran a jelölésrendszerben és a nyelvhasználatban is vissza fog köszönni a továbbiakban.

Ismertnek tekintjük az alábbi, numerikus analízisből jól ismert tételt.

4.3.2. Tétel (Szinguláris értékek szerinti felbontás)**.** *Ha M egy komplex értékű $m \times n$-es mátrix, akkor előállítható az alábbi alakban: $M = U\Sigma V^*$, ahol U egy $m \times m$-es unitér mátrix, V egy $n \times n$-es unitér mátrix, míg Σ egy $m \times n$-es diagonális mátrix, amelynek főátlójában nemnegatív valós elemek vannak.*

A következő két tétel fontos szerepet fog játszani a továbbiakban. A 3.1.4. posztulátum szerint az összetett kvantumrendszerek állapottere az őt alkotó rendszerek állapottereinek tenzorszorzata. Ezért hasznos az alábbi felbontás:

4.3.3. Tétel (Schmidt-felbontás)**.** *Legyen H_1 egy n-dimenziós, míg H_2 egy m-dimenziós Hilbert-tér. Tegyük fel, hogy $n \geq m$. Ekkor minden v $H_1 \otimes H_2$-beli vektorra léteznek olyan $\{u_1, ..., u_n\} \subset H_1$ és $\{v_1, ..., v_m\} \subset H_2$ ortonormált rendszerek, hogy $\alpha = \sum_{i=1}^{m} \alpha_i u_i \otimes v_i$, ahol az α_i skalárok nem-negatívak és teljesen meghatározottak a v vektor által.*

Bizonyítás. Rögzítsük H_1-ben az $\{e_1, ...e_n\}$ illetve H_2-ben az $\{f_1, ..., f_m\}$ ortonormált bázisokat. Ekkor egy elemi tenzort, $e_i \otimes f_j$-t azonosítsuk az $e_i f_j^T$ mátrixszal, ahol f_j^T az f_j vektor transzponáltja. Ezután az azonosítás után egy általános eleme a tenzorszorzatnak ilyen alakban írható:

$$v = \sum_{1 \leq i \leq n, 1 \leq j \leq m} \beta_{ij} e_i \otimes f_j.$$

Legyen $M_v = (\beta_{ij})_{ij}$ $n \times m$-es mátrix. Erre a mátrixra alkalmazva a numerikus analízisből ismert szinguláris értékek szerinti felbontást, adódik a

$$M_v = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T$$

felbontás, ahol U és V unitér mátrixok, míg Σ diagonális mátrix. Legyen $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$, ahol U_1 $n \times m$-es mátrix, így $M_v = U_1 \Sigma V^T$. Legyen $\{u_1, ..., u_m\}$ az első m oszlopvektora az U_1 mátrixnak, $\{v_1, ..., v_m\}$ oszlopvektorai a V mátrixnak és $\alpha_1, ...\alpha_m$ diagonálelemei a Σ mátrixnak. Ekkor az M_v mátrix ilyen alakra hozható:

$$M_v = \sum_{k=1}^{m} \alpha_k u_k v_k^T,$$

vagyis

$$v = \sum_{k=1}^{m} \alpha_k u_k \otimes v_k,$$

ami bizonyítja az állítást. □

Az u_k, v_k bázisokat Schmidt-bázisnak nevezzük, míg a nemnulla α_k konstansok számát a v állapotvektor Schmidt-számának nevezzük. A Schmidt-felbontásnak fontos következménye, amit úton-útfélen alkalmazni fogunk, az az észrevétel, mely szerint ha $|\varphi\rangle$ egy AB összetett kvantumrendszer állapota (vagyis az AB rendszer tiszta állapotban van), akkor $S(A) = S(B)$. Ezt könnyen beláthatjuk a Schmidt-felbontással; ekkor $\rho^A = \sum_k \alpha_k^2 |v_k\rangle\langle v_k|$, hasonlóan $\rho^B = \sum_k \alpha_k^2 |u_k\rangle\langle u_k|$, vagyis mindkét sűrűségoperátornak ugyanazok a sajátértékei, amiből már következik, hogy megegyezik az entrópiájuk: lásd 4.1.1. és 4.3.1 definíciókat.

Egy másik igen fontos és gyakran használt technika a kvantum-információelméletben a *purifikáció* vagy tisztítás.

4.3.4. Tétel (Purifikáció). *Minden A kvantumrendszerhez, melynek sűrűségoperátora ρ^A, létezik olyan R kvantumrendszer, hogy az $|AR\rangle$ kvantumrendszer tiszta állapotban van és marginálisai A és R.*

Bizonyítás. Tekintsünk egy A tetszőleges rendszert és a hozzá tartozó ρ^A sűrűségoperátort. Legyen ρ^A a diagonális bázisban a következő alakú: $\rho^A = \sum_i p_i |i^A\rangle\langle i^A|$. Ekkor az R segédrendszer állapottere legyen ugyanannyi dimenziós, mint A állapottere és $|i^R\rangle$ legyen R állapotterében az ortonormált bázis. Ekkor értelmezzük $|AR\rangle$-t az alábbi módon:

$$|AR\rangle = \sum_i \sqrt{p_i} |i^A\rangle |i^R\rangle.$$

Az így értelmezett $|AR\rangle$ állapotvektornak önmagával vett tenzorszorzata egy tiszta állapot, amelynek marginálisai triviálisan A és R. □

4.4. A Neumann entrópia alaptulajdonságai

4.4.1. Tétel (A Neumann entrópia alaptulajdonságai). *(1) Az entrópia nemnegatív. Az $S(\rho)$ entrópia akkor és csak akkor 0, ha ρ egy tiszta állapot.*

(2) A d-dimenzós Hilbert-téren értelmezett sűrűségoperátorok entrópiájának maximuma $\log d$. Az entrópia $\log d$ akkor és csak akkor, ha $\rho = I/d$, ahol I a d-dimenziós egységmátrix.

(3) Legyen AB egy tiszta állapotú, összetett kvantumrendszer. Ekkor $S(A) = S(B)$.

