ВВЕДЕНИЕ В КВАНТОВУЮ ТЕОРИЮ ИНФОРМАЦИИ

А. С. Холево

Москва – 2013

Оглавление

I **7**

1 Статистическая структура **9**
1.1 Классические и квантовые системы 9
1.2 Гильбертово пространство 11
1.3 Квантовые состояния 12
1.4 Двухуровневые системы 14
1.5 Анализ понятия “наблюдаемая” 15
1.6 Экстремальные наблюдаемые 17
1.7 Переполненные системы векторов 19
1.8 Переполненные системы для q-бита 20
1.9 Томография квантового состояния 22
1.10 Теорема Наймарка 23

2 Составные квантовые системы **25**
2.1 Наводящие соображения 25
2.2 Тензорное произведение гильбертовых пространств 26
2.3 Разложение Шмидта и очищение 28
2.4 Парадокс ЭПР. Неравенство Белла 29
2.5 Квантовая псевдотелепатическая игра 32
2.6 Корреляционные неравенства и операторные алгебры 34

3 Применения сцепленных состояний **35**
3.1 Квантовое состояние как информационный ресурс 35
3.2 Сверхплотное кодирование 36
3.3 Квантовая телепортация 38
3.4 Квантовые алгоритмы 40
3.4.1 Алгоритм Саймона 40
3.4.2 Замечания об алгоритме Шора 43
3.4.3 Алгоритм Гровера 43

4 Классически-квантовые каналы **45**
4.1 Классическая теория информации 45
4.1.1 Энтропия и сжатие данных 45
4.1.2 Пропускная способность канала с шумом 47

4.2 Оптимальное различение квантовых состояний 51
4.2.1 Постановка задачи 51
4.2.2 Различение по максимуму правдоподобия 52
4.2.3 Максимум информации 56
4.3 Сжатие квантовой информации 60
4.4 Квантовая теорема кодирования 63
4.5 Квантовая граница информации 66
4.6 Доказательство прямой теоремы 70

Предисловие

Квантовая теория информации (КТИ) – новая, быстро развивающаяся научная дисциплина, которая изучает общие закономерности передачи, хранения и преобразования информации в системах, подчиняющихся законам квантовой механики. Квантовая теория информации использует математический аппарат матричного и операторного анализа, некоммутативной теории вероятностей и статистики для исследования потенциальных возможностей таких систем, а также разрабатывает принципы их рационального и помехоустойчивого дизайна. КТИ стимулирует развитие экспериментальной физики, значительно расширяющее возможности целенаправленного манипулирования состояниями микросистем и потенциально важное для новых эффективных приложений. В настоящее время работы в области квантовой информатики, включающей КТИ, экспериментальные и технологические разработки, ведутся в научно-исследовательских центрах всех развитых стран.

Настоящий курс лекций вводят в круг основных понятий КТИ и отражает ряд ее принципиальных достижений. Появление идей квантового компьютинга, квантовой криптографии и новых коммуникационных протоколов позволило говорить не только об ограничениях, но и о новых возможностях, заключенных в использовании специфически квантовых ресурсов, таких как сцепленность (запутанность) квантовых состояний, квантовый параллелизм, дополнительность между измерением и возмущением. Необычные возможности квантовых систем передачи и преобразования информации проиллюстрированы на примерах сверхплотного кодирования, квантовой телепортации и эффективных квантовых алгоритмов. Часть I соответствует содержанию первого семестра. Часть II (второй семестр) будет посвящена фундаментальному понятию квантового канала связи и его энтропийным и информационным характеристикам.

В лекциях приведены необходимые предварительные сведения из классической теории информации и дается введение в статистическую структуру квантовой теории, поэтому для их понимания достаточно владения основными общематематическими дисциплинами. Настоящий курс лекций осуществляется в рамках сотрудничества с Российским Квантовым Центром. Комментарии, предложения, замечания просьба присылать по адресу: ah@icqt.org.

Часть I

Глава 1

Статистическая структура квантовой теории

Прежде чем перейти собственно к квантовой теории информации, необходимо изложить предварительные сведения о *статистической структуре квантовой теории*. Цель состоит не только в том, чтобы ввести определения и зафиксировать обозначения, но и в том, чтобы глубже разобраться в основах квантовой теории и ее вероятностной интерпретации (более полное изложение этих вопросов слушатель найдет в [9]).

Мы будем иметь дело с конечномерными квантовыми системами. С одной стороны, уже в этом случае, причем наиболее наглядно, проявляются радикальные отличия квантовой статистики. С другой, именно системы с конечным числом уровней представляют интерес с точки зрения квантового компьютинга (впрочем, в квантовой теории передачи информации большое внимание привлекают и "системы с непрерывными переменными", которые описываются бесконечномерными пространствами).

1.1 Классические и квантовые системы

Классическая система характеризуется наличием *фазового пространства* Ω, точки которого ω описывают детерминированные состояния системы. Для простоты далее рассматривается случай конечного множества Ω, $d = |\Omega|$. (Статистическим) *состоянием* называется распределение вероятностей на Ω:

$$P = \{p_1, \ldots, p_d\}\,; \quad p_\omega \geq 0, \sum_\omega p_\omega = 1.$$

Вещественная случайная величина:

$$X = \{x_1, \ldots, x_d\}\,; \quad \bar{x}_\omega = x_\omega$$

Математическое ожидание случайной величины X в состоянии P:

$$\mathsf{E}_P X = \sum_\omega p_\omega x_\omega.$$

Для плавного перехода к квантовым системам полезно ввести представление классических величин диагональными матрицами

$$P = \mathrm{diag}\,[p_\omega]\,, \quad X = \mathrm{diag}\,[x_\omega]\,, \quad \mathsf{E}_P X = \mathrm{Tr} PX,$$

где Tr – след матрицы.

Квантовая система описывается d-мерным пространством $\mathbb{C}^d$. Квантовое состояние задается матрицей плотности

$$S = [s_{ij}]_{i,j=1,\ldots,d}\,, \quad S^* = S \geq 0, \quad \mathrm{Tr} S = 1.$$

Вещественная квантовая наблюдаемая задается эрмитовой матрицей

$$X = [x_{ij}]_{i,j=1,\ldots,d}\,, \quad X^* = X.$$

Математическое ожидание наблюдаемой X в состоянии S дается *статистическим постулатом* Борна:

$$\mathsf{E}_S X = \mathrm{Tr} SX. \tag{1.1}$$

При таком подходе обнаруживается аналогия в статистическом описании классических и квантовых систем: сначала приготавливается состояние (P или S), затем производится измерение случайной величины или наблюдаемой X. Как приготовление, так и измерение несут в себе случайность, в результате чего исход измерения случаен, причем его математическое ожидание дается формулой (1.1). При этом для каждой квантовой величины – состояния или наблюдаемой, представляемой эрмитовой матрицей – существует ортонормированный базис из собственных векторов, в котором эта величина представляется диагональной матрицей. Фундаментальное отличие классического описания состоит в том, что оно использует только коммутирующие величины, $XY = YX$. В самом деле, все диагональные матрицы коммутируют между собой. В известном смысле верно и обратное:

Т е о р е м а 1. *Эрмитовы матрицы $A^{(1)}, \ldots, A^{(m)}$ попарно коммутируют тогда и только тогда, когда они совместно диагонализуемы, т.е. существует ортонормированный базис из общих для них собственных векторов.*

Доказательство. Проведем доказательство для двух матриц A, B, которое обобщается очевидным образом. Переходя к базису, в котором A диагональна, мы можем считать, что $A = \mathrm{diag}[a_j], B = [b_{jk}]$. Из условия $AB - BA = 0$ получаем $(a_j - a_k)b_{jk} = 0$. Таким образом, $a_j \neq a_k$ влечет $b_{jk} = 0$. Группируя вместе одинаковые a_j, получаем, что матрицы A, B можно представить в блочно-диагональном виде $A = \mathrm{diag}[a'_j I_j], B = \mathrm{diag}[B_j]$, где все a'_j различны, I_j – единичные матрицы, размерности которых d_j

равны кратности a'_j, а B_j – эрмитовы $d_j \times d_j$-матрицы. Теперь в каждом блоке B_j можно перейти к базису, в котором B_j диагональна, при этом вид матрицы A не изменится.□

Некоммутирующие матрицы $X, Y; XY \neq YX$, описывают несовместимые наблюдаемые, т.е. такие, которые невозможно точно измерить одновременно. Существование несовместимых наблюдаемых – это проявление квантового свойства *дополнительности*. Физические измерения над микрообъектами производятся при помощи макроскопических экспериментальных устройств, предполагающих сложную и специфичную пространственно - временную организацию окружающей среды. Различные способы такой организации, соответствующие измерениям различных наблюдаемых, могут быть взаимно исключающими (несмотря на то, что относятся к одинаково приготовленному микрообъекту), то есть *дополнительными*. Аналогичные соображения относятся и к приготовлению квантовых состояний. Дополнительность – это первое фундаментальное отличие квантовой системы от классической. Существуют и промежуточные "гибридные" системы, сочетающие черты классического и квантового описания (системы с правилами суперотбора). Математической моделью таких систем являются алгебры матриц или операторов (алгебры фон Неймана).

1.2 Гильбертово пространство

Пусть $\mathcal{H}$ - d-мерное комплексное векторное пространство размерности $\dim \mathcal{H} = d < \infty$, со скалярным произведением $\langle\phi|\psi\rangle, \phi, \psi \in \mathcal{H}$, удовлетворяющее аксиомам унитарного пространства; следуя скорее физической, нежели математической традиции, мы считаем, что $\langle\phi|\psi\rangle$ линейно по второму аргументу ψ и антилинейно по первому ϕ. Мы будем использовать дираковские обозначения: вектор ψ из $\mathcal{H}$ (который удобно представлять себе как вектор-столбец) обычно будет обозначаться $|\psi\rangle$; соответственно, $\langle\psi|$ обозначает вектор сопряженного пространства (эрмитово сопряженный вектор-строку). При этом $\langle\phi|\psi\rangle$ естественно обозначает скалярное произведение. Эти обозначения позволяют удобно записывать операторы, например, $A = |\psi\rangle\langle\phi|$ — оператор ранга 1, действующий на вектор $|\chi\rangle$ по формуле $A|\chi\rangle = |\psi\rangle\langle\phi|\chi\rangle$. Если $\langle\psi|\psi\rangle = 1$, то $|\psi\rangle\langle\psi|$ — проектор на единичный вектор $|\psi\rangle$.

Пусть $\{e_i\}_{i=1,\ldots,d}$ – ортонормированный базис (о.н.б.) в $\mathcal{H}$. Произвольный вектор $\psi \in \mathcal{H}$ может быть представлен в виде

$$|\psi\rangle = \sum_{i=1}^{d} |e_i\rangle\langle e_i|\psi\rangle, \tag{1.2}$$

что эквивалентно равенству

$$\sum_{i=1}^{d} |e_i\rangle\langle e_i| = I, \tag{1.3}$$

где I – единичный оператор в $\mathcal{H}$.

З а д а ч а 1. Запишите матричное представление для операторов в $\mathcal{H}$, аналогичное представлению векторов (1.2).

Дополнительным преимуществом обозначений Дирака является возможность записи вместо векторов их меток, например, можно писать просто $|i\rangle$ вместо $|e_i\rangle$.

Иногда мы будем рассматривать *вещественное* гильбертово (т. е. евклидово) пространство. Фундаментальное отличие комплексного случая проявляется в существовании *поляризационного тождества*

$$\beta(\phi,\psi)=\frac{1}{4}\sum_{k=0}^{3}(-i)^k\beta(\phi+i^k\psi,\phi+i^k\psi), \tag{1.4}$$

позволяющего восстановить все значения формы $\beta(\phi,\psi)$, линейной по второму аргументу и антилинейной по первому, по ее диагональным значениям $\beta(\psi,\psi),\psi\in\mathcal{H}$ (в вещественном случае подобное восстановление возможно лишь для симметричных форм). Благодаря этому, например, для доказательства операторного равенства $A=B$ достаточно установить равенство всех диагональных матричных элементов $\langle\psi|A\psi\rangle=\langle\psi|B\psi\rangle,\psi\in\mathcal{H}$.

Сведения об операторах в конечномерном гильбертовом пространстве, которые используются в дальнейшем, приведены в Приложении.

1.3 Квантовые состояния

Состояние квантово-механической системы, представляющее собой статистический ансамбль одинаково приготовленных экземпляров системы, описывается *оператором плотности* (матрицей плотности в фиксированном базисе), т.е. оператором S в $\mathcal{H}$, удовлетворяющим условиям $S\geq 0$, $\operatorname{Tr}S=1$. Пусть $\mathcal{S}(\mathcal{H})$ — выпуклое множество всех операторов плотности. Выпуклая комбинация операторов плотности описывает смешивание соответствующих статистических ансамблей. Смесь $S=pS_1+(1-p)S_2$ получается, если взять ансамбли систем, приготовленных в состояниях S_1 и S_2 и смешать их в пропорции $p:1-p$.

В выпуклых множествах особо важны *крайние точки*, не представимые в виде нетривиальной смеси других точек, т.е. $S=pS_1+(1-p)S_2$, $0<p<1$, влечет $S=S_1=S_2$. С точки зрения статистической интерпретации, крайние точки множества состояний, называемые *чистыми состояниями*, соответствуют процедурам приготовления без участия классической случайности. В классической модели они, очевидно, описываются вырожденными распределениями, сосредоточенными в одной из точек фазового пространства. Соответствующие диагональные матрицы являются (одномерными) проекторами, $P^2=P$. Отметим также, что классическому равномерному распределению $P=\{1/d,\ldots,1/d\}$ соответствует квантовое *хаотическое* состояние $S=1/d\,I$.

В квантовом статистическом ансамбле есть два вида случайности: во-первых, устранимая в принципе случайность, обусловленная флуктуациями

классических параметров процедуры приготовления, и во-вторых, неуничтожимая квантовая случайность, присутствующая в любом чистом состоянии.

Теорема 2. *Крайние точки множества квантовых состояний* $\mathfrak{S}(\mathcal{H})$, *называемые чистыми состояниями, суть (одномерные) проекторы,* $S^2 = S$, *и только они.*

Доказательство. Рассмотрим спектральное разложение эрмитова оператора S

$$S = \sum_{i=1}^{d} s_i |e_i\rangle\langle e_i|, \quad s_j \geq 0, \sum s_j = 1, \tag{1.5}$$

где s_i – собственные значения, $|e_i\rangle$ – собственные векторы оператора S, $d = \dim \mathcal{H}$. Если S – крайняя точка, то эта сумма содержит только одно ненулевое слагаемое, следовательно, S есть одномерный проектор. Обратно, пусть S – одномерный проектор и $S = pS_1 + (1-p)S_2$, где $0 < p < 1$. Возведем это выражение в квадрат и рассмотрим разность S и S^2:

$$pS_1(I - S_1) + (1-p)S_2(I - S_2) + p(1-p)(S_1 - S_2)^2 = S - S^2 = 0. \tag{1.6}$$

Сумма трех положительных операторов равна нулю, следовательно, каждое слагаемое должно равняться нулю. Но это означает, что $S_1 = S_2 = S$, т. е. S – крайняя точка. □

Обозначим $\mathrm{Ext}(\mathcal{S})$ множество крайних точек произвольного выпуклого множества $\mathcal{S}$. Отметим следующий общий результат:

Теорема 3 (Каратеодори). *Пусть* $\mathcal{S}$ *— выпуклое компактное подмножество* n*-мерного векторного пространства, тогда любая точка* $S \in \mathcal{S}$ *может быть представлена в виде выпуклой комбинации (смеси) не более чем* $n+1$ *крайних точек*:

$$S = \sum_{j=1}^{n+1} p_j S_j, \quad S_j \in \mathrm{Ext}(\mathcal{S}).$$

В качестве примера рассмотрим выпуклое множество $\mathfrak{P}_d$ всех классических состояний – распределений вероятностей $P = \{p_1, \ldots, p_d\}$ на множестве из d элементов. В силу условия $\sum_\omega p_\omega = 1$, множество $\mathfrak{P}_d$ может быть погружено в $\mathbb{R}^n$, $n = d-1$. Его крайними точками являются вырожденные распределения, для которых вероятности p_ω равны нулю, за исключением одной, равной 1. Всего имеется d таких точек, и любое распределение из $\mathfrak{P}_d$ единственным образом представляется в виде их выпуклой комбинации с коэффициентами p_ω. Такое множество называется *симплексом*, и единственность представления является характеристическим свойством этого выпуклого множества.

Задача 2. Доказать, что если $\dim \mathcal{H} = d$, то $\mathcal{S}(\mathcal{H})$ погружается в вещественное пространство размерности $n = d^2 - 1$. Если же $\mathcal{H}$ евклидово (вещественное) пространство, то $n = d(d+1)/2 - 1$.