(4) Legyenek p_i *valószínűségek és a* ρ_i *sűrűségoperátorok képei legyenek ortogonális alterek. Ekkor*

$$S(\sum_i p_i\rho_i) = H(p_i) + \sum_i p_i S(\rho_i).$$

(5) Tegyük fel, hogy p_i *valószínűségek,* $|i\rangle$ *ortogonális egységvektorok egy Hilbert-téren, és* ρ_i *egy másik Hilbert-téren értelmezett sűrűségoperátorok tetszőleges halmaza. Ekkor*

$$S(\sum_i p_i|i\rangle\langle i| \otimes \rho_i) = H(p_i) + \sum_i p_i S(\rho_i).$$

Bizonyítás. (1) Nyilvánvaló.

(2) Alkalmazzuk a Klein-egyenlőtlenséget (4.2.2. tétel) a tetszőleges ρ és I/d sűrűségoperátorokra:

$$0 \leq S(\rho||I/d) = -S(\rho) + \log d.$$

(3) Schmidt tételéből tudjuk, hogy ekkor AB marginálisainak ugyanazok a sajátértékei. Ekkor az entrópiájuk is megegyezik: $S(A) = S(B)$.

(4) Legyenek λ_i^j és $|e_i^j\rangle$ a sajátértékei és sajátvektorai a ρ_i sűrűségoperátornak. Vegyük észre, hogy $p_i\lambda_i^j$ és $|e_i^j\rangle$ sajátértékei és sajátvektorai a $\sum_i p_i\rho_i$ sűrűségoperátornak:

$$S(\sum_i p_i\rho_i) = -\sum_{ij} p_i\lambda_i^j \log p_i\lambda_i^j$$

$$= -\sum_i p_i \log p_i - \sum_i p_i \sum_j \lambda_i^j \log \lambda_i^j$$

$$= H(p_i) + \sum_i p_i S(\rho_i).$$

(5) A (4)-es pontból azonnal következik. □

A klasszikus eset analógiájára, hasonló módon a kvantum esetben is definiálhatók a feltételes-, valamint kölcsönös-entrópia illetve kölcsönös információ fogalmai.

4.4.2. Definíció. Az A és B marginálisokkal rendelkező AB összetett rendszernek az entrópiája legyen A és B kölcsönös entrópiája, vagyis: $S(A,B) = -tr(\rho^{AB}\log(\rho^{AB}))$, ahol ρ^{AB} az összetett rendszer sűrűségmátrixa. A klasszikus eset mintájára definiáljuk a *feltételes entrópiát* és a *kölcsönös információt*:

$$S(A|B) = S(A,B) - S(B)$$

$$S(A:B) = S(A) + S(B) - S(A,B) = S(A) - S(A|B) = S(B) - S(B|A).$$

Számos a klasszikus esetben megszokott entrópia-tulajdonság a kvantum esetben már nem igaz. A Shannon entrópiára igaz volt az alábbi összefüggés: $H(X) \leq H(X,Y)$. Ezt az egyenlőtlenséget elég kézenfekvőnek érezzük, hiszen az X-ből szerezhető átlagos információmennyiség nem lehet több az (X,Y)-ból megszerezhető átlagos információmennyiségnél. Ez az egyenlőtlenség viszont kvantum

esetben nem teljesül. Tekintsük a $(|00\rangle + |11\rangle)/\sqrt{2}$ állapotvektor által meghatározott tiszta állapotot. Ennek marginálisai legyenek A és B. Mivel (A, B) tiszta állapot, ezért $S(A, B) = 0$, de mind A-nak és B-nek az $I/2$ a sűrűségoperátora, aminek az entrópiája 1. Ebből adódik, hogy az $S(A : B) = S(B) - S(B|A)$ mennyiség lehet negatív is.

4.5. Mérések és az entrópia

Egy kvantumrendszer szinte sosem tekinthető a külvilágtól és a megfigyelőktől elzárt rendszernek. Gyakran méréseket végzünk a rendszeren, hogy annak belső állapotáról információt szerezzünk. Ebben a pontban azt vizsgáljuk meg, hogy bizonyos mérések hogyan változtatják meg egy kvantumrendszer entrópiáját.

4.5.1. Definíció. Tekintsük P_i projekciók ($P_i^2 = P_i, P_i = P_i^*$) egy halmazát. Ekkor a P_i projekciók által meghatározott *projektív mérés* után a ρ kvantumrendszer a

$$\rho' = \sum_i P_i \rho P_i$$

kvantumállapotba kerül.

4.5.2. Tétel (A projektív mérések növelik az entrópiát)**.** *Legyen P_i projekciók egy teljes halmaza ($\sum_i P_i = I$) és ρ sűrűségoperátor. Ekkor a projektív mérés után a $\rho' = \sum_i P_i \rho P_i$ kvantumállapot entrópiája nem kisebb az eredeti kvantumállapoténál:*

$$S(\rho') \geq S(\rho),$$

egyenlőség akkor és csak akkor, ha $\rho = \rho'$.

Bizonyítás. Ismételten a Klein-egyenlőtlenséget kell alkalmaznunk, 4.2.2. tétel, a ρ és ρ' sűrűségoperátorokra.

$$0 \leq S(\rho||\rho') = -S(\rho) - tr(\rho \log \rho').$$

Tehát elég azt megmutatni, hogy $S(\rho') = -tr(\rho \log \rho')$. Felhasználva, hogy $P_i^2 = P_i$, $\sum_i P_i = I$ és a nyom ciklikus tulajdonságát kapjuk, hogy:

$$-tr(\rho \log \rho') = -tr\Big(\sum_i P_i \rho \log \rho'\Big)$$

$$= -tr\Big(\sum_i P_i \rho \log \rho' P_i\Big).$$

Vegyük észre, hogy $\rho' P_i = P_i \rho P_i = P_i \rho'$, ami azt mutatja, hogy P_i felcserélhető ρ'-vel és ezért $\log \rho'$-vel is, tehát:

$$-tr(\rho \log \rho') = -tr\Big(\sum_i P_i \rho P_i \log \rho'\Big)$$

$$= -tr(\rho' \log \rho') = S(\rho'),$$

amit bizonyítani akartunk. □

4.6. Szubadditivitás és háromszög-egyenlőtlenség

Már említettük, hogy a Shannon entrópiára ismert $H(X,Y) \geq H(X)$ összefüggés kvantum megfelelője nem teljesül. Ennek gyengítéseként fogható fel a "háromszög-egyenlőtlenség".

4.6.1. Tétel (Az entrópia szubadditív és teljesíti a háromszög-egyenlőtlenséget)**.** *Legyen A és B kvantumrendszerek összetett kvantumállapota ρ^{AB}. Ekkor:*

(1) $S(A,B) \leq S(A)+S(B)$; és egyenlőség akkor és csak akkor áll, ha A és B nem korreláltak, azaz $\rho^{AB} = \rho^A \otimes \rho^B$.

(2) $S(A,B) \geq |S(A) - S(B)|$

Bizonyítás. (1) A Klein-egyenlőtlenséget alkalmazzuk, 4.2.2. tétel, a $\rho = \rho^{AB}$ és $\sigma = \rho^A \otimes \rho^B$ kvantumállapotokra: $S(\rho) \leq -tr(\rho \log \sigma)$, ahol

$$-tr(\rho \log \sigma) = -tr(\rho^{AB}(\log \rho^A + \log \rho^B))$$

$$= -tr(\rho^A \log \rho^A) - tr(\rho^B \log \rho^B)$$

$$= S(A) + S(B).$$

Vagyis azt kaptuk, hogy $S(A,B) \leq S(A) + S(B)$, amit bizonyítani kellett. Látható, hogy egyenlőség akkor és csak akkor áll fenn, ha $\rho = \sigma$ a Klein-egyenlőtlenségben, vagyis $\rho^{AB} = \rho^A \otimes \rho^B$.

(2) Az A és B kvantumrendszerekhez választható olyan R segédrendszer, amely tiszta állapotba viszi A-t és B-t. Ezt az eljárást purifikációnak (tisztításnak) nevezzük. Ennek jogosságát a 4.3. szakaszban bizonyítottuk. A szubadditivitást alkalmazva adódik:

$$S(R) + S(A) \geq S(A,R).$$

Mivel ABR tiszta állapotban van, ezért a Schmidt-tételt alkalmazva kapjuk, hogy marginálisainak entrópiája egyenlő, azaz: $S(A,R) = S(B)$ és $S(R) = S(A,B)$. Ezeket az előző egyenlőtlenségbe behelyettesítve kapjuk:

$$S(A,B) \geq S(B) - S(A).$$

Az A és B rendszerek szimmetriájából adódik $S(A,B) \geq S(A) - S(B)$. □

4.7. Konkavitás

4.7.1. Tétel (A Neumann entrópia konkáv)**.** *Legyenek p_i nemnegatív valós számok, hogy $\sum_i p_i = 1$. Továbbá ρ_i jelöljön sűrűségoperátorokat. Ekkor teljesül a következő egyenlőtlenség:*

$$S\Big(\sum_i p_i \rho_i\Big) \geq \sum_i p_i S(\rho_i).$$

Bizonyítás. Tegyük fel, hogy ρ_i egy A rendszer állapotait írja le. Vezessünk be egy B *segédrendszert,* amelynek sűrűségoperátora úgy néz ki, hogy a főátló i-edik eleme p_i, minden máshol pedig 0 áll. Ekkor AB együttes sűrűségoperátora legyen:

$$\rho^{AB} = \sum_i p_i \rho_i \otimes |i\rangle\langle i|.$$

Ennek marginálisai nyilván A és B, hiszen a diagonális blokkok összege $\sum_i p_i \rho_i$, míg a blokkok nyomai valóban a B rendszert határozzák meg: $\sum_i p_i |i\rangle\langle i|$. Vagyis azt kaptuk, hogy:

$$S(A) = S\Big(\sum_i p_i \rho_i\Big),$$

$$S(B) = S\Big(\sum_i p_i |i\rangle\langle i|\Big) = H(p_i).$$

Alkalmazva a ρ^{AB} összetett rendszerre a 4.4.1. tétel (5)-ös pontját nyerjük, hogy:

$$S(A,B) = H(p_i) + \sum_i p_i S(\rho_i).$$

Felhasználva a szubadditivitást $S(A,B) \leq S(A) + S(B)$ kapjuk, hogy:

$$S\Big(\sum_i p_i \rho_i\Big) \geq \sum_i p_i S(\rho_i),$$

ami éppen a konkavitás. □

Érdemes itt megállni egy pillanatra. A kvantum-információelméletben nagyon fontos és gyakran használt az imént bemutatott trükk, vagyis egy *segédrendszer* készítésével történő bizonyítás. Milyen intuíció van mögötte? Az A rendszer p_i valószínűséggel ρ_i állapotban van. Az i értékét természetesen nem tudjuk. Célunk tehát egy olyan sűrűségoperátort csinálni, ami el tudja tárolni i értékét: ha az A rendszer a ρ_i állapotban lenne, akkor a B rendszer $|i\rangle\langle i|$ állapotban lenne és ha mérést végeznénk a B rendszeren, akkor megtudnánk, hogy A melyik állapotban van. Tehát ezért jó választás a B rendszer.

A következő tétel azt mutatja, hogy "mennyire" konkáv az entrópia.

4.7.2. Tétel (Felső becslés kevert állapotok entrópiájára). *Legyen* $\rho = \sum_i p_i \rho_i$, *ahol* p_i *nemnegatív valós számok, melyekre* $\sum_i p_i = 1$ *és* ρ_i *pedig sűrűségoperátorok. Ekkor*

$$S(\rho) \leq \sum_i p_i S(\rho_i) + H(p_i),$$

egyenlőség akkor és csak akkor, ha mindegyik ρ_i *képe páronként ortogonális alterek.*