Спектральное разложение (1.5) показывает, что в случае множества квантовых состояний (как и для других выпуклых множеств с гладкой границей) теорема Каратеодори дает завышенное значение n. С другой стороны, для множества классических состояний, представляющего собой симплекс, эта теорема дает точное значение. Это наводит на мысль интерпретировать квантовую теорию как классическую вероятностную модель, в статистической структуре которой зашифрованы некие неклассические ограничения (теорию со скрытыми параметрами). Для одиночной квантовой системы такая точка зрения возможна, но до сих пор не оказалась плодотворной. При переходе же к составным системам она приводит к неустранимым противоречиям с физическими принципами локальности и причинности (см. далее п. 2.4).

1.4 Двухуровневые системы

Простейшей классической системой является *бит* – система с двумя чистыми состояниями. Статистические состояния представляются диагональными матрицами

$$P = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}, \quad 0 \le p \le 1.$$

и множество всех классических состояний представляет собой единичный отрезок.

Наиболее простым, но важным примером квантовой системы является *q-бит* — двухуровневая квантовая система, $\dim \mathcal{H} = 2$. Будем использовать канонический базис: $|\uparrow\rangle = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$, $|\downarrow\rangle = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. Удобно ввести базис Паули в вещественном пространстве эрмитовых 2×2-матриц:

$$I = \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

В частности, всякий оператор плотности $S \in \mathcal{S}(\mathcal{H})$ представляется как

$$S(\vec{a}) = \frac{1}{2} \begin{bmatrix} 1+a_z & a_x - ia_y \\ a_x + ia_y & 1-a_z \end{bmatrix} = \frac{1}{2}(I + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z). \qquad (1.7)$$

Условие $\det S \ge 0$ накладывает следующее ограничение на *параметры Стокса* (a_x, a_y, a_z):

$$a_x^2 + a_y^2 + a_z^2 \le 1.$$

Таким образом, $\mathcal{S}(\mathcal{H})$ как выпуклое множество изоморфно единичному шару в $\mathbb{R}^3$.

Чистые состояния характеризуются условием $a_x^2 + a_y^2 + a_z^2 = 1$ и составляют *сферу Блоха*. Вводя углы Эйлера θ и ϕ так, что $a_z = \cos\theta$ и $a_x + ia_y = \sin\theta e^{i\phi}$, имеем $S(\vec{a}) = |\vec{a}\rangle\langle\vec{a}|$, где $\vec{a} = (a_x, a_y, a_z)$ и

$$|\vec{a}\rangle = \begin{bmatrix} \cos(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} \end{bmatrix}. \qquad (1.8)$$

Таким образом,

$$|\vec{a}\rangle\langle\vec{a}| = \frac{1}{2}(I + \sigma(\vec{a})), \tag{1.9}$$

где

$$\sigma(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z \tag{1.10}$$

эрмитов унитарный оператор со свойствами $\sigma(\vec{a})^2 = I$, $\operatorname{Tr}\sigma(\vec{a}) = 0$. В квантовых системах со спином $1/2$ вектор $\vec{a}$ описывает ансамбль (пучок частиц) со спином в направлении $\vec{a}$ [1]. Хаотическим является смешанное состояние с $a_x = a_y = a_z = 0$ (все направления спинов равновероятны), описываемое оператором плотности $S = I/2$.

Из (1.9) получается спектральное разложение эрмитова оператора

$$\sigma(\vec{a}) = |\vec{a}\rangle\langle\vec{a}| - |\vec{-a}\rangle\langle\vec{-a}|. \tag{1.11}$$

Собственные векторы $|\pm\vec{a}\rangle$, отвечающие собственным значениям ± 1, образуют о.н.б. Наблюдаемая (1.11), принимающая значения ± 1, описывает проекцию спина на направление $\vec{a}$. Таким образом, спин электрона – это векторный оператор с некоммутирующими компонентами $\sigma_x, \sigma_y, \sigma_z$. Эксперимент Штерна-Герлаха, описывающий приготовление состояний $S(\vec{a})$ и измерение наблюдаемых $\sigma(\vec{b})$, детально описан в лекциях Фейнмана [7].

З а д а ч а 3. Пользуясь свойствами матриц Паули, покажите, что математическое ожидание наблюдаемой $\sigma(\vec{b})$, $|\vec{b}| = 1$ в состоянии $S(\vec{a})$, $|\vec{a}| \leq 1$ равно

$$\operatorname{Tr} S(\vec{a})\sigma(\vec{b}) = \vec{a}\cdot\vec{b} \tag{1.12}$$

1.5 Статистический анализ понятия "наблюдаемая"

Во всяком физическом эксперименте присутствуют две основные стадии: приготовление состояния и измерение (наблюдаемых величин). Даже если приготовляется чистое квантовое состояние, где нет классической стохастичности, результат измерения в данном ансамбле все равно может быть случаен. Итак, мы измеряем случайную величину, распределение $\mu_S^M(x)$ которой зависит от приготовления ансамбля S и от измерительного прибора M. Естественно ожидать, что смешивание ансамблей приводит к такому же смешиванию распределений, т.е. если $S = \sum_j p_j S_j$, то $\mu_S^M(x) = \sum_j p_j \mu_{S_j}^M(x)$. Другими словами, вероятности исходов измерения должны быть *аффинными* функциями состояния. Этого на первый взгляд слабого ограничения оказывается достаточно для вывода обобщенной статистической формулы Борна.

[1] Физическими реализациями q-бита являются спин электрона, поляризация фотона, атом с двумя активными уровнями.

Теорема 4[9]. *Пусть $S \to \mu_S$ отображение множества квантовых состояний $\mathcal{S}(\mathcal{H})$ в вероятностные распределения на некотором конечном множестве исходов $\mathcal{X}$. Если отображение аффинно, то существует семейство эрмитовых операторов $\{M_x\}$ в $\mathcal{H}$ такое, что*

$$M_x \geq 0, \quad \sum_{x\in\mathcal{X}} M_x = I \tag{1.13}$$

и

$$\mu_S(x) = \operatorname{Tr} SM_x. \tag{1.14}$$

Семейство, обладающее свойствами (1.13), называется *разложением единицы*[2] в $\mathcal{H}$.

Набросок доказательства. Эрмитов оператор $A = A^*$ в $\mathcal{H}$ имеет (неединственное) представление $A = t_1S_1 - t_2S_2$, где $t_i \geq 0$, а $S_i \in \mathcal{S}(\mathcal{H})$. В самом деле, существование представления $A = A_+ - A_-$ ($A_\pm \geq 0$) вытекает из спектрального разложения оператора A, а далее

$$A = \operatorname{Tr} A_1 \frac{A_1}{\operatorname{Tr} A_1} - \operatorname{Tr} A_2 \frac{A_2}{\operatorname{Tr} A_2},$$

т.е. линейная оболочка $\operatorname{Lin}\mathcal{S}(\mathcal{H})$ совпадает с вещественным линейным пространством $\mathcal{B}(\mathcal{H})_h$ всех эрмитовых операторов в $\mathcal{H}$. Пусть $f(S)$ — аффинная функция на $\mathcal{S}(\mathcal{H})$, продолжим ее на эрмитовы операторы A, полагая $f(A) = t_1f(S_1) - t_2f(S_2)$. Надо проверить, что благодаря аффинности, такое продолжение однозначно и вещественно линейно (задача 4). Далее, f однозначно и комплексно линейно продолжается на алгебру всех операторов $\mathcal{B}(\mathcal{H})$.

Следовательно, если $A = [a_{ij}]$ в некотором базисе, то $f(A) = \sum_{ij} m_{ij}a_{ij} = \operatorname{Tr} AM$, где M — некоторый оператор. Применяя это к функциям $\mu_S(x)$, получаем $\mu_S(x) = \operatorname{Tr} SM_x$. Поскольку $\operatorname{Tr} SM_x$ — распределение вероятностей для всех S, то отсюда следуют соотношения (1.13). □

Основываясь на этом результате, в дальнейшем примем следующее

Определение. *Квантовой наблюдаемой* со значениями в $\mathcal{X}$ называется разложение единицы $M = \{M_x\}_{x\in\mathcal{X}}$ в гильбертовом пространстве системы $\mathcal{H}$.

В стандартных текстах по квантовой механике под наблюдаемой понимают ортогональное разложение единицы, т. е. такое что

$$M_x^2 = M_x, \quad M_xM_y = 0, \quad x \neq y.$$

Задача 5. Ортогональное разложение единицы характеризуется свойством: все M_x — проекторы: $M_x = M_x^2$.

[2]Другое название: вероятностная (положительная) операторно-значная мера (POVM).

Ортогональное разложение единицы в $\mathcal{H}$ мы будем называть *четкой наблюдаемой* .

Пусть $x \in \mathcal{X} \subset \mathbb{R}$ вещественные числа. Тогда четкой наблюдаемой однозначно сопоставляется эрмитов оператор

$$\sum_{x\in\mathcal{X}} xE_x = X.$$

В силу взаимной однозначности соответствия, его также будем называть вещественной четкой наблюдаемой. Математическое ожидание такой наблюдаемой дается обычной формулой Борна (1.1), тогда как распределение вероятностей – формулой

$$\mu_S(x) = \operatorname{Tr} SE_x.$$

Чтобы пояснить используемую терминологию, а также статистический смысл неортогональных разложений единицы, рассмотрим о.н.б. $\{|\omega\rangle\}$ в $\mathcal{H}$ и операторы, диагональные в этом базисе. Оператор плотности

$$S = \sum_\omega s_\omega |\omega\rangle\langle\omega|, \quad s_\omega \geq 0, \quad \sum s_\omega = 1$$

задает классическое состояние — распределение вероятностей на "фазовом пространстве" $\Omega = \{\omega\}$. Эрмитов оператор $X = \sum\limits_\omega x_\omega |\omega\rangle\langle\omega|$ может быть записан в виде

$$X = \sum_x xE_x, \qquad E_x = \sum_{\omega: x_\omega = x} |\omega\rangle\langle\omega|.$$

Классическим наблюдаемым X соответствуют случайные величины x_ω на Ω. Проекторам E_x отвечают индикаторы подмножеств Ω, на которых $x_\omega = x$, а ортогональному разложению единицы — разбиение пространства Ω.

Рассмотрим неортогональное разложение единицы с элементами $M_x = \sum\limits_\omega M(x|\omega)|\omega\rangle\langle\omega|$. Тогда собственные числа удовлетворяют условиям $0 \leq M(x|\omega) \leq 1$ и

$$\sum_x M(x|\omega) = 1, \quad \omega \in \Omega, \tag{1.15}$$

т.е. определяют переходные вероятности из Ω в $\mathcal{X}$. Таким образом, в классическом случае разложения единицы описывают рандомизованные ("нечеткие") наблюдаемые, задающие распределение вероятностей исходов x в каждой точке ω фазового пространства. Для четких наблюдаемых, удовлетворяющих условию $M_x^2 = M_x$, эти вероятности принимают значения 0 или 1.

1.6 Смеси наблюдаемых. Экстремальные наблюдаемые

Пусть $\{M^j\}$ – семейство наблюдаемых с одним и тем же множеством исходов $\mathcal{X}$. Для данного распределения вероятностей $\{p_j\}$ можно естественным

образом определить *смесь* $M = \{M_x; x \in \mathcal{X}\}$ этих наблюдаемых по формуле

$$M_x = \sum_j p_j M_x^j; \quad x \in \mathcal{X}.$$

Таким образом, множество $\mathfrak{M}_\mathcal{X}$ всех наблюдаемых с заданным пространством исходов $\mathcal{X}$ становится выпуклым множеством. Аналогично смесям состояний, смеси наблюдаемых описывают измерения с флуктуирующими классическими параметрами. Крайние точки выпуклого множества наблюдаемых $\mathfrak{M}_\mathcal{X}$ будем называть *экстремальными* наблюдаемыми. Подобно чистым состояниям, они описывают статистику "чистых" измерений, свободную от классической случайности.

Следующий результат описывает нетривиальное соотношение между такими наблюдаемыми без классической случайности и четкими наблюдаемыми.

Теорема 5.*Всякая четкая наблюдаемая $M \in \mathfrak{M}_\mathcal{X}$ экстремальна. Обратно, всякая экстремальная наблюдаемая $M \in \mathfrak{M}_\mathcal{X}$ с коммутирующими компонентами, $[M_x, M_{x'}] \equiv 0$, является четкой наблюдаемой.*

Доказательство. Пусть M – четкая наблюдаемая. Предположим, что $M = pM^1 + (1-p)M^2$, $0 < p < 1$. Тогда, аналогично (1.6)

$$pM_x^1(I - M_x^1) + (1-p)M_x^2(I - M_x^2) + p(1-p)(M_x^1 - M_x^2)^2 = 0. \qquad (1.16)$$

откуда $M_x^1 \equiv M_x^2 \equiv M_x$, и M – крайняя точка.

Пусть теперь $[M_x, M_{x'}] \equiv 0$, тогда по теореме 1 M_x одновременно диагонализуемы, и можно считать, что $M_x = \mathrm{diag}[M(x|\omega)]$. Покажем, что если M экстремальная наблюдаемая, то $M(x|\omega) = 0$ или 1 для всех x, ω, т.е. M – четкая наблюдаемая. Пусть $0 < M(x_0|\omega_0) < 1$, тогда в силу условия (1.15) найдется $x_1 \neq x_0$, такой что $0 < M(x_1|\omega_0) < 1$. Определим две новые наблюдаемые $M^\pm$, полагая $M^\pm(x_0|\omega_0) = M(x_0|\omega_0) \pm \epsilon$, $M^\pm(x_1|\omega_0) = M(x_1|\omega_0) \mp \epsilon$ и оставляя прочие $M(x|\omega)$ без изменения. Тогда $M = 1/2M^+ + 1/2M^-$, т.е. M не экстремальна. □

Из доказанной теоремы следует, что в классическом случае экстремальные наблюдаемые совпадают с четкими, что дает им понятную характеризацию как наблюдаемых без случайности в процедуре измерения. В квантовой статистической модели все не так просто. Множество крайних точек квантовых наблюдаемых исчерпывается четкими наблюдаемыми только в случае двух исходов измерения (они играют особую роль в различных аксиоматических подходах; мы будем называть их *тестами*). Это следует из теоремы, так как любой тест имеет коммутирующие компоненты $\{M_0, M_1 = I - M_0\}$. Таким образом, любой экстремальный тест вполне определяется проектором $P = M_0$.

Однако в случае более чем двух исходов, $|\mathcal{X}| > 2$, всегда существуют нечеткие экстремальные квантовые наблюдаемые! Наиболее интересный класс будет рассмотрен в следующем разделе.

1.7 Переполненные системы векторов

Система векторов $\{|\psi_i\rangle\} \subset \mathcal{H}$ называется *переполненной*, если

$$\sum_j |\psi_i\rangle\langle\psi_j| = I.$$

Очевидным примером является всякий ортонормированный базис. В общем случае векторы могут быть ненормированными и линейно зависимыми. Тем не менее имеет место представление (вообще говоря, неоднозначное) векторов и операторов через переполненную систему, именно

$$|\psi\rangle = \sum_j |\psi_j\rangle\langle\psi_j|\psi\rangle,$$

$$A = \sum_j |\psi_j\rangle\langle\psi_j| A \sum_k |\psi_k\rangle\langle\psi_k| = \sum_{j,k} |\psi_j\rangle\langle\psi_k|\langle\psi_j|A|\psi_k\rangle.$$

З а д а ч а 6. Система $\{|\psi_j\rangle\}$ является переполненной тогда и только тогда, когда
1) система полна, т.е. $\{|\psi_j\rangle\}^\perp = \{0\}$;
2) матрица $P = [\langle\psi_j|\psi_k\rangle]$ идемпотентна, т.е. $P = P^2$.

Пусть $\{|\phi_j\rangle\}$ — произвольная полная (не обязательно ортонормированная) система векторов. Тогда *оператор Грама*

$$G = \sum_j |\phi_j\rangle\langle\phi_j|$$

невырожден. Система векторов $|\psi_j\rangle = G^{-1/2}|\phi_j\rangle$ является переполненной.

Переполненная система в подпространстве гильбертова пространства возникает при проецировании ортонормированного базиса пространства на подпространство. Более того, далее будет доказана теорема, принадлежащая М.А. Наймарку, из которой следует, что *всякая* переполненная система получается таким образом.

С каждой переполненной системой связано разложение единицы, т.е. наблюдаемая

$$M_j = |\psi_j\rangle\langle\psi_j|. \tag{1.17}$$

В частности, для любой полной системы $\{|\phi_j\rangle\}$ набор операторов

$$M_j = G^{-1/2}|\phi_j\rangle\langle\phi_j|G^{-1/2} \tag{1.18}$$

задает наблюдаемую, в определенном смысле "измеряющую" состояния $|\phi_j\rangle\langle\phi_j|$[3].

Т е о р е м а 6. *Наблюдаемая (1.17) является экстремальной тогда и только тогда, когда операторы M_x линейно независимы.*

[3] А. С. Холево, Об асимптотически оптимальном различении гипотез в квантовой статистике. ТВП, 1978, т.23, N2, 429-432. В англоязычной литературе это называется *square-root measurement*.

Доказательство. Пусть M – крайняя точка и предположим, что

$$\sum_x c_x|\psi_x\rangle\langle\psi_x| = 0. \tag{1.19}$$

Взяв достаточно малое $\epsilon > 0$, определим

$$M_x^{\pm} = (1 \pm \epsilon c_x)M_x \geq 0, \quad x \in \mathcal{X}.$$

Тогда $M^{\pm}$ являются наблюдаемыми и, по построению, $M = \frac{1}{2}M^+ + \frac{1}{2}M^-$. Но M – крайняя точка, значит, $M_x^+ = M_x^- = M_x$. Итак, из (1.19) следует $c_x = 0$, т. е. компоненты M линейно независимы.

Обратно, пусть

$$|\psi_x\rangle\langle\psi_x| = pM_x^1 + (1-p)M_x^2, \quad 0 < p < 1,$$

– разложение M в смесь, тогда

$$0 \leq pM_x^1 \leq |\psi_x\rangle\langle\psi_x|.$$

Умножая справа и слева на эрмитов оператор $I - \frac{|\psi_x\rangle\langle\psi_x|}{\langle\psi_x|\psi_x\rangle}$, получаем

$$(I - \frac{|\psi_x\rangle\langle\psi_x|}{\langle\psi_x|\psi_x\rangle})M_x^1(I - \frac{|\psi_x\rangle\langle\psi_x|}{\langle\psi_x|\psi_x\rangle}) = 0,$$

откуда

$$\sqrt{M_x^1}(I - \frac{|\psi_x\rangle\langle\psi_x|}{\langle\psi_x|\psi_x\rangle}) = 0,$$

и поэтому

$$M_x^1(I - \frac{|\psi_x\rangle\langle\psi_x|}{\langle\psi_x|\psi_x\rangle}) = 0.$$

Отсюда получаем $M_x^1 = \lambda_x|\psi_x\rangle\langle\psi_x|$ с $\lambda_x = \langle\psi_x|M_x^1|\psi_x\rangle/\langle\psi_x|\psi_x\rangle^2$. Тогда $\sum_x \lambda_x|\psi_x\rangle\langle\psi_x| = I$, т. е.

$$\sum_x (\lambda_x - 1)|\psi_x\rangle\langle\psi_x| = 0.$$

В силу линейной независимости, $\lambda_x = 1$, и $M_x^1 = M_x$ для всех x, следовательно, M – крайняя точка. □

1.8 Переполненные системы для q-бита

Теорема 7. *Пусть $\vec{a}_j$; $j = 1, \ldots, m$ система единичных векторов в $\mathbb{R}^3$, такая что $\sum_{j=1}^m \vec{a}_j = 0$. Тогда векторы $\sqrt{2/m}|\vec{a}_j\rangle$; $j = 1, \ldots, m$ образуют переполненную систему в пространстве q-бита $\mathcal{H}$, так что*

$$\frac{2}{m}\sum_{j=1}^m |\vec{a}_j\rangle\langle\vec{a}_j| = I. \tag{1.20}$$

Соответствующая наблюдаемая экстремальна тогда и только тогда, когда векторы $\vec{a}_j - \vec{a}_1$; $j = 2, \ldots, m$ *линейно независимы.*

Доказательство. Первое утверждение, т.е. соотношение (1.20), непосредственно следует из (1.9). Также используя это соотношение получаем, что линейная зависимость операторов $|\vec{a}_j\rangle\langle\vec{a}_j|$, равносильная, в силу теоремы 6, неэкстремальности наблюдаемой, означает, что

$$0 = \sum_{j=1}^{m} c_j |\vec{a}_j\rangle\langle\vec{a}_j| = \frac{1}{2}\left[\sum_{j=1}^{m} c_j I + \sigma(\sum_{j=1}^{m} c_j \vec{a}_j)\right].$$

Отсюда

$$\sum_{j=1}^{m} c_j = 0, \quad \sum_{j=1}^{m} c_j \vec{a}_j = 0$$

или

$$\sum_{j=2}^{m} c_j (\vec{a}_j - \vec{a}_1) = 0.$$

□

Примеры *симметричных* систем единичных векторов в $\mathbb{R}^3$, удовлетворяющих условиям теоремы, а также соответствующие им переполненные системы и наблюдаемые приведены ниже.

$m = 2 : \vec{a}_{1,2} = (0, 0, \pm 1)$. В этом случае имеем о.н.б. $|\vec{a}_1\rangle = |\uparrow\rangle, |\vec{a}_2\rangle = |\downarrow\rangle$ и ортогональное разложение единицы

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, M_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

$m = 3$: равноугольная конфигурация трех векторов в вещественной плоскости $\vec{a}_1 = (0, 0, 1), \vec{a}_{2,3} = (\pm\sqrt{3}/2, 0, -1/2)$ ("Мерседес-Бенц"). Соответствующая переполненная система в $\mathcal{H}$

$$|\psi_1\rangle = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |\psi_{2,3}\rangle = \sqrt{\frac{2}{3}} \begin{bmatrix} 1/2 \\ \pm\sqrt{3}/2 \end{bmatrix}$$

и неортогональное разложение единицы

$$M_1 = \frac{2}{3} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_{2,3} = \frac{2}{3} \begin{bmatrix} 1/4 & \pm\sqrt{3}/4 \\ \pm\sqrt{3}/4 & 3/4 \end{bmatrix}.$$

$m = 4$: конфигурация тетраэдра $\vec{a}_1 = (0, 0, 1), \vec{a}_2 = (\sqrt{8}/3, 0, -1/3), \vec{a}_{2,3} = (-\sqrt{2}/3, \pm\sqrt{6}/3, -1/3)$. Соответствующая переполненная система в $\mathcal{H}$

$$|\psi_1\rangle = \sqrt{\frac{1}{2}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\psi_2\rangle = \sqrt{\frac{1}{2}} \begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}, \quad |\psi_{3,4}\rangle = \sqrt{\frac{1}{2}} \begin{bmatrix} 1/\sqrt{3}\,(-1/2 \mp i\sqrt{3}/2) \\ \sqrt{2}/\sqrt{3}\,(-1/2 \pm i\sqrt{3}/2) \end{bmatrix}$$

и неортогональное разложение единицы

$$M_1 = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_2 = \frac{1}{2}\begin{bmatrix} 1/3 & \sqrt{2}/3 \\ \sqrt{2}/3 & 2/3 \end{bmatrix}, \quad M_{2,3} = \frac{1}{2}\begin{bmatrix} 1/3 & \sqrt{2}/3\left(-1/2 \mp i\sqrt{3}/2\right) \\ c.c. & 3/4 \end{bmatrix}.$$

В случаях $m = 3, 4$ получаем нечеткие экстремальные наблюдаемые. Такого рода наблюдаемые не имеют аналога в классической статистике.

1.9 Томография квантового состояния

В последнем случае $m = 4 = d^2$, поэтому линейно независимые операторы $M_j; j = 1, 2, 3, 4$ образуют базис в пространстве эрмитовых операторов. Таким образом, вероятности

$$\mu_S(j) = \operatorname{Tr} SM_j = \langle\psi_j|S|\psi_j\rangle$$

однозначно определяют состояние S. В общем случае, наблюдаемая в $\mathcal{H}$, $\dim\mathcal{H} = d$, обладающая таким свойством, называется *информационно-полной*. Экстремальная наблюдаемая вида

$$M_j = d^{-1}|\psi_j\rangle\langle\psi_j|; j = 1, \ldots, d^2,$$

где $|\psi_j\rangle$ – единичные векторы, называется *симметричной информационно-полной* (SIC-POVM), если

$$\operatorname{Tr} M_j M_k = \text{const},$$

причем константа оказывается равной $[d^2(d+1)]^{-1}$. Существование SIC-POVM показано аналитически, либо численно, для $d \leq 67$. Имеется гипотеза, что они существуют во всех размерностях [4].

З а д а ч а 7. Покажите, что любое состояние S восстанавливается по формуле

$$S = \sum_{j=1}^{d^2}[d(d+1)\mu_S(j) - 1]M_j.$$

Восстановление состояния по статистике измерений (одного или целого ряда) называют *томографией* квантового состояния. Например, формула (1.7) показывает, что состояние q-бита восстанавливается по средним значениям компонент спина $a_x = \operatorname{Tr} S\sigma_x, a_y = \operatorname{Tr} S\sigma_y, a_z = \operatorname{Tr} S\sigma_z$. Тем более, это позволяют сделать вероятности для 3-х ортонормированных базисов операторов $\sigma_x, \sigma_y, \sigma_z$. Эти базисы обладают свойством "равнонаклоненности" (mutual unbiasedness):

$$|\langle e_j|h_k\rangle|^2 = \text{const} \tag{1.21}$$

для всех j, k. В общем случае, базисы в $\mathcal{H}$, $\dim\mathcal{H} = d$, обладающие таким свойством, называются *равнонаклоненными* (MUB). Доказывается, что количество попарно равнонаклоненных базисов не превосходит $d+1$, причем

[4] http://en.wikipedia.org/wiki/SIC-POVM

константа равна $1/d$. Существование $d+1$ равнонаклоненных базисов доказано для размерностей вида p^k, где p – простое число; имеется гипотеза, что в других размерностях они не существуют. Измерения в равнонаклоненных базисах удобны для томографии квантовых состояний.

1.10 Теорема Наймарка

Геометрический смысл неортогональных разложений единицы проясняет следующая теорема.

Теорема 8. *Пусть $\{M_x\}_{x\in\mathcal{X}}$ — разложение единицы в гильбертовом пространстве $\mathcal{H}$, $\dim\mathcal{H}=d$, $|\mathcal{X}|=n$. Существует гильбертово пространство $\widetilde{\mathcal{H}}$, $\dim\widetilde{\mathcal{H}}\le n\cdot d$, изометрический оператор $V:\mathcal{H}\to\widetilde{\mathcal{H}}$ и ортогональное разложение единицы $\{E_x\}$ в $\widetilde{\mathcal{H}}$, такие, что*

$$M_x = V^*E_xV.$$

Изометрический оператор — это оператор, сохраняющий скалярное произведение, следовательно все углы, расстояния и объем. Для любых $|\phi\rangle,|\psi\rangle\in\mathcal{H}$ выполняется $\langle\phi|V^*V|\psi\rangle=\langle\phi|\psi\rangle$, т.е. $V^*V=I$. Изометрическое вложение V позволяет отождествить $\mathcal{H}$ с подпространством $V\mathcal{H}$ пространства $\widetilde{\mathcal{H}}$ и считать, что $\mathcal{H}\subset\widetilde{\mathcal{H}}$. Тогда M_x можно рассматривать просто как ограничение E_x на $\mathcal{H}$:

$$E_x=\begin{bmatrix} M_x & \cdots \\ \cdots & \cdots \end{bmatrix}$$

Заметим, что теорема имеет место и в случае общего разложения единицы в бесконечномерном гильбертовом пространстве.

Набросок доказательства. Рассмотрим векторную сумму $\mathcal{H}_n$ n копий пространства $\mathcal{H}$, состоящую из векторов

$$|\Psi\rangle=\begin{bmatrix} |\psi_1\rangle \\ \cdots \\ |\psi_n\rangle \end{bmatrix},\quad \psi_j\in\mathcal{H},$$

в которой определим псевдоскалярное произведение формулой

$$\langle\Psi|\Psi'\rangle=\sum_x\langle\psi_x|M_x|\psi'_x\rangle.$$

Соответствующая квадратичная форма может быть вырождена. Обозначим $\mathcal{H}_0=\{\Psi\in\mathcal{H}_n:\langle\Psi|\Psi\rangle=0\}$ и рассмотрим фактор-пространство $\mathcal{H}_n/\mathcal{H}_0$. В нем определено настоящее скалярное определение. Это и будет $\widetilde{\mathcal{H}}$. (Заметим, что размерность $n\cdot d$ пространства $\mathcal{H}_n$ могла лишь уменьшиться при факторизации). Определим

$$V|\psi\rangle=\begin{bmatrix} |\psi\rangle \\ \cdots \\ |\psi\rangle \end{bmatrix}\equiv|\Psi\rangle.$$

З а д а ч а 8. После факторизации эта формула корректно определяет оператор V из $\mathcal{H}$ в $\widetilde{\mathcal{H}}$.

Этот оператор изометричен, т.к.

$$\langle\psi|V^*V\psi'\rangle = \sum_x \langle\psi|M_x|\psi'\rangle = \langle\psi|\psi\rangle,$$

поскольку $\sum M_x = I$. Теперь введем ортогональное разложение единицы, полагая в $\mathcal{H}_n$

$$E_y|\Psi\rangle = \begin{bmatrix} 0 \\ |\psi_y\rangle \\ 0 \end{bmatrix}.$$

При этом $\langle\psi|V^*E_yV|\psi'\rangle = \langle\psi|M_y|\psi'\rangle$. □

Из этой теоремы следует, что всякая переполненная системы является проекцией ортонормированного базиса. Далее с ее помощью будет прояснен статистический смысл нечетких наблюдаемых.

Глава 2

Составные квантовые системы

2.1 Наводящие соображения

Рассмотрим две классические системы, с фазовыми пространствами Ω_1, Ω_2 в статистических состояниях $P_1 = \{p_i\}, P_2 = \{q_j\}$, соответственно. Фазовым пространством составной системы является *декартово произведение* фазовых пространств подсистем $\Omega_1 \times \Omega_2$. Состояние составной системы, в котором эти подсистемы рассматриваются как независимые, описывается произведением распределений $P_1 \times P_2 = \{p_i q_j\}$. Коррелированные подсистемы описываются совместным распределением $P_{12} = \{p_{ij}\}$, при этом *маргинальные распределения* подсистем даются частичными суммами

$$p_i = \sum_j p_{ij}, \quad q_j = \sum_i p_{ij}.$$

Пусть теперь состояния двух квантовых систем описываются матрицами плотности: $S_1 = [s_{ik}], S_2 = [r_{jl}]$. Тогда состояние составной системы, в котором эти подсистемы рассматриваются как независимые, описывается *тензорным произведением* матриц $S_1 \otimes S_2 = [s_{ik} r_{jl}]$. Коррелированные квантовые системы описываются произвольными матрицами с составными индексами: $S_{12} = \left[s_{(ij)(kl)}\right]$. При этом *частичные состояния* подсистем даются частичными следами

$$S_1 = \left[\sum_k s_{ik} r_{jk}\right], \quad S_2 = \left[\sum_i s_{ij} r_{ik}\right]. \tag{2.1}$$

Понятие квантовой *сцепленности*[1] возникает уже при рассмотрении чистых состояний. Для классических систем чистые состояния исчерпываются

[1] Англ. “entanglement”, в российской физической литературе переводится как “запутанность”.

распределениями, вырожденными в точках фазового пространства, поэтому если составная система находится в чистом состоянии, то ее подсистемы также находятся в чистых состояниях $P_{12} = \delta_i \times \delta_j$.

Чистое состояние квантовой системы является одномерным проектором, т.е. $S_{12} = [c_{ij}\bar{c}_{kl}]$. Если $c_{ij} \neq c_i c_j$, то состояние *сцепленное*. В этом случае частичные состояния S_1, S_2, полученные по формуле (2.1) уже *не чистые*. Выходит так, что статистичность в каждой из подсистем возникает из ее "окружения"!

Для более детального рассмотрения нам понадобится соответствующий математический аппарат.

2.2 Тензорное произведение гильбертовых пространств

Своеобразие и необычные возможности квантовой теории информации в значительной мере обусловлены свойствами составных квантовых систем. Пусть $\mathcal{H}_i$ $(i = 1, 2)$ гильбертовы пространства двух квантовых систем со скалярными произведениями $\langle\cdot|\cdot\rangle_i$. Их совокупность описывается тензорным произведением гильбертовых пространств, которое строится следующим образом. Пусть задано билинейное отображение

$$|\psi_1\rangle,\, |\psi_2\rangle \longrightarrow |\psi_1\rangle \otimes |\psi_2\rangle \equiv |\psi_1 \otimes \psi_2\rangle \tag{2.2}$$

пары пространств $\mathcal{H}_i$ $(i = 1, 2)$ в некоторое гильбертово пространство $\mathcal{H}$, причем

1. векторы-произведения $|\psi_1 \otimes \psi_2\rangle$ линейно порождают $\mathcal{H}$;

2. скалярное произведение на порождающих элементах дается соотношением
$$\langle\phi_1 \otimes \phi_2|\psi_1 \otimes \psi_2\rangle = \langle\phi_1|\psi_1\rangle_1 \langle\phi_2|\psi_2\rangle_2.$$

Данными требованиями пространство $\mathcal{H}$ определяется однозначно с точностью до унитарной эквивалентности, и называется *тензорным произведением* $\mathcal{H}_1 \otimes \mathcal{H}_2$ гильбертовых пространств.