Bizonyítás. Először tiszta állapotokra bizonyítunk, vagyis minden $\rho_i = |\psi_i\rangle\langle\psi_i|$ valamilyen $|\psi_i\rangle$ egységhosszú vektorra a Hilbert-térből. Tegyük fel, hogy a ρ_i egy A rendszer állapotai. Legyen B egy *segédrendszer* a $|i\rangle$ bázisban p_i valószínűséggel. Az összetett rendszert definiáljuk az alábbi módon:

$$|AB\rangle = \sum_i \sqrt{p_i} |\psi_i\rangle |i\rangle.$$

Mivel $|AB\rangle$ tiszta állapotban van, ezért a Schmidt-tétel alapján kapjuk, hogy

$$S(B) = S(A) = S\Big(\sum_i p_i |\psi_i\rangle\langle\psi_i|\Big) = S(\rho).$$

Tegyük fel, hogy projektív mérést végzünk a B rendszeren az $|i\rangle$ bázisban. Legyen a mérés eredménye:

$$\rho^{B'} = \sum_i p_i |i\rangle\langle i|.$$

Ekkor a 4.5.2. tétel alapján $S(\rho) = S(B) \leq S(B') = H(p_i)$. Mivel mindegyik ρ_i tiszta állapotban van, ezért $S(\rho_i) = 0$ minden i-re. Ebből kapjuk, hogy

$$S(\rho) \leq \sum_i p_i S(\rho_i) + H(p_i),$$

egyenlőség akkor és csak akkor, ha $B = B'$, vagyis ha $|\psi_i\rangle$ ortogonális állapotok.

A kevert állapotok esetében mindegyik állapotot fel tudjuk bontani tiszta állapotok konvex kombinációjára: $\rho_i = \sum_j p_j^i |e_j^i\rangle\langle e_j^i|$, ahol az $|e_j^i\rangle$ ortogonális állapotvektorok. Ezek alapján $\rho = \sum_{ij} p_i p_j^i |e_j^i\rangle\langle e_j^i|$. Alkalmazva a tiszta állapotokra vonatkozó eredményt és, hogy $\sum_j p_j^i = 1$ minden i-re kapjuk, hogy

$$S(\rho) \leq -\sum_{ij} p_i p_j^i \log(p_i p_j^i)$$

$$= -\sum_i p_i \log p_i - \sum_i p_i \sum_j p_j^i \log p_j^i$$

$$= H(p_i) + \sum_i p_i S(\rho_i),$$

amit bizonyítani kellett. Az egyenlőség feltételei pedig ugyanazok, mint a tiszta állapotok eseténél. □

4.8. Erős szubadditivitás

Az erős szubadditivitás kvantum megfelelője igaz, ám ennek bizonyítása némi előkészületet igényel.

4.8.1. Definíció. Legyen $f(A, B)$ valós értékű, mátrixokon értelmezett függvény. Ekkor azt mondjuk, hogy f *kölcsönösen konkáv* A-ban és B-ben, ha minden $0 \leq \lambda \leq 1$ esetén

$$f(\lambda A_1 + (1-\lambda)A_2, \lambda B_1 + (1-\lambda)B_2) \geq \lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2).$$

A definícióból azonnal következik, hogy ha $f(A, B)$ kölcsönösen konkáv függvény, akkor konkáv is rögzített A vagy B mellett. A megfordítás viszont már nem szükségszerűen igaz.

A kölcsönös konkavitás mintájára értelemszerűen definiálhatjuk a kölcsönös konvexitás fogalmát.

Bizonyítás nélkül fel fogjuk használni a következő eredményt.

4.8.2. Tétel (Lieb-tétel). *Legyen X egy mátrix és $0 \leq t \leq 1$. Ekkor az*

$$f(A,B) = tr(X^* A^t X B^{1-t})$$

kölcsönösen konkáv az A, B pozitív mátrixokra.

4.8.3. Tétel (A relatív entrópia konvex). *A relatív entrópia, $S(\rho||\sigma)$ kölcsönösen konvex a változóiban.*

Bizonyítás. Tetszőleges A és X ugyanazon véges dimenziós Hilbert-téren ható mátrixokra legyen

$$I_t(A,X) = tr(X^* A^t X A^{1-t}) - tr(X^* X A).$$

Ebben a kifejezésben az első tag Lieb tétele, 4.8.2. tétel, miatt konkáv A-ban, míg a második tag lineáris A-ban, így $I_t(A,X)$ konkáv A-ban. Legyen

$$I(A,X) = \frac{d}{dt}\bigg|_{t=0} I_t(A,X) = tr(X^*(\log A)XA) - tr(X^* X(\log A)A).$$

Vegyük észre, hogy $I_0(A,X) = 0$. Felhasználva, hogy $I_t(A,X)$ konkáv A-ban kapjuk,

$$\begin{aligned} I(\lambda A_1 + (1-\lambda)A_2, X) &= \lim_{\Delta\to 0} \frac{I_\Delta(\lambda A_1 + (1-\lambda)A_2, X)}{\Delta} \\ &\geq \lambda \lim_{\Delta\to 0} \frac{I_\Delta(A_1,X)}{\Delta} + (1-\lambda)\lim_{\Delta\to 0}\frac{I_\Delta(A_2,X)}{\Delta} \\ &= \lambda I(A_1,X) + (1-\lambda)I(A_2,X). \end{aligned}$$

Vagyis $I(A,X)$ konkáv függvény A-ban. Legyenek az alábbi blokkmátrixok

$$A = \begin{bmatrix} \rho & 0 \\ 0 & \sigma \end{bmatrix}, X = \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix}$$

Egyszerű számolás mutatja, hogy $I(A,X) = -S(\rho||\sigma)$. Az $I(A,X)$ konkavitásából következik, hogy $S(\rho||\sigma)$ kölcsönösen konvex. □

A relatív entrópia konvexitásának egyszerű következménye a következő:

4.8.4. Tétel (A feltételes entrópia konkavitása). *AB legyen egy összetett kvantumrendszer, melynek marginálisai A és B. Ekkor a $S(A|B)$ feltételes entrópia konkáv ρ^{AB}-ben.*

Bizonyítás. Legyen d az A rendszer dimenziója. Ekkor

$$\begin{aligned} S\left(\rho^{AB}\middle|\middle|\frac{I}{d}\otimes\rho^B\right) &= -S(A,B) - tr\left(\rho^{AB}\log\left(\frac{I}{d}\otimes\rho^B\right)\right) \\ &= -S(A,B) - tr(\rho B \log \rho^B) + \log d \\ &= -S(A|B) + \log d. \end{aligned}$$