Задача 9. Пусть $\{e_1^j\}, \{e_2^k\}$ — ортонормированные базисы в $\mathcal{H}_1, \mathcal{H}_2$, тогда $\{e_1^j \otimes e_2^k\}$ — ортонормированный базис в $\mathcal{H}_1 \otimes \mathcal{H}_2$ и $\dim \mathcal{H} = \dim \mathcal{H}_1 \cdot \dim \mathcal{H}_2$.

Например, для системы из двух кубитов о.н.б. является

$$|\uparrow\uparrow\rangle = |\uparrow\rangle_1 \otimes |\uparrow\rangle_2,\, |\uparrow\downarrow\rangle = |\uparrow\rangle_1 \otimes |\downarrow\rangle_2,\, |\downarrow\uparrow\rangle = |\downarrow\rangle_1 \otimes |\uparrow\rangle_2,\, |\downarrow\downarrow\rangle = |\downarrow\rangle_1 \otimes |\downarrow\rangle_2. \tag{2.3}$$

Таким образом, реализуя $\mathcal{H}_{1,2}$ как пространства $\mathbb{C}^{d_{1,2}}$ числовых последовательностей $\{c_j^1\}, \{c_k^2\}$, получим реализацию $\mathcal{H}$ в виде пространства матриц

$[c_{jk}]$. Заметим, что всякий вектор $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ однозначно записывается в виде

$$|\psi\rangle = \sum_{k=1}^{d_2} |\psi_k\rangle \otimes |e_k^2\rangle,$$

другими словами

$$|\psi\rangle = \begin{bmatrix} |\psi_1\rangle \\ \dots \\ |\psi_{d_2}\rangle \end{bmatrix}, \tag{2.4}$$

где компоненты $|\psi_k\rangle \in \mathcal{H}_1$, так что в общем случае $\mathcal{H}_1 \otimes \mathcal{H}_2$ изоморфно прямой сумме $d_2 = \dim \mathcal{H}_2$ слагаемых $\mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_1$.

Для операторов $X_{1,2}$ в пространствах $\mathcal{H}_{1,2}$ зададим их тензорное произведение в пространстве $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, полагая

$$(X_1 \otimes X_2)(\psi_1 \otimes \psi_2) = X_1\psi_1 \otimes X_2\psi_2,$$

и продолжая по линейности. В представлении (2.4) тензорного произведения $\mathcal{H}_1 \otimes \mathcal{H}_2$ произвольный оператор действует как блочная $d_2 \times d_2$-матрица $X = [X_{kl}]$, элементами которой являются операторы в $\mathcal{H}_1$.

З а д а ч а 10. Если S_j — операторы плотности в $\mathcal{H}_1$, то $S_1 \otimes S_2$ — оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Пусть оператор T действует в $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. *Частичный след* оператора T (по второму сомножителю) обозначим $\mathrm{Tr}_{\mathcal{H}_2} T$; это оператор в $\mathcal{H}_1$, ассоциированный с формой

$$\langle\phi|(\mathrm{Tr}_{\mathcal{H}_2} T)|\psi\rangle = \sum_k \langle \phi \otimes e_2^k | T | \psi \otimes e_2^k \rangle, \quad \phi, \psi \in \mathcal{H}.$$

З а д а ч а 11. Определение корректно (не зависит от выбора ортонормированного базиса $\{e_2^k\}$). Если $T = T_1 \otimes T_2$, то $\mathrm{Tr}_{\mathcal{H}_2}(T_1 \otimes T_2) = (\mathrm{Tr}\, T_2) T_1$.

Рассмотрим теперь важное следствие из теоремы Наймарка, дающее статистическую интерпретацию произвольного разложения единицы и устанавливающее согласованность обобщенного и стандартного определений квантовой наблюдаемой.

С л е д с т в и е . *Пусть $\{M_j\}$ — разложение единицы в $\mathcal{H}$, тогда найдется гильбертово пространство $\mathcal{H}_0$, единичный вектор $\psi_0 \in \mathcal{H}_0$ и ортогональное разложение единицы $\{E_j\}$ в $\mathcal{H} \otimes \mathcal{H}_0$, такие, что*

$$M_j = \mathrm{Tr}_{\mathcal{H}_0}(I \otimes |\psi_0\rangle\langle\psi_0|) E_j.$$

Доказательство. Согласно теореме Наймарка, $M_j = V^* \widetilde{E_j} V$, где $V : \mathcal{H} \to \widetilde{\mathcal{H}}$ — изометрическое вложение. Отождествим $\mathcal{H}$ с подпространством $\widetilde{\mathcal{H}}$. Расширяя, если необходимо, пространство $\widetilde{\mathcal{H}}$, можно считать, что $\dim \widetilde{\mathcal{H}} = \dim \mathcal{H} \cdot d_0$, и значит

$$\widetilde{\mathcal{H}} = \mathcal{H} \oplus \dots \oplus \mathcal{H} = \mathcal{H} \otimes \mathcal{H}_0,$$

где $\mathcal{H}_0 = \mathbb{C}^{d_0}$, причем $\mathcal{H}$ отождествляется с первым слагаемым в прямой сумме, или с подпространством $\mathcal{H} \otimes |\psi_0\rangle$, где

$$|\psi_0\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Имеем для $\phi, \psi \in \mathcal{H}$:

$$\langle\phi|M_j|\psi\rangle = \langle\phi \otimes \psi_0|E_j|\psi \otimes \psi_0\rangle = \langle\phi| \operatorname{Tr}_{\mathcal{H}_0}(I \otimes |\psi_0\rangle\langle\psi_0|)E_j|\psi\rangle.$$

□

Итак, всякую наблюдаемую можно реализовать в виде стандартной наблюдаемой в составной системе за счет добавления вспомогательной системы, находящейся в фиксированном чистом состоянии $S_0 = |\psi_0\rangle\langle\psi_0|$. Такой способ реализации естественно назвать *квантовой рандомизацией*.

В классической статистике рандомизация, т. е. использование внешней случайности (рулетки) при принятии решений, хотя и может оказаться полезным приемом (например, в теории игр), никогда не увеличивает информации о состоянии наблюдаемой системы. В разделе 4.2.2 мы покажем, что в квантовой механике это уже не так: парадоксальным образом, квантовая рандомизация позволяет извлекать больше информации о наблюдаемой системе, нежели содержится в стандартных наблюдаемых, не использующих вспомогательной системы.

2.3 Разложение Шмидта и очищение

Рассмотрим состояние S_{12} составной системы в гильбертовом пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2$. Чистое состояние S_{12} называется *сцепленным*, если оно не представимо в виде тензорного произведения $S_1 \otimes S_2$.

Таким образом, всякий единичный вектор $|\psi\rangle_{12} \in \mathcal{H}_1 \otimes \mathcal{H}_2$, который является нетривиальной суперпозицией векторов-произведений, порождает чистое сцепленное состояние. Примером является *максимально сцепленное состояние*, которое порождается вектором

$$|\psi_{12}\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |e_j^1\rangle \otimes |e_j^2\rangle \tag{2.5}$$

в пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2$, где $d = \dim \mathcal{H}_1 = \dim \mathcal{H}_2$, а $\{e_j^{1,2}\}$ – ортонормированные базисы в $\mathcal{H}_{1,2}$. Частичными состояниями в $\mathcal{H}_1, \mathcal{H}_2$ являются хаотические состояния I/d.

В квантовой теории информации часто используется следующий простой, но неожиданный результат[2]:

[2]См., например, G. Lindblad, "Quantum entropy and quantum measurements," *Lect. Notes Phys.* **378**, Quantum Aspects of Optical Communication, Ed. by C. Benjaballah, O. Hirota, S. Reynaud, 1991, 71-80.

Теорема 9*[Разложение Шмидта].Пусть* $S_{12} = |\psi\rangle\langle\psi|$ *– чистое состояние в гильбертовом пространстве* $\mathcal{H}_1 \otimes \mathcal{H}_2$*, и пусть* $S_1 = \mathrm{Tr}_{\mathcal{H}_2} S_{12}$*,* $S_2 = \mathrm{Tr}_{\mathcal{H}_1} S_{12}$ *– частичные состояния. Тогда* S_1 *и* S_2 *имеют одни и те же ненулевые собственные значения* λ_j*. Более того,*

$$|\psi\rangle = \sum_j \sqrt{\lambda_j} |e_j^1\rangle \otimes |e_j^2\rangle, \tag{2.6}$$

где $\{e_j^{1,2}\}$ *– ортонормированные собственные векторы операторов* S_1 *и* S_2 *соответственно.*

Доказательство. Пусть $\{e_j^1\}$ – ортонормированный базис в $\mathcal{H}_1$ из собственных векторов оператора S_1, тогда имеет место разложение

$$|\psi\rangle = \sum_j |e_j^1\rangle \otimes |h_j^2\rangle, \tag{2.7}$$

с некоторыми векторами $|h_j^2\rangle \in \mathcal{H}_2$. Вычисление частичного следа оператора $|\psi\rangle\langle\psi|$ по $\mathcal{H}_2$ дает

$$\sum_{j,k} \langle h_j^2 | h_k^2 \rangle |e_k^1\rangle\langle e_j^1| = \sum_j \lambda_j |e_j^1\rangle\langle e_j^1| \equiv S_1, \tag{2.8}$$

и поэтому $\langle h_j^2 | h_k^2 \rangle = \lambda_j \delta_{jk}$. Таким образом, полагая $|e_j^2\rangle = \frac{1}{\sqrt{\lambda_j}} |h_j^2\rangle$ при $\lambda_j > 0$, получаем ортонормированную систему, которую можно дополнить до базиса в $\mathcal{H}_2$, состоящего из собственных векторов оператора S_2.□

Имеет место следующее обращение предыдущего утверждения:

Теорема 10*[Очищение состояний]. Пусть* S_1 *– состояние в* $\mathcal{H}_1$*, тогда найдутся гильбертово пространство* $\mathcal{H}_2$ *той же размерности, что и* $\mathcal{H}_1$*, и чистое состояние* $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$*, такие, что* $S_1 = \mathrm{Tr}_{\mathcal{H}_2} |\psi\rangle\langle\psi|$.

Для любого чистого состояния $|\psi'\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$*, обладающего этим же свойством, найдется унитарный оператор* U_2 *в* $\mathcal{H}_2$*, такой, что* $|\psi'\rangle = (I_1 \otimes U_2)|\psi\rangle$.

Доказательство. Диагонализуем S_1 и определим $|\psi\rangle$ по формуле (2.6) с произвольным базисом $\{e_j^2\}$ в гильбертовом пространстве $\mathcal{H}_2$, изоморфном $\mathcal{H}_1$. Любой другой вектор $|\psi'\rangle$ имеет разложение (2.6) с другим базисом в $\mathcal{H}_2$. Остается заметить, что любые два базиса в гильбертовом пространстве связаны унитарным преобразованием. □

2.4 Парадокс ЭПР. Неравенство Белла

Ключевой пример необычного (с классической точки зрения) поведения составной квантовой системы рассмотрели Эйнштейн, Подольский и Розен (ЭПР) в 1935 г. В современной форме, использующей спиновые степени свободы, его представил Бом в 1950-х, а значительное прояснение внес Белл в

работах 1960-х годов. Рассмотрим составную систему из двух q-битов, например, две частицы со спином $1/2$, каждая из которых описывается гильбертовым пространством $\mathcal{H}$ с $\dim \mathcal{H} = 2$. В начальный момент частицы взаимодействуют таким образом, что конечное состояние их спинов, называемое *синглетным*, описывается вектором

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big[|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle\Big],$$

где векторы

$$|\uparrow\rangle = \begin{bmatrix}1\\0\end{bmatrix}, |\downarrow\rangle = \begin{bmatrix}0\\1\end{bmatrix}$$

описывают состояния каждой частицы со спином, направленным, соответственно, в положительном и отрицательном направлении оси z. Обычно пишут

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\Big].$$

Каждая из компонент описывает состояние с разнонаправленными спинами, а $|\psi\rangle$ — их суперпозиция, которую невозможно представить в виде произведения векторов состояний, относящихся к разным частицам. Синглетное состояние — канонический пример сцепленного состояния двух квантовых систем, т.е. состояния, не представимого в виде тензорного произведения чистых состояний.

Затем частицы разлетаются вдоль оси y на макроскопическое расстояние, а сцепленное спиновое состояние сохраняется. В частности, полный спин остается равным 0. Если теперь измерением спина фиксировать состояние первой частицы, то вторая частица оказывается в определенном состоянии с противоположным направлением спина. Таким образом, интерпретируя понятие квантового состояния, приходится выбирать между следующими альтернативами:

1) в квантовой механике, подобно классической, состояние описывает “реальные” внутренние свойства системы. При этом, чтобы объяснить, как вторая частица “узнает” о выборе направления измеряемого спина для первой частицы, приходится допустить мгновенное дальнодействие, противоречащее физическому “принципу локальности”;

2) вектор состояния – это лишь выражение информационного содержания процедуры приготовления системы, включающее прошлое взаимодействие подсистем. В этом случае никакого противоречия с локальностью не возникает, но приходится отказаться от полноты механистического описания состояния как “совокупности внутренних свойств”.

Внимательное рассмотрение этого мысленного эксперимента приводит к более глубокому и неожиданному выводу, на который обратил внимание Белл: если пытаться описывать корреляции измерений спинов двух частиц классически и в соответствии с принципом локальности, то оказывается невозможным достичь такого характера и уровня коррелированности, который соответствует предсказаниям квантовой механики. Более

того, этот уровень коррелированности может быть количественно сформулирован и проверен экспериментально. Дадим точную формулировку. Пусть вектор $\vec{a} = (a_x, a_y, a_z)$ задает некоторое направление, тогда $\sigma(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z$ — наблюдаемая спина в направлении $\vec{a}$. Оператор $\sigma(\vec{a})$ имеет собственные значения ± 1 (спин вдоль и против направления $\vec{a}$). Таким образом

$$\sigma(\vec{a}) = \overbrace{|\psi(\vec{a})\rangle\langle\psi(\vec{a})|}^{S(\vec{a})} - \overbrace{|\psi(-\vec{a})\rangle\langle\psi(-\vec{a})|}^{S(-\vec{a})}.$$

Напомним, что $\vec{a}$ имеет углы Эйлера (θ, ϕ), при этом вектор (1.8) отвечает чистому состоянию со спином в направлении $\vec{a}$. Соответствующий оператор плотности

$$S(\vec{a}) = \frac{I + \sigma(\vec{a})}{2}.$$

Рассмотрим эксперимент, в котором производятся совместные измерения наблюдаемой $\sigma(\vec{a})$ для одной системы и $\sigma(\vec{b})$ – для другой.

З а д а ч а 12. Для синглетного состояния двух q-битов корреляция спинов дается формулой

$$\langle\psi|\sigma(\vec{a}) \otimes \sigma(\vec{b})|\psi\rangle = -\vec{a} \cdot \vec{b}. \tag{2.9}$$

Оказывается, что такая корреляция не может быть смоделирована никакой классической моделью составной системы, удовлетворяющей принципу локальности. Это вытекает из следующего неравенства Клаузера–Хорна–Шимони–Хольта:

Пусть X_j, Y_k $(j, k = 1, 2)$ — случайные величины на произвольном вероятностном пространстве Ω, такие что $|X_j| \le 1$, $|Y_k| \le 1$. Тогда для любого распределения вероятностей на Ω корреляции этих величин удовлетворяют неравенству

$$|\mathsf{E}X_1Y_1 + \mathsf{E}X_1Y_2 + \mathsf{E}X_2Y_1 - \mathsf{E}X_2Y_2| \le 2, \tag{2.10}$$

где E — соответствующее математическое ожидание.

Доказательство получается усреднением элементарного неравенства

$$-2 \le X_1Y_1 + X_1Y_2 + X_2Y_1 - X_2Y_2 \le 2.$$

Принцип локальности, или, лучше сказать, *разделимости* в данной модели заключается в том, что физическая наблюдаемая для первой системы описывается одной и той же случайной величиной (X_1 в случае первых двух корреляций, X_2 в другом случае) независимо от того, какая величина — Y_1 или Y_2 измеряется во второй системе. Это условие кажется настолько естественным, что оно даже трудно уловимо. Однако именно оно запрещает мгновенное влияние измерения, проводящегося в одной системе, на измерения в другой системе. Если от него отказаться, то интересующие нас четыре физические корреляции могут быть любыми величинами из отрезка $[-1, 1]$.

Вернемся теперь к системе из двух q-битов и рассмотрим четыре эксперимента, когда в первом q-бите измеряется наблюдаемая спина $\sigma(\vec{a}_j)$ $(j = 1, 2)$, а во втором $\sigma(\vec{b}_k)$ $(k = 1, 2)$, где направления $\vec{a}_j, \vec{b}_k$ $(j, k = 1, 2)$ образуют конфигурацию, изображенную на рисунке.