Mivel $S(A|B) = \log d - S(\rho^{AB}||I/d\otimes\rho^B)$, ezért a feltételes entrópia konkavitása következik a relatív entrópia kölcsönös konvexitásából. □

4.8.5. Tétel (Erős szubadditivitás). *Tetszőleges A, B, C kvantumrendszerekre teljesül az alábbi két egyenlőtlenség:*

$$S(A) + S(B) \leq S(A, C) + S(B, C)$$

$$S(A, B, C) + S(B) \leq S(A, B) + S(B, C).$$

Bizonyítás. Valójában a két egyenlőtlenség ekvivalens egymással. Először a feltételes entrópia alkalmazásával belátjuk az első egyenlőtlenséget, majd utána megmutatjuk, hogy az első egyenlőtlenségből következik a második. $T(\rho^{ABC})$ legyen az a függvény, amelyet így értelmezünk:

$$T(\rho^{ABC}) = S(A) + S(B) - S(A, C) - S(B, C) = -S(C|A) - S(C|B).$$

A feltételes entrópia konkavitása miatt $T(\rho^{ABC})$ konvex függvénye ρ^{ABC}-nek. Legyen $\rho^{ABC} = \sum_i p_i |i\rangle\langle i|$ a diagonális alakja ρ^{ABC}-nek. A T konvexitása miatt $T(\rho^{ABC}) \leq \sum_i p_i T(|i\rangle\langle i|)$. De $T(|i\rangle\langle i|) = 0$, mivel a Schmidt-tétel miatt $S(A, C) = S(B)$ és $S(B, C) = S(A)$. Tehát $T(\rho^{ABC}) \leq 0$ és ezért

$$S(A) + S(B) - S(A, C) - S(B, C) \leq 0,$$

ami pont az első egyenlőtlenséggel ekvivalens.

A második egyenlőtlenség bizonyításához vezessünk be egy R segédrendszert, ami az ABC kvantumrendszert tiszta állapotba viszi (purifikálja). Ekkor az első egyenlőtlenséget alkalmazva:

$$S(R) + S(B) \leq S(R, C) + S(B, C).$$

Mivel ABCR tiszta állapotban van, ezért a Schmidt-tétel szerint $S(R) = S(A, B, C)$ és $S(R, C) = S(A, B)$, vagyis

$$S(A, B, C) + S(B) \leq S(A, B) + S(B, C),$$

amit bizonyítani akartunk. □

4.9. Az erős szubadditivitás: néhány következmény

Ebben a pontban az erős szubadditivitás néhány egyszerű, ám annál fontosabb következményét látjuk be, amik szükségesek a szakdolgozat fő tételének, a Holevo-korlátnak bizonyításához.

Érdekes észrevétel, hogy az erős szubadditivitás mind a klasszikus, mind a kvantum esetben igaz, habár teljesen más okokból. A klasszikus megfelelője a $S(A) + S(B) \leq S(A, C) + S(B, C)$ egyenlőtlenségnek már abból a tényből is következik, hogy $H(A) \leq H(A, C)$ és $H(B) \leq H(B, C)$. A két egyenlőtlenség összege az erős szubadditivitást adja. Kvantum esetben, mint láttuk előfordulhat, hogy $S(A) > S(A, C)$ vagy $S(B) > S(B, C)$, de éppen az erős szubadditivitás miatt egyszerre a kettő nem teljesülhet.

Az erős szubadditivitás másik igen hasznos alakja:

$$0 \leq S(C|A) + S(C|B)$$

vagy a vele ekvivalens

$$S(C:A) + S(C:B) \leq 2S(C).$$

4.9.1. Tétel (Az erős szubadditivitás következményei). *(1) Legyen ABC egy összetett rendszer. Ekkor $S(A|B,C) \leq S(A|B)$.*

(2) Legyen ABC egy összetett rendszer. Ekkor $S(A:B) \leq S(A:B,C)$.

(3) Tegyük fel, hogy ABC egy összetett kvantumrendszer, továbbá Γ egy kvantumoperáció, amely az ABC rendszeren hat. Ekkor jelölje $S(A:B)$ a kölcsönös információját az A és B marginálisoknak, mielőtt még Γ-t alkalmaztuk volna ABC-re. Jelölje $S(A':B')$ az A, B marginálisok kölcsönös információját, miután Γ-t alkalmaztuk ABC-re. Ekkor $S(A':B') \leq S(A:B)$.

Bizonyítás. (1) A bizonyítás teljesen hasonlóan megy, mint a klasszikus esetben: lásd 2.3.2. tétel (7)-es pontját. Kvantum esetben ez így néz ki: $S(A|B,C) \leq S(A|B)$ ekvivalens azzal, hogy $S(A,B,C) - S(B,C) \leq S(A,B) - S(B)$. Ezt átrendezve kapjuk a $S(A,B,C) + S(B) \leq S(A,B) + S(B,C)$ egyenlőtlenséget, ami éppen az erős szubadditivitás.

(2) $S(A:B) \leq S(A:B,C)$ ekvivalens azzal, hogy $S(A) + S(B) - S(A,B) \leq S(A) + S(B,C) - S(A,B,C)$, ezt pedig átrendezve kapjuk az $S(A,B,C) + S(B) \leq S(A,B) + S(B,C)$ egyenlőtlenséget, ami az erős szubadditivitás.