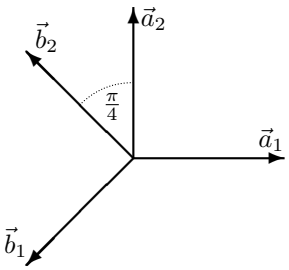

Рис. 2.1: Выбор векторов $\vec{a}_j$ и $\vec{b}_k$

При этом система приготавливается в одном и том же синглетном состоянии. Подстановка соответствующих значений корреляций из формулы (2.9) в левую часть формулы (2.10) дает значение $2\sqrt{2}$, нарушающее неравенство. Отсюда следует, что либо квантовая механика дает неправильные выражения для корреляций, либо для данной составной системы не существует классического вероятностного описания, удовлетворяющего условию локальности. После первого эксперимента (Аспе, 1981–1982) был проделан целый ряд аналогичных экспериментов по измерению ЭПР-корреляций, результаты которых с определенностью свидетельствуют в пользу квантовой механики.

2.5 Квантовая псевдотелепатическая игра

Квантовые корреляции (сцепленность) – новый информационный ресурс, *не сводимый* к классическим корреляциям. "Квантовое превосходство" в гротескной форме демонстрирует *игра Мермина-Переса*: игроки A и B играют против крупье C. C выбирает клетку (i, j) в матрице 3×3 и сообщает номер строки i игроку A, а номер столбца j – игроку B. A должен расставить ± 1 в своей строке т.ч. произведение $= 1$; B – ± 1 в своем столбце т.ч. произведение $= -1$. AB выигрывают, если выбранные ими элементы в клетке i, j совпадут. A и B могут выработать общую стратегию до начала игры, но после им не разрешено общаться: A не знает j, B не знает i. Например: $C \longrightarrow A$: строка 2, $C \longrightarrow B$: столбец 3

$$A: \begin{bmatrix} \dots & \dots & \dots \\ 1 & -1 & -1 \\ \dots & \dots & \dots \end{bmatrix} \qquad B: \begin{bmatrix} \dots & \dots & -1 \\ \dots & \dots & 1 \\ \dots & \dots & 1 \end{bmatrix}$$

AB проигрывают.

Классическая стратегия: Игроки A и B могли бы заранее выбрать фиксированную $3 \times 3-$матрицу с элементами ± 1, однако матрицы, удовлетворяющей сформулированным ограничениям, не существует. Из ограничения на A (соотв. B) произведение всех матричных элементов должно равняться 1 (соотв. -1). Они могут принять рандомизованную стратегию, но вероятность успеха всегда будет < 1.

Квантовая стратегия: Однако если A и B могут заранее создать сцепленное состояние и заранее выбрать схему квантовых измерений, каждый в своей лаборатории, то существует способ обеспечить выигрыш с вероятностью 1!

Рассмотрим *состояние Белла* $S = |\Psi\rangle_{AB}\langle\Psi|$,

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_A \otimes |\uparrow\rangle_B + |\downarrow\rangle_A \otimes |\downarrow\rangle_B\right) = \frac{1}{\sqrt{2}}\left(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle\right). \qquad (2.11)$$

Приготовленное сцепленное состояние является тензорным произведением двух состояний Белла для двух пар q-битов: A_1B_1 и A_2B_2 :

$$|\Psi\rangle = |\Psi\rangle_{A_1B_1} \otimes |\Psi\rangle_{A_2B_2}.$$

q-биты A_1 и A_2 посылаются игроку A, а q-биты B_1 и B_2 – игроку B до объявления C.

A и B также условливаются, что после получения номеров i и j, они производят измерения спинов, каждый в своих q-битах, в соответствии с таблицей

$$\begin{bmatrix} \sigma_0 \otimes \sigma_z & \sigma_z \otimes \sigma_0 & \sigma_z \otimes \sigma_z \\ \sigma_x \otimes \sigma_0 & \sigma_0 \otimes \sigma_x & \sigma_x \otimes \sigma_x \\ -\sigma_x \otimes \sigma_z & -\sigma_z \otimes \sigma_x & \sigma_y \otimes \sigma_y \end{bmatrix}$$

и записывают результаты измерения в соответствующие клетки.

Обозначая X_{ij} наблюдаемую на пересечении $i-$й строки и $j-$го столбца, имеем:

1. $X_{ij} = X_{ij}^*$ и $X_{ij}^2 = \sigma_0 \otimes \sigma_0 \equiv I$, т.ч. X_{ij} имеют собственные значения ± 1;

2. в каждой строке i операторы $X_{ij}; j = 1, 2, 3$ коммутируют, т.е. являются совместимыми наблюдаемыми, более того $X_{i1}X_{i2}X_{i3} = I$. Поэтому для любого $i = 1, 2, 3$, указанного C, игрок A может совместно измерить наблюдаемые $X_{ij}; j = 1, 2, 3$, получив результаты $+1$ или -1, подчиняющиеся ограничению для A. Тогда A помещает эти результаты в строку i. Аналогичное описание применимо к игроку B и любому указанному столбцу j.

3. чудесным образом, номера, помещенные A и B на пересечении $i-$й строки и $j-$го столбца обязательно совпадут! Это следует из равенства

$$\left(X_{ij}^A \otimes X_{ij}^B\right)|\Psi\rangle = |\Psi\rangle; \quad i, j = 1, 2, 3,$$

где X_{ij}^A (соотв. X_{ij}^B) – оператор X_{ij} в системе $A = A_1A_2$ (соотв. $B = B_1B_2$). Это равенство говорит, что если вся система $A_1A_2B_1B_2$ приготовлена в состоянии $|\Psi\rangle$, то произведение результатов измерений A и B в любой клетке ij будет равно 1, т.е. результаты совпадут.

Квантовая стратегия удовлетворяет всем правилам игры. Именно использование квантовых информационных технологий позволяет получить результат, недостижимый классическими средствами. С точки зрения классического наблюдателя дело обстоит так, как будто между A и B существует нематериальная связь. Игры типа описанной выше, были экспериментально реализованы и продемонстрировали “квантовое превосходство”.

2.6 Корреляционные неравенства и операторные алгебры

Если бы четыре корреляции в (2.10) принимали произвольные не зависящие друг от друга значения, то границу 2 в правой части неравенства следовало бы заменить на 4. Таким образом, квантовая локальность является ограничением, которое приводит к меньшему значению.

Квантовые корреляционные неравенства (Цирельсон). Пусть X_j, Y_k $(j,k = 1,2)$ вещественные четкие квантовые наблюдаемые, т.ч. $|X_j| \leq 1$, $|Y_k| \leq 1$, $X_jY_k = Y_kX_j$. Тогда для любого квантового состояния S

$$|\mathsf{E}_S X_1Y_1 + \mathsf{E}_S X_1Y_2 + \mathsf{E}_S X_2Y_1 - \mathsf{E}_S X_2Y_2| \leq 2\sqrt{2}. \tag{2.12}$$

Для системы из двух q-битов AB равенство в (2.12) достигается для наблюдаемых

$$X_j = \sigma(a_j) \otimes I_B, \quad Y_k = I_A \otimes \sigma(b_k)$$

и состояния Белла (2.11).

Адекватным математическим аппаратом для описания всевозможных корреляционных неравенств оказывается современная теория операторных пространств, получившая также название "квантовый функциональный анализ". В частности, знаменитая гипотеза Конна о конечномерной аппроксимируемости в II_1-факторах оказывается равносильной "гипотезе Цирельсона" о совпадении множеств корреляций между подсистемами составной системы, реализуемых в тензорной и алгебраической (локальная теория поля) моделях составных квантовых систем (в отличие от несовпадения множеств классически- и квантово-реализуемых корреляций, которое демонстрируется неравенствами типа (2.10))[3].

[3] M. Junge, M. Navascues, C. Palazuelos, D. Perez-Garcia, V. B. Scholz, R. F. Werner, Connes' embedding problem and Tsirelson's problem, J. Math. Phys. 52, 012102 (2011)

Глава 3

Применения сцепленных состояний

3.1 Квантовое состояние как информационный ресурс

В этом разделе нам потребуются элементарные сведения об эволюциях квантовой системы. В дальнейшем, в части II этот вопрос будет рассмотрен углубленно и с общих позиций теории открытых квантовых систем. Пока же достаточно знать следующее:

1) Обратимые эволюции квантовой системы описываются унитарными операторами U: вектор исходного чистого состояния ψ преобразуется в результате такой эволюции в $U\psi$. Соответственно, оператор плотности S преобразуется в USU^*.

2) Важнейший пример необратимой эволюции — изменение состояния в результате измерения. Простейшее идеальное квантовое измерение связывается с ортонормированным базисом $|e_x\rangle$, векторы которого индексированы возможными исходами измерения x. Если система перед измерением находится в состоянии S, то в результате такого измерения она переходит с вероятностью $\langle e_x|Se_x\rangle$ в состояние $|e_x\rangle\langle e_x|$. Весь статистический ансамбль после измерения разбивается на подансамбли, соответствующие различным исходам x, и описывается состоянием

$$S' = \sum_x |e_x\rangle\langle e_x|Se_x\rangle\langle e_x|, \tag{3.1}$$

вообще говоря отличным от исходного. Таким образом, квантовое измерение включает неустранимое воздействие на наблюдаемую систему, которое изменяет ее состояние, даже если исходы наблюдения "не считываются". В этом принципиальное отличие квантовых "наблюдаемых" от классических случайных величин, наблюдение которых не изменяет статистический ансамбль, а сводится к простому отбору его представителей.

Квантовое состояние приготавливается макроскопическими устройствами. Изменяя параметры устройства, мы изменяем параметры состояния, и таким образом получаем возможность "записывать" классическую информацию в квантовом состоянии. Простейший квантовый канал связи математически задается семейством (выходных или сигнальных) состояний S_x, где параметр x пробегает входной алфавит. Отображение $x \to S_x$ в сжатой форме содержит описание физического процесса, порождающего состояние S_x. Например, пусть $x = 0, 1$, причем S_1 когерентное состояние поля излучения лазера, а S_0 вакуумное состояние. В этом случае мы имеем канал с двумя чистыми неортогональными состояниями.

Для того чтобы извлечь классическую информацию, содержащуюся в квантовом состоянии, необходимо произвести измерение. В приведенном выше примере такую роль играет любой приемник лазерного излучения с возможной последующей обработкой результатов измерения. Если измерение задается базисом $|e_y\rangle$, то условная вероятность получить исход y, при условии, что был послан сигнал x, дается формулой

$$P(y|x) = \langle e_y|S_x e_y\rangle. \tag{3.2}$$

Таким образом, для фиксированного измерения мы получаем обычный канал связи. Это дает возможность поставить вопрос о максимальном количестве классической информации, которое может быть передано по данному квантовому каналу связи и о его пропускной способности. Этот вопрос будет детально рассмотрен в главе 4. Отметим здесь лишь один факт, имеющий принципиальное значение:

Пропускная способность любого квантового канала ограничена сверху величиной $\log \dim \mathcal{H}$*, причем эта величина достигается для "идеального" канала, сигнальные состояния которого образованы векторами о.н.б. в пространстве* $\mathcal{H}$*, а измерение задается этим же о.н.б.* Таким образом, размерность гильбертова пространства является мерой максимального информационного ресурса квантовой системы.

3.2 Сверхплотное кодирование

Рассмотрим теперь следующий вопрос. Нелокальный, с классической точки зрения, характер ЭПР-корреляций наводит на мысль попытаться использовать их для мгновенной передачи информации. Покажем, что этого невозможно достичь, находясь в рамках квантовой механики (с точки зрения которой ЭПР-корреляции не противоречат локальности). Рассмотрим две квантовые системы A и B, в пространствах $\mathcal{H}_A$ и $\mathcal{H}_B$ соответственно, которые находятся в сцепленном состоянии S_{AB}. В случае, представляющем интерес, системы пространственно разделены, хотя формально это ни в чем не выражается. Система A получает классическую информацию, содержащуюся в значениях параметра x, которая может быть использована для выполнения произвольных унитарных операций U_x в пространстве $\mathcal{H}_A$. При этом состояние системы AB переходит в $S_x = (U_x \otimes I_B)S_{AB}(U_x \otimes I_B)^*$,

таким образом, классическая информация записывается в квантовом состоянии составной системы. В свою очередь, над системой B может быть произведено произвольное измерение, описываемое о.н.б. $|e_y\rangle$ в $\mathcal{H}_B$. Легко видеть, что результирующая переходная вероятность (3.2) не зависит от x, а значит количество передаваемой информации в самом деле равно нулю.

Хотя ЭПР-корреляции сами по себе не позволяют передавать информацию, оказывается, что наличие таких корреляций между системами позволяет увеличить максимальное количество классической информации, передаваемой от A к B, вдвое, если между системами имеется идеальный квантовый канал связи, т. е. возможность безошибочно передать любое квантовое состояние. Таким образом, ЭПР-корреляции выступают как "катализатор" при передаче классической информации через квантовый канал связи, и с этот точки зрения, также представляют собой особого рода информационный ресурс.

Рассмотрим системы A и B, каждая из которых представляет собой q-бит, между которыми имеется идеальный квантовый канал связи. Из того что было сказано выше, вытекает, что максимальное количество классической информации, которое может быть передано от A к B, равно $\log 2 = 1$ бит, и получается при кодировании бита в два ортогональных вектора, например,

$$0 \to |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad 1 \to |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Протокол "сверхплотного кодирования," предложенный Беннетом и Виснером в 1992 г., имеет в своей основе простой математический факт: базис Белла

$$|e_+\rangle = |00\rangle + |11\rangle, \quad |e_-\rangle = |00\rangle - |11\rangle, \quad |h_+\rangle = |10\rangle + |01\rangle, \quad |h_-\rangle = |10\rangle - |01\rangle$$

в системе из двух q-битов AB (мы используем канонический базис $|0\rangle, |1\rangle$ в пространстве одного q-бита и для краткости опускаем нормировочный множитель $1/\sqrt{2}$) может быть получен из одного вектора $|e_+\rangle$ действием "локальных" унитарных операторов, т. е. операторов, действующих нетривиально только в пространстве q-бита A, например

$$|e_-\rangle = (\sigma_z \otimes I)|e_+\rangle, \quad |h_+\rangle = (\sigma_x \otimes I)|e_+\rangle, \quad |h_-\rangle = -i(\sigma_y \otimes I)|e_+\rangle.$$

Таким образом, если AB изначально находится в *сцепленном* состоянии $|e_+\rangle$, участник A может закодировать 2 бита классической информации в 4 состояния базиса Белла, производя только локальные операции, а затем (физически) послать свой q-бит B по идеальному квантовому каналу. Тогда, производя измерение в базисе Белла, участник B получает 2 бита классической информации. Конструкции протоколов сверхплотного кодирования и телепортации допускают обобщение на случай пространства произвольной конечной размерности.

3.3 Квантовая телепортация

До сих пор говорилось о передаче классической информации через квантовый канал связи. Такая информация может быть "записана" в квантовом состоянии и передана через физический канал. Однако квантовое состояние и само по себе является информационным ресурсом постольку, поскольку имеет статистическую неопределенность. Оказывается, что информация, содержащаяся в неизвестном квантовом состоянии, имеет качественные отличия от классической, и поэтому заслуживает специального термина *квантовая информация*. Наиболее ярким отличием квантовой информации является невозможность копирования (no cloning). Очевидно, что классическая информации может воспроизводиться в любом количестве. Но физический прибор, который бы выполнял аналогичную задачу для квантовой информации, противоречит принципам квантовой механики, так как преобразование

$$|\psi\rangle \to \underbrace{|\psi\rangle \otimes \cdots \otimes |\psi\rangle}_{n}$$

является нелинейным, и не может быть осуществлено унитарным оператором. Конечно, это можно сделать каждый раз специальным прибором для данного конкретного состояния (и даже для фиксированного набора ортогональных состояний), но не существует универсального прибора, который бы размножал произвольное квантовое состояние.

Каким образом может быть передано квантовое состояние? Очевидно, что можно просто физически переслать саму систему. Гораздо более интересный и нетривиальный способ — *телепортация* квантового состояния, при которой сама система физически не передается, а передается лишь классическая информация[1]. При этом существенным дополнительным ресурсом, который вновь играет роль "катализатора," является ЭПР-корреляция между входом и выходом канала связи. Заметим, что свести передачу произвольного квантового состояния к только передаче классической информации, не используя дополнительного квантового ресурса, невозможно: поскольку классическая информация копируема, это означало бы возможность копирования и квантовой информации.

Пусть имеются две квантовые системы A и B, описывающие, соответственно, вход и выход канала связи. На вход A поступает произвольное состояние $|\psi\rangle$; можно описать процедуру, при которой исходное состояние B перейдет в $|\psi\rangle$, а входное $|\psi\rangle$ с необходимостью разрушится (иначе мы имели бы копирование).

В простейшей (и основной) версии системы A и B являются двухуровневыми (q-битами).

1. Перед началом передачи система AB приготовляется в состоянии $|00\rangle + |11\rangle$.

[1] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channel," *Phys. Rev. Lett.*, vol 70, 1895-1899 1993.