(3) Ahogy azt korábban láttuk a 3.4. fejezetben a Γ kvantumoperáció hatása ABC-n leírható egy $I \otimes U$ transzformációval való konjugálással, ahol I identikus az A állapotterén, U unitér transzformáció a BC állapotterén. Γ hatása ABC-re éppen olyan, mintha az BC állapottéren egy U_{BC} unitér báziscsrét végeztünk volna el. Ha a vesszőzött állapotok jelölik a Γ utáni állapotokat, akkor kezdetben $S(A:B) = S(A:B,C)$, mivel B és C korrelálatlanok. Mivel a BC állapotterén egy unitér báziscserét végzünk, az A rendszer állapotterét helyben hagyjuk, így a sajátértékek nem változnak, tehát Γ hatása után $S(A:B,C) = S(A':B',C')$. Felhasználva az előző pontot: $S(A':B') \leq S(A':B',C')$. Mindezeket összetéve kapjuk, hogy $S(A':B') \leq S(A:B)$. □

A 2.3.2. tétel után láttuk, hogy a Shannon-féle kölcsönös információ nem mindig szubadditív, így a Neumann-féle kölcsönös információ sem az. Viszont a feltételes entrópia már szubadditív, sőt ennél több is igaz:

4.9.2. Tétel (A feltételes entrópia szubadditivitása). *Legyen $ABCD$ egy összetett kvantumrendszer. Ekkor a feltételes entrópia együttesen szubadditív mindkét változóban:*

$$S(A,B|C,D) \leq S(A|C) + S(B|D).$$

Ha ABC egy összetett kvantumrendszer, akkor a feltételes entrópia mind az első, mint a második változóban szubadditív:

$$S(A,B|C) \leq S(A|C) + S(B|C)$$

$$S(A|B,C) \leq S(A|B) + S(A|C).$$

Bizonyítás. Alkalmazzuk az erős szubadditivitást az $ABCD$ összetett rendszerre; kapjuk, hogy

$$S(A,B,C,D) + S(C) \leq S(A,C) + S(B,C,D).$$

Mindkét oldalhoz $S(D)$-t hozzáadva:

$$S(A,B,C,D) + S(C) + S(D) \leq S(A,C) + S(B,C,D) + S(D).$$

A jobb oldal utolsó két tagját becsüljük felül az erős szubadditivitással:

$$S(A,B,C,D) + S(C) + S(D) \leq S(A,C) + S(B,D) + S(C,D).$$

Ezt az egyenlőtlenséget átrendezve kapjuk

$$S(A,B|C,D) \leq S(A|C) + S(B|D),$$

ami éppen az együttes szubadditivitás.

A feltételes entrópia szubadditivitása az első változó szerint, $S(A,B|C) \leq S(A|C) + S(B|C)$, triviálisan ekvivalens az erős szubadditivitással. A második változó szerinti szubadditivitás már nem ilyen egyszerű. Azt szeretnénk megmutatni, hogy $S(A|B,C) \leq S(A|B) + S(A|C)$. Ez ekvivalens a következő egyenlőtlenséggel:

$$S(A,B,C) + S(B) + S(C) \leq S(A,B) + S(B,C) + S(A,C).$$

Legalább az egyike az alábbi egyenlőtlenségeknek teljesül a 4.9.1. tétel előtti megjegyzés szerint: $S(C) \leq S(A,C)$ vagy $S(B) \leq S(A,B)$. Tegyük fel, hogy $S(C) \leq S(A,C)$. Ehhez hozzáadva az erős szubadditivitást: $S(A,B,C) + S(B) \leq S(A,B) + S(B,C)$, éppen a bizonyítandó állítást nyerjük. □

5. fejezet

A Holevo-korlát és következményei

A klasszikus információelmélet központi problémája, hogy hogyan lehet klasszikus információt (biteket) átküldeni egy csatornán, ami a klasszikus fizika törvényei szerint működik. A kvantuminformációelmélet ugyanezt a kérdést a kvantummechanika felől közelíti meg: hatékonyabban lehet-e információt küldeni kvantumrendszerekkel kvantummechanikai csatornán? Lehet-e kvantumrendszerekkel úgy üzeneteket küldeni, hogy ne tudják azt lehallgatni? Számos ilyen és ehhez hasonló kérdést lehetne föltenni, de ebben a fejezetben csak azzal foglalkozunk, hogy mennyi információt lehet egy kvantumrendszerből "kinyerni". A Holevo-korlát erre a "kinyerhető" információmennyiségre ad egy felső becslést.

5.1. Kvantumállapotok megkülönböztetése és az elérhető információ

Az alábbi példával érzékeltetjük a klasszikus és kvantum információ közötti különbséget. Legyen Aladár és Béla két olyan személy, akik az alábbi módon kommunikálnak. Aladárnak rendelkezésére áll egy klasszikus üzenetek kibocsátására alkalmas gép, mely az $X = 0,1,...,n$ jeleket továbbítja $p_0, p_1, ..., p_n$ valószínűséggel. Béla célja, hogy minél jobb valószínűséggel ki tudja találni az Aladár által küldött X értékét. Ennek érdekében Aladár készít egy ρ_X kvantumállapotot, melyet egy rögzített $\rho_0, \rho_1, ..., \rho_n$ halmazból választ. Ezt a kiválasztott ρ_X állapotot elküldi Bélának, aki kvantummérést végez a kapott állapoton és megteszi a tippjét X értékéről az Y mérési eredmény függvényében.

Béla információmennyiségét X-ről jól kifejezi a kölcsönös információ, azaz $H(X:Y)$. Nyilván Béla csak akkor tud teljes biztonsággal következtetni X-re, ha $H(X:Y) = H(X)$. Mivel mindig igaz, hogy $H(X:Y) \leq H(X)$, ezért valóban jó mértéke a "kinyerhető" információnak $H(X:Y)$-nak $H(X)$-hez való közelsége. Tehát Béla olyan mérést igyekszik majd választani, hogy $H(X)$-hez minél közelebb "vigye" $H(X:Y)$-t.

5.1.1. Definíció. Béla azon információmennyisége, amelyet X-ről szerezhet az Y mérési eredmény hatására: $H(X:Y)$. Ezen $H(X:Y)$-k maximumát véve az összes lehetséges mérésen kapjuk az *elérhető információ* fogalmát.

Az elérhető információ tehát egy olyan mérőszám, amely azt fejezi ki, hogy Béla milyen biztonsággal tud X értékére következtetni.

A klasszikus információelméletben nem véletlenül nem került elő az elérhető információ fogalma. Klasszikus esetben két állapot megkülönböztetése lehet, hogy nehéz (gondoljunk egy zavaros kézírásra), de ezt semmilyen elv nem zárja ki. Nem úgy, mint a kvantum esetben: ortogonális állapotoktól eltekintve nem lehet kvantum állapotokat teljes biztonsággal megkülönböztetni. Ennek bizonyítását korábban, a 3.3. szakaszban már elvégeztük.

Ezt a jelenséget fogalmazzuk újra mostani tudásunk szerint. Ha Aladár p valószínűséggel egy $|\varphi\rangle$ állapotot, $(1-p)$ valószínűséggel egy $|\varphi\rangle$-re nem merőleges $|\psi\rangle$ állapotot készít, akkor Béla elérhető információja biztosan kisebb lesz, mint $H(p)$. Klasszikus esetben, ha Aladár a 0 bitet p valószínűséggel küldi, míg az 1 bitet $1-p$ valószínűséggel, akkor nincsen semmilyen elvi akadálya, hogy Béla meg tudja különböztetni a két állapotot, így Béla elérhető információja megegyezik a $H(p)$ entrópiával.

Még érzékletesebbé tehetjük a kvantum esetet az alábbi példával. Képzeljük el, hogy Aladár a 0 és 1 állapotokat két különböző diszkrét valószínűségi változó szerint készíti el. Legyenek ezek a $(p,1-p)$ és a $(q,1-q)$ eloszlások. Ekkor Béla feladata az volna, hogy a kapott 0 vagy 1 állapot alapján következtessen, hogy Aladár azt melyik valószínűségi változó szerint készítette el. Nyilván ezt nem tudja megtenni teljes biztonsággal.

5.2. A Holevo-korlát

A kvantum-információelmélet egyik legalapvetőbb tétele a Holevo-korlát, mely felső korlátot ad az elérhető információra. A tételt 1973-ban publikálta Alexander Holevo orosz matematikus.

5.2.1. Tétel (A Holevo-korlát)**.** *Tegyük fel, hogy Aladár elkészíti a ρ_X kvantumállapotot, ahol $X = 0,1,...,n$ $p_0, p_1, ..., p_n$ valószínűségekkel. Ekkor Béla a ρ_X állapoton kvantummérést végez el az $E_y = E_0, ..., E_m$ halmazbeli POVM elemekkel, melynek eredménye Y. Ekkor bármilyen mérést is végez Béla az elérhető információra ez a felső becslés adható:*

$$H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x),$$

ahol $\rho = \sum_x p_x \rho_x$.

A Holevo-tétel jobb oldalán szereplő mennyiség olyan fontos, hogy külön neve is van: a Holevo-féle χ mennyiség, amit gyakran csak χ-vel jelölnek.

Bizonyítás. A bizonyítás alapja egy konstrukció, amely három kvantum rendszert használ: P,Q és M rendszereket. Aladár a Q rendszert adja Bélának amelyen Béla mérést végez, ezt az M segédrendszerrel írjuk le. P-re úgy gondolhatunk, mint egy előkészületi rendszerre. Definíció szerint P-nek legyen $|x\rangle$ ortonormált bázisa, melyek az $X = 0,1,...,n$ értékeknek felelnek meg. M-re gondolhatunk úgy, mint

Béla mérőműszerére, amely az $|y\rangle$ bázisban van, amelyek az $1, ..., n$ lehetséges mérési eredményeknek felelnek meg. Az egész rendszer a kezdeti állapotban definíció szerint legyen a következő:

$$\rho^{PQM} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|,$$

ahol a tenzor fölbontást PQM sorrendjében végezzük el. Intuitívan a PQM rendszer ezen állapota azt jelenti, hogy Aladár az x értéket választotta p_x valószínűséggel, majd elkészítette az ennek megfelelő ρ_x kvantum állapotot, amit odaadott Bélának, akinek a mérőműszere a $|0\rangle$ kezdeti állapotban van.

Hogyan tudjuk leírni Béla egy mérését? Erre kiválóan alkalmas a már jól ismert kvantum operációk nyelve. Legyen tehát Γ kvantum operáció, amely a Q és M rendszereken hat. A Γ hatása az, hogy az E_y halmazból mérést végez Q-n, aminek eredményét eltárolja az M rendszerben:

$$\Gamma(\sigma \otimes |0\rangle\langle 0|) = \sum_y \sqrt{E_y}\sigma\sqrt{E_y} \otimes |y\rangle\langle y|,$$

ahol σ Q-nak tetszőleges állapota és $|0\rangle$ a mérőműszer kezdőállapota. Könnyű ellenőrizni, hogy az így definiált Γ kvantum operáció. Ez azon múlik,hogy a $v \otimes |0\rangle \mapsto \sum_y \sqrt{E_y}v \otimes |y\rangle$ lineáris izometrikus beágyazása-e a $Q \otimes \mathbb{C}|0\rangle$ térnek a $Q \otimes M$ térbe. Ezt pedig az alábbi egyenlőség-lánc mutatja: $\sum_y |\sqrt{E_y}v|^2 = \sum_y \langle\sqrt{E_y}v, \sqrt{E_y}v\rangle = \sum_y \langle E_y v, v\rangle = \langle\sum_y E_y v, v\rangle = \langle v, v\rangle = |v|^2$. Mivel minden izometria kiterjeszthető unitéren ezért Γ valóban kvantum operáció.

A Γ alkalmazása után a PQM rendszer marginálisait a P',Q' és R' hármas jelölje. Ekkor $S(P : Q) = S(P : Q, M)$, mivel M kezdetben korrelálatlan Q-val és P-vel. A 4.9.1. tétel szerint Γ nem növeli a kölcsönös információt: $S(P : Q, M) \geq S(P' : Q', M')$. Ugyancsak a 4.9.1. tétel szerint $S(P' : Q', M') \geq S(P' : M')$. Vagyis összességében azt kaptuk, hogy

$$S(P : Q) \geq S(P' : M').$$

Ez pedig lényegében a Holevo-korlát! Kis algebrai átalakítás után mindjárt nyilvánvalóvá válik! Először tekintsük az egyenlőtlenség bal oldalát. Vegyük észre, hogy

$$\rho^{PQ} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x,$$

amiből következik, hogy $S(P) = H(p_x)$, és $S(Q) = S(\rho)$. Továbbá $S(P, Q) = H(p_x) + \sum_x p_x S(\rho_x)$ a 4.4.1. tétel (5)-ös pontja szerint. Mindezeket egybevetve adódik,

$$S(P : Q) = S(P) + S(Q) - S(P, Q) = S(\rho) - \sum_x p_x S(\rho_x),$$

ami éppen a Holevo-féle χ mennyiség! Most számoljuk ki $S(P' : M')$-t. Vegyük észre, hogy

$$\rho^{P'Q'M'} = \sum_{xy} p_x |x\rangle\langle x| \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|.$$

$P'Q'M'$ rendszer $P'M'$ marginálisát szeretnénk kiszámolni. Ehhez vegyük észre, hogy az (X,Y) pár együttes eloszlására teljesül: $p(x,y) = p_x p(y|x) = p_x tr(\rho_x E_y) = p_x tr(\sqrt{E_y}\rho_x\sqrt{E_y})$. Ebből következik, hogy

$$\rho^{P'M'} = \sum_{xy} p(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|.$$

Tehát $S(P' : M') = H(X : Y)$. És éppen ezt akartuk bizonyítani. Ezzel befejeztük a Holevo-korlát bizonyítását.