2. C посылает A произвольное чистое состояние

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

Совокупность трех систем CAB описывается состоянием

$$(a|0\rangle + b|1\rangle) \otimes (|00\rangle + |11\rangle) = a|000\rangle + b|100\rangle + a|011\rangle + b|111\rangle.$$

3. Затем

   (a) A производит некоторое обратимое преобразование состояния системы CA;

   (b) A производит измерение (с 4 исходами, что составляет 2 бита классической информации). Преобразование и измерение будут описаны ниже.

4. A посылает результат измерения B по классическому каналу связи.

5. В зависимости от полученного результата измерения B производит некоторое преобразование и получает это произвольное $|\psi\rangle$.

Производимые преобразования являются характерными примерами логических операций, используемых в квантовом компьютинге. На 3-м шаге над системой CA производится операция CNOT (контролирумое “нет”):

$$|00\rangle \to |00\rangle, \quad |01\rangle \to |01\rangle, \quad |10\rangle \to |11\rangle, \quad |11\rangle \to |10\rangle,$$

при которой состояние первого q-бита сохраняется, а состояние второго q-бита не изменяется, либо изменяется на противоположное, в зависимости от состояния первого q-бита. При этом базис переходит в базис, следовательно, в 4-х мерном пространстве CA этому преобразованию соответствует унитарный оператор. Затем к q-биту C применяется операция Адамара H с унитарной матрицей

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Тогда

$$|0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \qquad |1\rangle \to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle),$$

т.е. базис поворачивается на угол $\pi/4$.

Начальное состояние всей системы CAB есть

$$a|000\rangle + b|100\rangle + a|011\rangle + b|111\rangle.$$

После действия CNOT на CA получаем

$$a|000\rangle + b|110\rangle + a|011\rangle + b|101\rangle.$$

Потом H действует на C

$$a(|000\rangle + |100\rangle) + b(|010\rangle - |110\rangle) + a(|011\rangle + |111\rangle) + b(|001\rangle - |101\rangle).$$

Выделяя состояние системы CA, получаем

$$|00\rangle(a|0\rangle + b|1\rangle) + |01\rangle(a|1\rangle + b|0\rangle) + |10\rangle(a|0\rangle - b|1\rangle) + |11\rangle(a|1\rangle - b|0\rangle).$$

Теперь производится измерение в системе CA, проецирующее на один из 4-х базисных векторов $|00\rangle, \ldots, |11\rangle$. Результат измерения $00, 01, 10, 11$ посылается от A к B по классическому (идеальному) каналу связи. В зависимости от полученного результата B применяет к своему состоянию один из унитарных операторов

$$I = \sigma_0, \quad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad -i\sigma_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

преобразующих состояние B в $a|0\rangle + b|1\rangle$.

Возможность телепортации состояния поляризации фотона была продемонстрирована экспериментально Цайлингером в 1997 г. С тех пор были проведены десятки экспериментов, включая телепортацию состояний массивных частиц (впервые в 2004 г.)

3.4 Квантовые алгоритмы

Идея квантового компьютера была предложена Фейнманом в 1981 г. для моделирования квантовомеханических систем. Вопрос: не может ли квантовое устройство решать какие-либо задачи более эффективно, чем классический компьютер, был впервые затронут в книге Ю.И. Манина "Вычислимое и невычислимое", 1980 г.) Простейшие, но довольно искусственные примеры таких задач рассмотрели Дейч и Джоза. Их усовершенствованием является алгоритм Саймона, который лежит в основе и алгоритма Шора, эффективно решающего важную и практически интересную (по крайней мере, с точки зрения криптографии) задачу разложения большого натурального числа на простые множители.

3.4.1 Алгоритм Саймона

Обозначим $B = \{0, 1\}$, $B^n = B^{\times n}$. Пусть задано отображение $f : B^n \to B^n$. Известно, что функция f является периодической, то есть $f(x) = f(y) \Leftrightarrow y = x \oplus \xi$, где $\xi \in B^n$ — двоичный (булев) вектор. Здесь $\oplus$ обозначает покомпонентное двоичное сложение векторов. Мы предполагаем $\xi \neq 0$, случай перестановки $\xi = 0$ может быть рассмотрен аналогично.

Требуется найти период ξ за наименьшее возможное число шагов (принимая за шаг каждый акт вычисления функции f). Классическое решение задачи сводится к перебору и требует число шагов $O(2^{n/2})$, растущее экспоненциально с n. (После вычисления s значений функции f, сравнивая

значения во всевозможных парах точек, мы можем исключить не более $s(s-1)/2$ из 2^n-1 значений ξ, так что в худшем случае $s(s-1)/2 = 2^n-1$, откуда $s \sim 2^{n/2}$). Можно доказать, что и применение вероятностных алгоритмов, которые дают правильный ответ лишь с заданной вероятностью $1-\varepsilon$, не позволяет добиться ускорения.

Квантовый алгоритм требует всего $O(n)$ шагов, если считать за шаг квантовое вычисление функции f. При этом решение носит вероятностный характер. Для описания квантового алгоритма нам понадобится n-мерное обобщение операции Адамара $H_n = \underbrace{H \otimes \cdots \otimes H}_{n}$. Рассмотрим *квантовый регистр* — физическую систему из n q-битов; информация будет задаваться состоянием этой системы. Если x — набор нулей или единиц длины n, то векторы $|x\rangle$ образуют о.н.б. Действие H_n в этом базисе задается формулой

$$H_n|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y \in B^n} (-1)^{x \cdot y} |y\rangle,$$

где $x \cdot y$ — скалярное произведение векторов $x \in B^n$, $y \in B^n$ по модулю 2. Оператор H_n унитарный, эрмитов и $H_n^2 = I$.

Алгоритм Саймона состоит из следующих шагов:

1. Сначала квантовый регистр приготовляется в основном состоянии $|0\rangle = |00\ldots\rangle$, затем применяется операция Адамара:

   $$|00\ldots\rangle \xrightarrow{H_n} \frac{1}{\sqrt{2^n}} \sum_{y \in B^n} |y\rangle.$$

   В результате получается суперпозиция всевозможных базисных состояний с одинаковыми коэффициентами.

2. Затем к этой суперпозиции применяется унитарный оператор, обратимо вычисляющий функцию f:

   $$\left(\sum_x |x\rangle \right) \otimes |z\rangle \xrightarrow{U_f} \sum_x |x\rangle \otimes |z \oplus f(x)\rangle,$$

   где $\oplus$ обозначает сложение по модулю 2, т.е. логическую операцию "XOR". Предполагается, что такой унитарный оператор дан "свыше" (поэтому его принято называть "оракулом"). Отметим, что в алгоритме Шора соответствующее вычисление описывается эффективно. В принципе, он может быть составлен из некоторых элементарных одно- и двух-кубитных операций, если известно, как само отображение f составлено из элементарных логических операций. Здесь $|z\rangle$ состояние *вспомогательного регистра*, который введен, чтобы сделать операцию вычисления функции обратимой. Если исходно этот регистр находится в основном состоянии $|0\rangle$, то

   $$\left(\sum_x |x\rangle \right) \otimes |00\ldots\rangle \longrightarrow \sum_x |x\rangle \otimes |f(x)\rangle.$$

3. Вновь применяя операцию Адамара, получаем вектор состояния

$$\frac{1}{2^n}\sum_{x\in B_n}\sum_{y\in B_n}(-1)^{x\cdot y}|y\rangle\otimes|f(x)\rangle.$$

4. Поскольку $\xi\neq 0$, отображение f принимает 2^{n-1} разных значений. Измеряя оба регистра, получаем 2^{n-1} разных исходов $(y,f(x))$ с вероятностью

$$\left(\frac{1}{2^n}\right)^2[(-1)^{x\cdot y}+(-1)^{(x+\xi)\cdot y}]^2,$$

равной $2^{-(2n-2)}$, если $y\cdot\xi=0$, и 0 в противном случае.

Таким образом, получается случайный равномерно распределенный вектор $y(\omega)$ из булевой "гиперплоскости" $y\cdot\xi=0$. Если повторить эту процедуру $n-1$ раз, то с положительной вероятностью полученные векторы будут линейно независимы, что позволяет найти вектор ξ.

Лемма 1. *Пусть $y_1(\omega),\ldots,y_{n-1}(\omega)$ вероятностно независимые, равномерно распределенные случайные векторы из гиперплоскости $y\cdot\xi=0$. Тогда*

$$\mathsf{P}\{y_1(\omega),\ldots,y_{n-1}(\omega)\ \ \}\geq e^{-1}.$$

Доказательство. Вектор $y(\omega)$ принимает 2^{n-1} равновероятных значений. Если $y_1(\omega),\ldots,y_{k-1}(\omega)$ линейно независимы, то имеется 2^{k-1} их различных линейных комбинаций. Поэтому получаем следующие значения условных вероятностей

$$\mathsf{P}\{y_k(\omega)\,\text{линейно независим от}\,y_1(\omega),\ldots,y_{k-1}(\omega)|y_1(\omega),\ldots,y_{k-1}(\omega)\text{линейно независимы}\}=$$
$$=\frac{2^{n-1}-2^{k-1}}{2^{n-1}}=1-\frac{1}{2^{n-k}}.$$

Тогда

$$\begin{aligned}&\mathsf{P}\{y_1(\omega),\ldots,y_k(\omega)\text{ линейно независимы}\}\\ &=\mathsf{P}\{y_k(\omega)\text{ линейно независим от }y_1(\omega),\ldots,y_{k-1}(\omega);\\ &\qquad\qquad y_1(\omega),\ldots,y_{k-1}(\omega)\text{ линейно независимы}\}\\ &=\left(1-\frac{1}{2^{n-k}}\right)\mathsf{P}\{y_1(\omega),\ldots,y_{k-1}(\omega)\text{ линейно независимы}\}.\end{aligned}$$

Следовательно

$$\begin{aligned}&\mathsf{P}\{y_1(\omega),\ldots,y_{n-1}(\omega)\text{ линейно независимы}\}\\ &\qquad=\left(1-\frac{1}{2^{n-1}}\right)\cdots\left(1-\frac{1}{2}\right)\\ &\qquad=\exp\left[\sum_{k=1}^{n-1}\ln\left(1-\frac{1}{2^k}\right)\right]\geq\exp\left[-\sum_{k=1}^{n-1}\frac{1}{2^k}\right]\geq e^{-1}.\end{aligned}$$

5) Повторяем всю процедуру m раз, где $(1-e^{-1})^m \leq \varepsilon$. Тогда с вероятностью $1-\varepsilon$ получим по крайней мере $n-1$ линейно независимых булевых векторов, ортогональных ξ, а значит, и сам вектор ξ.

Квантовый алгоритм требует лишь $O(n)$ применений оператора U_f вместо $O(2^{n/2})$ вычислений значения f для классического алгоритма. За счет чего достигается такое радикальное ускорение? Очевидно, за счет того, что однократное применение оператора U_f дает состояние, которое в латентной форме содержит все значения функции f, и из которого интересующая нас информация может быть извлечена посредством квантового измерения. Такой прием называют "квантовым параллелизмом". Важно, однако, подчеркнуть, что в отличие от параллелизма в классическом компьютинге, речь отнюдь не идет об одновременном вычислении всех значений функции.

3.4.2 Замечания об алгоритме Шора

Алгоритм, предложенный Шором в 1994 г., эффективно решает задачу нахождения множителя большого натурального числа $N \sim 2^n$. Задача факторизации — разложения на множители — одна из фундаментальных проблем математики, имеющая далеко не только академический интерес: трудность решения этой задачи лежит в основе надежности криптографии с открытым ключом. Наилучший из известных в настоящее время алгоритмов имеет экспоненциальную сложность $O(2^{cn^{1/3}\log^{2/3} n})$. Есть (но не доказано) предположение, что полиномиальное решение этой задачи не существует.

Квантовый алгоритм Шора имеет полиномиальную сложность $O(n^2 \log n \log\log n)$. Представление о его эффективности дает следующая грубая оценка: задача факторизации числа $N \sim 2^{800}$ не решаема за разумное время на классическом компьютере, тогда как применение квантового алгоритма при тактовой частоте 1 Мгц потребовало бы пару дней. Алгоритм использует сведение задачи факторизации к нахождению периода функции $f(x) = a^x (\mathrm{mod}\, N)$, где a выбирается случайным образом. Можно показать, что в большинстве случаев период r является четным и число $a^{r/2} \pm 1$ имеет общий множитель с N, который находится с помощью классического алгоритма Евклида. Алгоритм Шора включает детальное описание эффективного выполнения операции U_f. Нахождение периода $f(x)$ использует квантовую модификацию быстрого преобразования Фурье (роль которого в более простой задаче Саймона выполняло преобразование Адамара H_n). Подробнее об алгоритме Шора и квантовых вычислениях см. в [5, 2].

3.4.3 Алгоритм Гровера

Этот алгоритм решает задачу поиска. Более точно, предполагается, что задана булева функция $F : B^n \to B$, такая что $F(x_0) = 1$, $F(x) = 0$, $x \neq x_0$. Требуется найти x_0, причем вычисление значения функции F в любой заданной точке принимается за один шаг. Классический алгоритм сводится к перебору значений x и проверки для них равенства $F(x) = 1$,

что в наименее благоприятном случае требует $N \sim 2^n$ шагов. Квантовый алгоритм Гровера позволяет решить задачу за $\approx \sqrt{N} = 2^{n/2}$ шагов, при этом решение носит вероятностный характер.

Предполагается, что в гильбертовом пространстве, натянутом на базис $|x\rangle, x \in B^n$, задан "оракул" – унитарный оператор U_F, такой что

$$U_F|x\rangle = |x\rangle, x \neq x_0, \quad U_F|x_0\rangle = -|x_0\rangle.$$

Введем обозначения

$$|\bar{x}_0\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \neq x_0} |x\rangle, \quad \theta_0 = \arcsin \frac{1}{\sqrt{N}}.$$

Алгоритм состоит из следующих шагов:

1. К основному состоянию применяется операция Адамара

$$|0\rangle \xrightarrow{H_n} \frac{1}{\sqrt{N}} \sum_x |x\rangle = |\psi(\theta_0)\rangle,$$

где введено обозначение

$$|\psi(\theta)\rangle = \sin\theta|x_0\rangle + \cos\theta|\bar{x}_0\rangle.$$

Эта операция переводит вектор $|0\rangle$ в вектор, лежащий в плоскости, натянутой на базис $|x_0\rangle, |\bar{x}_0\rangle$.

2. К полученному состоянию применяется унитарный оператор $U = H_n J H_n U_F$, где J — оператор, действующий по формулам $J|0\ldots0\rangle = |0\ldots0\rangle, \quad J|x\rangle = -|x\rangle, \;\; x \neq 0$.

Задача 13. Проверьте, что

$$U|\psi(\theta)\rangle = |\psi(\theta+\varphi)\rangle,$$

где

$$\sin\varphi = 2\frac{\sqrt{N-1}}{N},$$

т. е. U осуществляет поворот на угол φ в плоскости, натянутой на натянутой на базис $|x_0\rangle, |\bar{x}_0\rangle$.

После применения оператора U m раз, где $m \approx (\pi/4)\sqrt{N}$, конечное состояние $|\psi(\theta_m)\rangle, \theta_m = \theta_0 + m\varphi \approx \pi/2$ становится очень близким к искомому: $|\psi(\theta_m)\rangle \approx |x_0\rangle$, причем тем ближе, чем больше N.

В этом алгоритме квантовый параллелизм проявляется в том, что вычисления функции F в отдельных точках заменяются действием унитарного оператора U_F на суперпозицию базисных состояний, что и позволяет достичь полиномиального ускорения.

Глава 4

Классически-квантовые каналы

4.1 Основные понятия классической теории информации

4.1.1 Энтропия и сжатие данных

Пусть X дискретная случайная величина, принимающая значения в конечном множестве $\mathcal{X} = \{1, \ldots, |\mathcal{X}|\}$, и имеющая распределение вероятностей $p\{p_x\}$, так что значение $x \in \mathcal{X}$ появляется с вероятностью p_x. *Энтропия* случайной величины X определяется соотношением

$$H(X) = -\sum_{x \in \mathcal{X}} p_x \log p_x, \tag{4.1}$$

с соглашением $0 \log 0 = 0$ (далее log, как правило, обозначает двоичный логарифм).

Задача 14. $0 \leq H(X) \leq \log |\mathcal{X}|$, причем минимальное значение принимается на вырожденных распределениях, а максимальное — на равномерном.

Обычно $H(X)$ интерпретируется как мера неопределенности, изменчивости или информационного содержания случайной величины X. Поясним последнее утверждение.