□

5.3. A Holevo-korlát néhány következménye és egy példa

A Holevo-korlát sarokköve a kvantum-információelméletnek. Ebben a pontban néhány egyszerűbb tételt, állítást fogunk megmutatni ennek illusztrálására. A Holevo-féle χ mennyiségre könnyen adhatunk felső becslést, ami a Neumann-entrópia konkavitására adott felső becsléssel ekvivalens; 4.7.2. tétel:

$$S(\rho) - \sum_x p_x S(\rho_x) \leq H(X). \tag{5.3.1}$$

Egyenlőség 5.3.1-ben akkor és csak akkor áll, ha ρ_x képei ortogonális alterek a 4.4.1. tétel (4)-es és (5)-ös pontja szerint. Vagyis ha ρ_x képei nem ortogonális alterek, akkor az egyenlőtlenség éles, tehát ebben az esetben Béla nem tudja teljes biztonsággal megmondani Y mérési eredménye alapján X értékét. Ez a megállapítás van összhangban van a korábbi tétellel, mely szerint nem ortogonális állapotokat nem lehet teljes biztonsággal megkülönböztetni.

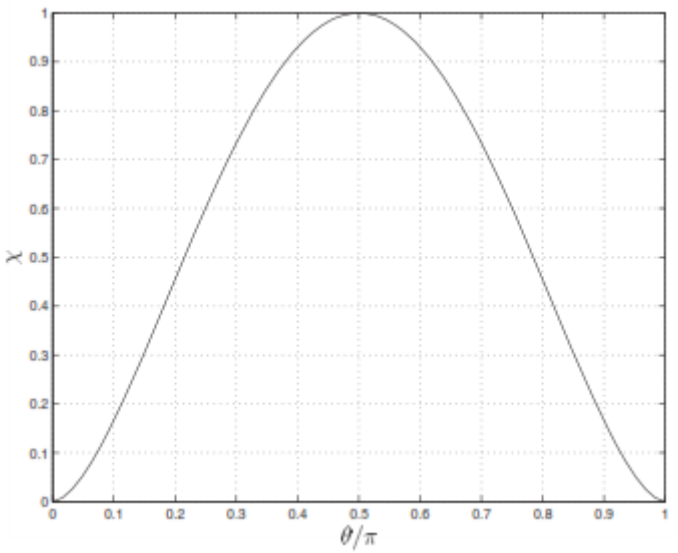
Most nézzük meg egy konkrét példán, hogy működik a tétel! Tegyük fel, hogy az alábbi két kvantum állapot valamelyikét küldi Aladár azonos, 1/2-1/2 valószínűséggel Bélának: $|0\rangle$ állapotot vagy $\cos\theta|0\rangle + \sin\theta|1\rangle$ qubitot küldi, ahol θ valós paraméter. Ekkor a $|0\rangle$,$|1\rangle$ bázisban így írhatjuk fel ρ-t:

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}.$$

Egyszerű számolás után kapjuk, hogy ρ sajátértékei: $(1 \pm \cos\theta)/2)$. Mivel a qubitok tiszta állapotban vannak, ezért $H(X : Y) \leq S(\rho)$, azaz

$$H(X : Y) \leq -\Big(\big((1+\cos\theta)/2\big)\log\big((1+\cos\theta)/2\big) + \big((1-\cos\theta\big)/2)\log\big((1-\cos\theta)/2\big)\Big).$$

5.1. ábra. A Holevo-féle χ mennyiség θ/π függvényében



Ez azt jelenti, hogy $H(X:Y)$ akkor maximális, ha $\theta = \pi/2$ és ekkor az értéke 1 bit. A maximumot tehát akkor éri el az elérhető információ, ha Aladár ortogonális állapotokat küld Bélának. θ-nak minden más értékére az elérhető információ szigorúan kisebb, mint 1 bit, ahogy az ábra is mutatja. Ekkor Béla nem tudja teljes biztonsággal eldönteni, hogy Aladár melyik állapotot küldte neki.

Igencsak meglepő a következő egyszerű állítás:

5.3.1. Tétel (n-szintű kvantum rendszer elérhető információja)**.** *n qubittal nem lehet több információt közölni, mint n klasszikus bit.*

Bizonyítás. Használjuk a Holevo-korlátot és becsüljük felülről az elérhető információt:

$$H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x) \leq S(\rho),$$

mivel az entrópia nemnegatív a 4.4.1. tétel (1) pontja szerint. Mivel egy n qubit egy 2^n dimenziós komplex Hilbert-térben van, ezért a 4.4.1. tétel (2) pontja szerint $S(\rho) \leq \log(2^n) \leq n$, amit bizonyítani akartunk. □

Irodalomjegyzék

[1] Michael A. Nielsen és Isaac L. Chuang *Quantum Computation and Quantum Information* 2010: Cambridge University Press 1

[2] M. Ohya és Dénes Petz *Quantum Entropy and its use* 2004: Springer-Verlag, Heidelberg 1

[3] Dénes Petz *Quantum Information Theory and Quantum Statistics* 2008: Springer-Verlag, Heidelberg 1

[4] Mark Braverman *Information Theory in Computer Science* 2011: Princeton (Interneten elérhető jegyzet) http://www.cs.princeton.edu/courses/archive/fall11/cos597D/ 1