Рассмотрим случайный источник, который порождает последовательность независимых одинаково распределенных случайных величин с распределением p. Последовательность $w = (x_1, \ldots, x_n)$ букв алфавита $\mathcal{X}$ называется *словом* длины n. Общее количество таких слов $|\mathcal{X}|^n = 2^{n \log |\mathcal{X}|}$. Поэтому можно закодировать все эти слова, используя двоичные последовательности длины $n \log |\mathcal{X}|$, т.е. $n \log |\mathcal{X}|$ *бит*. Однако, используя то обстоятельство, что p в общем случае не-равномерное распределение, можно

предложить лучший способ кодирования. Возможность сжатия данных тесно связана со свойством *асимптотической равнораспределенности*, которое является прямым следствием закона больших чисел:

Теорема 11. *Если $X_1, \ldots, X_n, \ldots$ независимые и одинаково распределенные случайные величины с распределением $p = \{p_x\}$, то*

$$-\frac{1}{n}\sum_{i=1}^{n} \log p_{x_i} \longrightarrow H(X) \qquad \text{по вероятности.} \tag{4.2}$$

Таким образом, для любых $\delta, \epsilon > 0$ найдется n_0, такое что для всех $n \geq n_0$ имеет место

$$\mathsf{P}\left\{ \Big| -\frac{1}{n}\sum_{i=1}^{n} \log p_{x_i} - H(X) \Big| < \delta \right\} > 1 - \epsilon. \tag{4.3}$$

Замечая, что вероятность появления слова $w = (x_1, \ldots, x_n)$ равна

$$p_w = p_{x_1} \cdot \ldots \cdot p_{x_n} = 2^{-n\left(-\frac{1}{n}\sum_{i=1}^{n} \log p_{x_i}\right)} \tag{4.4}$$

мы теперь можем использовать соотношение (4.3) чтобы ввести понятие *типичного слова*: слово w, имеющее вероятность p_w, называется δ-типичным, если

$$2^{-n(H(X)+\delta)} < p_w < 2^{-n(H(X)-\delta)}. \tag{4.5}$$

Непосредственно устанавливаются следующие свойства типичных слов:

1. Существует не более $2^{n(H(X)+\delta)}$ типичных слов.

2. Для достаточно больших n существует, по крайней мере, $(1-\epsilon)2^{n(H(X)-\delta)}$ типичных слов.

3. Множество не-типичных слов имеет вероятность $\leq \epsilon$.

Теперь можно осуществить эффективное *сжатие данных*, используя все двоичные последовательности длины $n(H(X)+\delta)$ чтобы закодировать все δ-типичные слова и отбрасывая не-типичные (или кодируя их одним и тем же добавочным символом). Вероятность ошибки при таком кодировании будет меньше или равна ϵ. Обратно, любой код, использующий двоичные последовательности длины $n(H(X)-\delta)$, имеет асимптотически неисчезающую вероятность ошибки, стремящуюся к единице при $n \to \infty$ (задача 15).

Поскольку эффективное кодирование требует асимптотически $N \sim 2^{nH(X)}$ слов, энтропия $H(X)$ может быть интерпретирована как мера количества информации (в битах на передаваемый символ) в случайном источнике. Ясно, что для равномерного распределения $p_x = 1/|\mathcal{X}|$ энтропия $H(X) = H_{max}(X) = \log|\mathcal{X}|$ и сжатие невозможно.

4.1.2 Пропускная способность канала с шумом

Канал связи с шумом описывается вероятностями переходов $p(y|x)$ из входного алфавита $\mathcal{X}$ в выходной алфавит $\mathcal{Y}$, т. е. условными вероятностями того, что принят символ $y \in \mathcal{Y}$, при условии, что был послан символ $x \in \mathcal{X}$. Соответствующее уменьшение информационного содержания источника описывается *шенноновским количеством информации*:

$$I(X;Y) = H(X) - H(X|Y), \tag{4.6}$$

где $H(X) = -\sum_x p_x \log p_x$ энтропия источника (входа), а $H(X|Y)$ *условная энтропия* входа относительно выхода Y, которая описывает *потерю* информации в канале связи:

$$\begin{aligned} H(X|Y) = & \quad \sum_y p_y H(X|Y=y) = -\sum_y p_y \sum_x \frac{p_{x,y}}{p_y} \log \frac{p_{x,y}}{p_y} = \\ = & \; -\sum_{x,y} p_{x,y} \log p_{x,y} + \sum_y p_y \log p_y = H(X,Y) - H(Y). \end{aligned}$$

Здесь $H(X,Y)$ *совместная энтропия* пары случайных величин (X,Y), соответствующая совместному распределению $p_{x,y} = p(y|x)p_x$. Подставляя эту формулу в определение шенноновского количества информации (4.6), мы видим, что оно симметрично по X и Y, и поэтому может быть также названо *взаимной информацией*

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X), \tag{4.7}$$

где в последней формуле уже $H(Y)$ может быть интерпретирована, как информационное содержание выхода, а $H(Y|X)$ как его бесполезная составляющая, обусловленная *шумом.* Взаимная информация всегда неотрицательна: тот факт, что $H(X) \geq H(X|Y)$ легко вытекает из вогнутости функции $-x \log x$ (задача 16). Отсюда также вытекает свойство субаддитивности энтропии: $H(XY) \leq H(X) + H(Y)$. Далее, $I(X;Y) = 0$ тогда и только тогда, когда X и Y независимые случайные величины: $p_{x,y} = p_x \cdot p_y$.

Если посылается последовательность букв, и канал $p(y|x)$ действует независимо на каждую посланную букву, то он называется *каналом без памяти. Пропускная способность* такого канала определяется как

$$C = \max_{\{p_x\}} I(X;Y), \tag{4.8}$$

где максимум берется по всевозможным распределениям на входе $\{p_x\}$.

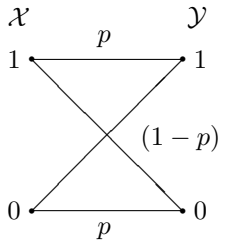

Рис. 4.1: Двоичный симметричный канал

В качестве примера рассмотрим *двоичный симметричный канал.* В этом случае $\mathcal{X}$ и $\mathcal{Y}$ состоят из двух букв $0, 1$, которые передаются без ошибки с вероятностью p. Вводя *двоичную энтропию*

$$h(p) = -p \log p - (1-p) \log(1-p), \tag{4.9}$$

взаимную информацию можно записать как $I(X;Y) = H(X) - h(p)$. Максимум этой величины, равный

$$C = 1 - h(p), \tag{4.10}$$

достигается на равномерном входном распределении: $p_0 = p_1 = 1/2$.

Применяя *блочное кодирование* для канала без памяти, когда канал используется для посылки n букв, получаем

$$x^n = \left\{ \begin{array}{ccc} x_1 & \longrightarrow & y_1 \\ x_2 & \longrightarrow & y_2 \\ \vdots & & \vdots \\ x_n & \longrightarrow & y_n \end{array} \right\} = y^n$$

где $p(y^n|x^n) = p(y_1|x_1)\cdot\ldots\cdot p(y_n|x_n)$. Пусть Y^n обозначает выход дискретного канала без памяти со входом X^n. Очевидно, что последовательность $C_n = \max_{X^n} I(X^n;Y^n)$ супераддитивна: $C_{n+m} \geq C_n + C_m$. Более того, можно доказать, что она аддитивна, используя следующую лемму:

Лемма 2.

$$I(X^n;Y^n) \leq \sum_{i=1}^{n} I(X_i;Y_i). \tag{4.11}$$

Доказательство. Имеет место *цепное правило* для условной энтропии:

$$H(X_1,\ldots,X_n) = \sum_{i=1}^{n} H(X_i|X_1,\ldots,X_{i-1}), \tag{4.12}$$

которое легко доказать по индукции, используя формулу:

$$H(X,Y) = H(X) + H(Y|X). \tag{4.13}$$

Тогда взаимная информация

$$\begin{aligned} I(X^n;Y^n) &= H(Y^n) - H(Y^n|X^n) = \\ &= H(Y^n) - \sum_{i=1}^{n} H(Y_i|Y_1,\ldots,Y_{i-1},X^n) = \\ &= H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i), \end{aligned}$$

поскольку для канала без памяти Y_i зависит только от X_i и, таким образом,

$$I(X^n;Y^n) \leq \sum_{i=1}^{n} (H(Y_i) - H(Y_i|X_i)) = \sum_{i=1}^{n} I(X_i;Y_i).$$

Беря максимум выражения (4.11), получаем аддитивность, в частности, $C_n = nC$.

Определение. *Кодом* (W, V) *размера* N для канала $p(y|x)$ называется совокупность N слов $w^{(1)}, \ldots, w^{(N)}$ длины n вместе с разбиением множества $\mathcal{Y}^n$ на N непересекающихся подмножеств $V^{(0)}, V^{(1)}, \ldots, V^{(N)} \subset \mathcal{Y}^n$. Подмножества $V^{(1)}, \ldots, V^{(N)}$ интерпретируются как области принятия решения: если на выходе принято значение $y^n \in V^{(j)}; j = 1, \ldots, N$, то принимается решение, что было послано слово $w^{(j)}$; если же принято $y^n \in V^{(0)}$, то никакого определенного решения не принимается. Таким образом, *максимальная вероятность ошибки* такого кода есть

$$P_e(W, V) = \max_{1 \le j \le N} \left(1 - p(V^{(j)}|w^{(j)})\right), \tag{4.14}$$

где $p(V^{(j)}|w^{(j)}) = \mathsf{P}(Y^n \in V^{(j)}|X^n = w^{(j)})$. *Средняя вероятность ошибки* равна

$$\overline{P}_e(W, V) = \frac{1}{N} \sum_{i=1}^{N} \left(1 - p(V^{(j)}|w^{(j)})\right) \le P_e(W, V), \tag{4.15}$$

и, как показывает следующая лемма, с точки зрения теории информации она асимптотически эквивалентна максимальной вероятности ошибки $P_e(W, V)$.

Лемма 3. *Пусть код размера* $2N$ *имеет среднюю вероятность ошибки* $\overline{P}_e(W, V) < \epsilon$. *Тогда найдется подкод размера* N, *имеющий максимальную вероятность ошибки* $P_e(W, V) < 2\epsilon$.

Доказательство. Предположим, что среди $2N$ слов имеется по крайней мере $N + 1$ слово с вероятностью ошибки $p(V^{(j)}|w^{(j)}) \ge 2\epsilon$, так что построить требуемый N-подкод невозможно. Тогда средняя ошибка $2N$-кода ограничена снизу величиной $\overline{P}_e(W, V) \ge \frac{1}{2N} 2\epsilon(N+1) > \epsilon$, что противоречит предположению. □

Лемма 4 (Неравенство Фано). *Пусть* X, Y *случайные величины и* $\widehat{X} = \widehat{X}(Y)$ *оценка случайной величины* X *с вероятностью ошибки* $P_e = P(\widehat{X}(Y) \ne X)$, *тогда*

$$H(X|Y) \le h(p_e) + p_e \log(|\mathcal{X}| - 1) \le 1 + p_e \log |\mathcal{X}|. \tag{4.16}$$

Доказательство. Пусть E индикатор ошибки оценивания,

$$E = \begin{cases} 0, & \text{если } \widehat{X}(Y) = X \\ 1, & \text{в противном случае.} \end{cases} \tag{4.17}$$

Аналогично соотношению $H(E|X) = H(E, X) - H(X)$ получаем

$$H(E|X, Y) = H(E, X|Y) - H(X|Y) = 0, \tag{4.18}$$

поскольку E является функцией $(X, \widehat{X})$, и поэтому имеет определенное значение при фиксированных значениях (X, Y). Поэтому

$$\begin{aligned} H(X|Y) &= H(E, X|Y)H(E|Y) + H(X|E, Y) \leq \\ &\leq H(E) + (1 - p_e)H(X|E = 0, Y) + p_e H(X|E = 1, Y) = \\ &= h(p_e) + p_e \log(|\mathcal{X}| - 1) \leq 1 + p_e \log |\mathcal{X}|, \end{aligned}$$

где был использован тот факт, что $H(X|E = 0, Y)$ также равно нулю, поскольку $E = 0$ означает, что мы знаем X, если известно Y.

Теорема 12 (Теорема кодирования для канала с шумом). *Пусть*

$$p_e(n, N) = \min_{W,V} \overline{P}_e(W, V)$$

минимальная средняя ошибка для всевозможных N-кодов со словами длины n. Тогда при $n \to \infty$

$$p_e(n, 2^{nR}) \begin{cases} \to 0 & \text{если } R < C \quad \text{(прямая теорема кодирования);} \\ \not\to 0 & \text{если } R > C \quad \text{(слабое обращение);} \\ \to 1 & \text{если } R > C \quad \text{(сильное обращение).} \end{cases}$$

Величина $R = \frac{\log N}{n}$ называется *скоростью передачи* и равна числу передаваемых битов на символ для данного кода.

Доказательство слабого обращения. Рассмотрим произвольный код размера N со словами $w^{(1)}, \ldots, w^{(N)}$ длины n и разбиение множества $\mathcal{Y}^n$ на $N + 1$ область принятия решения $V^{(0)}, V^{(1)}, \ldots, V^{(N)} \subset Y^n$. Обозначим Z случайную величину, принимающую значения $1, \ldots, N$ с равными вероятностями $\frac{1}{N}$ и пусть $\widehat{Z}(Y^n)$ оценка для Z, такая что $\widehat{Z}(Y^n) = j$, если $Y^n \in V^{(j)}$. Тогда согласно неравенству Фано

$$\begin{aligned} nC = C_n &\geq I(Z; Y^n) = H(Z) - H(Z|Y^n) \geq \\ &\geq \log N - 1 - \underbrace{P\{\widehat{Z}(Y^n) \neq Z\}}_{=\overline{P}_e(W,V)} \log N. \end{aligned}$$

Подставляя $N = 2^{nR}$ и оптимизируя по W, V, получаем

$$\begin{aligned} nC &\geq nR - 1 - p_e(n, 2^{nR})nR, \\ \frac{C}{R} &\geq (1 - p_e(n, 2^{nR})) - \frac{1}{nR}, \end{aligned}$$

и в пределе $n \to \infty$ при $R > C$:

$$\lim \inf_{n\to\infty} p_e(n, 2^{nR}) \geq 1 - \frac{C}{R} > 0.$$

Основная идея доказательства *прямой теоремы кодирования*, восходящая к работе Шеннона[1], состоит в использовании *случайного кодирования*. Рассмотрим N слов $w^{(1)}, \ldots, w^{(N)}$, выбираемых случайным образом независимо с распределением вероятностей $P\{w^{(j)} = (x_1, \ldots, x_n)\} = p_{x_1} \cdot \ldots \cdot p_{x_n}$, где однобуквенное распределение $\{p_x\}$ выбрано так, что оно максимизирует $I(X;Y)$. Заметим, что имеется примерно $2^{nH(X)}$ ($2^{nH(Y)}$) типичных слов на входе (на выходе), и в среднем $2^{nH(Y|X)}$ типичных слов на выходе для каждого входного слова w.

Для того, чтобы ошибка различения слов на выходе стремилась к нулю, надо, чтобы множества типичных слов на выходе, соответствующие разным словам на входе, асимптотически не пересекались, поэтому размер кода не должен превосходить

$$N \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}. \tag{4.19}$$

Таким образом, $N \approx 2^{nC}$. Конечно, это рассуждение в высшей степени эвристично; строгое доказательство, реализующее эту идею, можно найти, например, в [10], [11].

Теорема кодирования раскрывает, таким образом, операциональный смысл понятия пропускной способности как максимальной скорости асимптотически безошибочной передачи информации через данный канал связи.

4.2 Оптимальное различение квантовых состояний

4.2.1 Постановка задачи

В этом разделе мы рассмотрим статистическую задачу, которая позволит в дальнейшем перейти к изучению квантовых каналов связи.

Пусть квантовая система находится в одном из состояний S_j, $j = 1, \ldots, n$. Над системой можно производить произвольное измерение. Требуется найти оптимальную процедуру измерения, позволяющую наилучшим образом выяснить, в каком из этих состояний находится система. Такая постановка задачи характерна для теории связи и для математической статистики.

Измерение (приемник) будет описываться наблюдаемой, т. е. разложением единицы $M = \{M_k\}$. Вероятность принять решение k, при условии, что был послан сигнал j, при этом равна $p_M(k|j) = \operatorname{Tr} S_j M_k$. Если был послан сигнал j, то вероятность того, что было принято правильное решение, есть $p_M(j|j)$. Примем дополнительное предположение, что сигнал j появляется с вероятностью π_j (например, в случае равновероятных сигналов $\pi_j = 1/n$.)

[1] К. Шеннон, Статистическая теория передачи электрических сигналов, в кн. "Теория передачи электрических сигналов при наличии помех", М.: ИЛ, 1953, 7-87.

Тогда средняя вероятность правильного решения

$$\mathcal{P}\{M\} = \sum_{j=1}^{n} \pi_j p_M(j),$$

а средняя вероятность ошибки равна $1 - \mathcal{P}\{M\}$, и задача состоит в ее минимизации, или же в максимизации $\mathcal{P}\{M\}$. В статистике применяется и минимаксный критерий, когда минимизируется $\max_j p_M(j|\, j)$, но мы его не будем здесь затрагивать.

Другой важный критерий, который мы рассмотрим позже, – шенноновская информация. Согласно формуле (4.7) количество взаимной информации между входом J (j — номер входного состояния) и выходом K (k — номер решения) дается формулой

$$\mathcal{J}\{M\} = \underbrace{H(K)}_{\text{энтропия}} - \underbrace{H(K|J)}_{\substack{\text{условная}\\ \text{энтропия}}}$$

$$= -\sum_k \sum_j p_{_M}(k|j)\pi_j \log \sum_l p_{_M}(k|l)\pi_l + \sum_j \pi_j \sum_k p_{_M}(k|j) \log p_{_M}(k|j)$$

$$= \sum_j \pi_j \sum_k p_{_M}(k|j) \log \left[\frac{p_{_M}(k|j)}{\sum\limits_l p_{_M}(k|l)\pi_l} \right],$$

4.2.2 Различение по максимуму правдоподобия

Будем максимизировать вероятность правильного решения

$$\mathcal{P}\{M\} = \sum_{j=1}^{n} \pi_j \operatorname{Tr} S_j M_j = \operatorname{Tr}(\sum_{j=1}^{n} \underbrace{\pi_j S_j}_{W_j} M_j).$$

Множество наблюдаемых, по которым ведется оптимизация

$$\mathfrak{M}_n = \left\{ M = \{M_k\}_{k=1,\ldots,n} : M_k \geq 0, \sum_{k=1}^{n} M_k = I \right\}$$

— выпуклое. Смесь (выпуклая комбинация) наблюдаемых описывает статистику измерения, производимого прибором с флуктуирующими параметрами. Функция $\mathcal{P}\{M\}$ аффинна, т.е.

$$\mathcal{P}\left\{\sum p_\lambda M^\lambda\right\} = \sum p_\lambda \mathcal{P}\{M^\lambda\}.$$

Оптимизация аффинной функции, заданной на выпуклом множестве — типичная задача линейного программирования.

Теорема 13. *Средняя вероятность правильного решения $\mathcal{P}\{M\}$ достигает максимума в крайней точке множества $\mathfrak{M}_n$. Наблюдаемая M^0 оптимальна тогда и только тогда, когда найдется эрмитов оператор Λ^0 такой, что*

1) $(\Lambda^0 - W_k)M_k^0 = 0$;

2) $\Lambda^0 \geq W_k$.

При этом имеет место соотношение двойственности

$$\max\{\mathcal{P}\{M\} : M \in \mathfrak{M}_n\} = \min\{\operatorname{Tr}\Lambda : \Lambda \geq W_k, k = 1, \ldots, n\}. \tag{4.20}$$

Доказательство. Докажем достаточность условий теоремы.

Пусть наблюдаемая M^0 удовлетворяет этим условиям, $M \in \mathfrak{M}_n$ — произвольная наблюдаемая, тогда

$$\mathcal{P}\{M\} = \operatorname{Tr}\sum_k W_k M_k \overset{2)}{\leq} \operatorname{Tr}\sum_k \Lambda^0 M_k$$

$$= \operatorname{Tr}\Lambda^0 \overset{1)}{=} \operatorname{Tr}\sum_k W_k M_k^0 \mathcal{P}\{M^0\}.$$

Здесь был использован простой факт:

Задача 17. Для $B \geq 0$ в $\mathcal{B}(\mathcal{H})$ и A_1, A_2, таких что $A_1 \leq A_2$, имеет место $\operatorname{Tr} A_1 B \leq \operatorname{Tr} A_2 B$, причем равенство имеет место тогда и только тогда, когда $A_1 B = A_2 B$.

Докажем необходимость условий теоремы.

Положим $M_k = X_k^2$, где X_k эрмитовы операторы, удовлетворяющие условию $\sum_k X_k^2 = I$. Применяя метод Лагранжа, сводим задачу максимизации $\mathcal{P}\{M\}$ на множестве $\mathfrak{M}_n$ к нахождению максимума функции

$$\operatorname{Tr}\sum_k W_k X_k^2 - \operatorname{Tr}\Lambda(\sum_k X_k^2 - I), \tag{4.21}$$

где Λ эрмитов оператор, по всевозможным наборам эрмитовых операторов X_k. Пусть X_k^0 оптимальный набор, положим $X_k = X_k^0 + \epsilon Y_k$, и рассмотрим (4.21) как функцию от ϵ. Рассматривая коэффициенты при ϵ и ϵ^2, получаем условия

$$\operatorname{Tr}[(W_k - \Lambda)X_k^0 + X_k^0(W_k - \Lambda)]Y_k = 0,$$

$$\operatorname{Tr}(W_k - \Lambda)Y_k^2 \leq 0$$

для произвольных эрмитовых Y_k, т.е.

$$(W_k - \Lambda)X_k^0 + X_k^0(W_k - \Lambda) = 0, \quad \Lambda - W_k \geq 0.$$

Второе неравенство есть условие 2) теоремы. Полагая $M_k^0 = (X_k^0)^2$, получаем из первого соотношения $\mathrm{Tr}(\Lambda - W_k)M_k^0 = 0$, что вместе со вторым неравенством влечет условие 1).

З а д а ч а 18. Доказать, что операторный множитель Лагранжа Λ является единственным решением двойственной задачи в правой части (4.20).

Проиллюстрируем смысл и полезность этих условий на нескольких примерах. Рассмотрим сначала классический случай, когда операторы плотности состояний коммутируют.

П р и м е р 1. Пусть операторы W_k (пропорциональные S_k) коммутируют, тогда существует общий ортонормированный базис, где они все диагонализуются

$$W_k = \sum_\omega W_k(\omega)|\omega\rangle\langle\omega|.$$

Тогда можно взять

$$\Lambda^0 = \sum_\omega \max_k W_k(\omega)|\omega\rangle\langle\omega|,$$

где $\max\limits_k W_k(\omega)$ — верхняя огибающая функций $W_k(\omega)$; $k = 1, \ldots, n$; $M_k^0 = \sum\limits_\omega \mathbf{1}_{\Omega_k}(\omega)|\omega\rangle\langle\omega|$; $\mathbf{1}_{\Omega_k}$ обозначает индикатор подмножества Ω_k, и подмножества $\Omega_k \subset \{\omega : \Lambda^0(\omega) = W_k(\omega)\}$ образуют разбиение множества $\Omega = \{\omega\}$.

Это приводит к принципу *максимального правдоподобия* в классической статистике: k-е решение необходимо принимать для тех ω, для которых $W_k(\omega)$ максимально. Таким образом, в классическом случае оптимальная наблюдаемая всегда может быть выбрана нерандомизованной. Это прямо связано с тем фактом, что в коммутативном случае крайние точки множества $\mathfrak{M}_n$ отвечают ортогональным разложениям единицы (см. задачу 7).

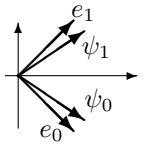

Рис. 4.2: Различение двух чистых состояний.

П р и м е р 2 (У п р а ж н е н и е 19). *Различение двух квантовых состояний.* Произвольная наблюдаемая с двумя значениями имеет вид $M = \{M_0, M_1\}$, $M_{0,1} \geq 0$, $M_1 = I - M_0$, причем стандартные наблюдаемые характеризуются условием $M_0^2 = M_0$, которое в точности соответствует крайним точкам "некоммутативного отрезка"$\mathfrak{M}_2 = \{0 \leq M_0 \leq I\}$ (задача 20). Таким образом, для различения двух состояний достаточно стандартных наблюдаемых.

Приведем явное решение. Пусть S_0, S_1 произвольные операторы плотности. Оператор Лагранжа

$$\Lambda = \pi_0 S_0 M_0 + \pi_1 S_1 M_1 = \pi_1 S_1 + (\pi_0 S_0 - \pi_1 S_1)M_0$$

эрмитов, поэтому $[M_0, \pi_0 S_0 - \pi_1 S_1] = 0$. Неравенство $\Lambda \geq \pi_1 S_1$ влечет $(\pi_0 S_0 - \pi_1 S_1)M_0 \geq 0$, а из $\Lambda \geq \pi_0 S_0$ вытекает

$$(\pi_0 S_0 - \pi_1 S_1)M_0 \geq (\pi_0 S_0 - \pi_1 S_1).$$

Очевидным решением является $M_0 = \mathbf{1}_{(0,\infty)}(\pi_0 S_0 - \pi_1 S_1)$, т. е. проектор на собственное подпространство оператора $\pi_0 S_0 - \pi_1 S_1$, отвечающий положительным собственным значениям. При этом

$$\max \mathcal{P}\{M\} = \mathrm{Tr}[\pi_1 S_1 + (\pi_0 S_0 - \pi_1 S_1)_+] = \frac{1}{2}[1 + \|\pi_0 S_0 - \pi_1 S_1\|_1],$$

где $\|T\|_1 = \mathrm{Tr}\,|T|$ –ядерная норма оператора T. Здесь $|T| = T_+ + T_-$, где $T_+(T_-)$ положительная (отрицательная) часть эрмитова оператора T, т. е. компонента его спектрального разложения, отвечающая положительной (отрицательной) части спектра.

Пусть $S_0 = |\psi_0\rangle\langle\psi_0|$, $S_1 = |\psi_1\rangle\langle\psi_1|$. В этом случае оптимум дается ортонормированным базисом $\{|e_0\rangle, |e_1\rangle\}$, так что $M_0 = |e_0\rangle\langle e_0|$, $M_1 = |e_1\rangle\langle e_1|$. Вектор $|e_0\rangle$ отвечает положительному собственному числу λ_0 оператора $\pi_0|\psi_0\rangle\langle\psi_0| - \pi_1|\psi_1\rangle\langle\psi_1|$, причем $\max \mathcal{P}\{M\} = \pi_1 + \lambda_0$. Диагонализуя оператор $\pi_0|\psi_0\rangle\langle\psi_0| - \pi_1|\psi_1\rangle\langle\psi_1|$, можно дать явное решение задачи (см. [8]). Пусть для простоты $\pi_0 = \pi_1 = 1/2$, тогда оптимальный базис расположен симметрично по отношению к $|\psi_0\rangle, |\psi_1\rangle$ (рис. 4.2) и

$$\max \mathcal{P}\{M^0\} = \frac{1}{2}\Big(1 + \sqrt{1 - |\,\langle\psi_1|\psi_0\rangle\,|^{\,2}}\Big).$$

Задача 21. Показать, что для различения n чистых состояний с линейно независимыми векторами $|\psi_j\rangle$; $j = 1, \ldots, n$, достаточно стандартных наблюдаемых. В этом случае оптимальная наблюдаемая дается векторами некоторой ортонормированной системы $|e_j\rangle$; $j = 1, \ldots, n$, см. [8].

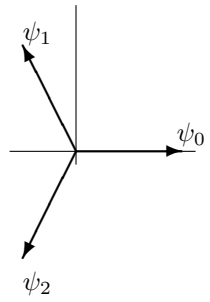

Рис. 4.3: *Векторы трех состояний*

Пример 3. На плоскости (рассматриваемой как вещественное подпространство двумерного унитарного пространства) рассмотрим "равноугольную" конфигурацию трех векторов (рис. 4.3)

$$|\psi_j\rangle = \begin{bmatrix} \cos\frac{2j\pi}{3} \\ \sin\frac{2j\pi}{3} \end{bmatrix}, \quad j = 0, 1, 2. \tag{4.22}$$

Соответствующие операторы плотности $S_j = |\psi_j\rangle\langle\psi_j|$, описывают состояния двухуровневой системы, например, плоскополяризованного фотона или частицы со спином 1/2.

Имеем

$$S_j = \begin{bmatrix} \cos^2\frac{2j\pi}{3} & \cos\frac{2j\pi}{3}\sin\frac{2j\pi}{3} \\ \cos\frac{2j\pi}{3}\sin\frac{2j\pi}{3} & \sin^2\frac{2j\pi}{3} \end{bmatrix}$$

$$= \frac{1}{2}\left(I + \begin{bmatrix} \cos\frac{4j\pi}{3} & \sin\frac{4j\pi}{3} \\ \sin\frac{4j\pi}{3} & \cos\frac{4j\pi}{3} \end{bmatrix}\right). \tag{4.23}$$

Поскольку

$$\sum_{j=0}^{2} e^{i\frac{4j\pi}{3}} = 0,$$

то

$$\sum_{j=0}^{2} S_j = \frac{3}{2} I,$$

то есть $M_k^0 = \frac{2}{3} S_k$ является разложением единицы.

Покажем, что в случае равновероятных состояний, $\pi_j = 1/3$, $\{M_k^0\}$ дает оптимальную наблюдаемую. Проверим условия теоремы. Поскольку $S_j^2 = S_j$, то

$$\Lambda^0 = \sum_{j=0}^{2} \frac{1}{3} S_j \frac{2}{3} S_j = \frac{2}{9} \sum_{j=0}^{2} S_j, = \frac{1}{3} I.$$

так что $I/3 = \Lambda^0 \geq S_j/3$ (условие 2)) и

$$\left(\Lambda^0 - \frac{1}{3} S_j \right) \frac{2}{3} S_j = \frac{1}{3} (I - S_j) S_j = 0$$

— условие 1) также выполнено.

Итак, $\max \mathcal{P}\{M\} = \operatorname{Tr} \Lambda^0 = 2/3$. Найдем теперь максимум по всевозможным стандартным наблюдаемым с тремя значениями. Нетривиальное ортогональное разложение единицы с тремя компонентами в двумерном пространстве имеет вид $M_0 = |e_0\rangle\langle e_0|$, $M_1 = |e_1\rangle\langle e_1|$, $M_2 = 0$, где $|e_0\rangle, |e_1\rangle$, – произвольный базис. Находя соответствующий максимум, получаем

$$\max_{M - \text{стандартные}} \mathcal{P}\{M\} = \frac{1 + \sqrt{3}/2}{3} < \frac{2}{3} = \max_{M \in \mathfrak{M}} \mathcal{P}\{M\}.$$

Таким образом, использование в квантовой статистике неортогональных разложений единицы в качестве наблюдаемых (т.е. использование квантовой рандомизации — дополнительной независимой квантовой системы в фиксированном состоянии) может приводить к выигрышу при различении состояний исходной системы! Подчеркнем, что в классическом случае никакая рандомизация не может улучшить качество процедуры различения состояний.

С геометрической точки зрения, причина состоит в том, что в квантовом случае не все крайние точки множества наблюдаемых $\mathfrak{M}_3$ (среди которых и находится наиболее информативная наблюдаемая), описываются ортогональными разложениями единицы.

4.2.3 Максимум информации

Пусть система находится в одном из m состояний $S_1, \ldots, S_m$, и над системой производится измерение наблюдаемой $M = \{M_k\}; k = 1, \ldots, n$, с целью получить максимальное количество информации. Число исходов измерения n заранее не фиксировано. A priori нет оснований требовать совпадения n и m. Множество всех наблюдаемых с конечным числом исходов обозначим $\mathfrak{M}$.

Таким образом, есть переходная вероятность $p_{_M}(k|j) = \operatorname{Tr} S_j M_k$, и шенноновское количество информации дается формулой

$$\mathcal{J}\{M\} = \sum_j \pi_j \sum_k p_{_M}(k|j)\Big[\log p_{_M}(k|j) - \log \sum_l p_{_M}(k|l)\pi_l\Big], \qquad (4.24)$$

где π_j — априорные вероятности состояний.

Лемма 5. *$\mathcal{J}\{M\}$ — выпуклая функция на $\mathfrak{M}$, т.е.*

$$\mathcal{J}\{pM^{(1)} + (1-p)M^{(2)}\} \le p\mathcal{J}\{M^{(1)}\} + (1-p)\mathcal{J}\{M^{(2)}\}.$$

В силу аффинной зависимости переходной вероятности от M, достаточно доказать, что $\mathcal{J}\{M\}$ является выпуклой функцией от переходной вероятности. Это вытекает из следующего общего свойства.

Лемма 6. *Шенноновское количество информации $\mathcal{J}\{M\}$ является выпуклой функцией от переходных вероятностей $p(k|j)$ и вогнутой функцией от априорных вероятностей π_j.*

Ограничимся доказательством первого утверждения, а второе оставим в качестве упражнения.

Доказательство. Рассмотрим множество переходных вероятностей $p(k|j) \ge 0$, $\sum\limits_k p(k|j) = 1$. Имеем

$$\mathcal{J}\{M\} = \sum_k \sum_j p(k|j)\pi_j \Big[\log p(k|j) - \log \sum_l p(k|l)\pi_l\Big].$$

Достаточно доказать выпуклость по переменным x для любого фиксированного k следующих функций

$$\sum_j \underbrace{p(k|j)}_{|||\,x_j}\pi_j \Big[\log \underbrace{p(k|j)}_{|||\,x_j} - \log(\sum_l \underbrace{p(k|l)}_{|||\,x_l}\pi_l)\Big],$$

поскольку количество информации является суммой слагаемых вида

$$f(x) = \sum_j \pi_j x_j \Big[\log x_j - \log \sum_l x_l \pi_l\Big].$$

Дифференцируя по x_j, получаем

$$\frac{\partial f(x)}{\partial x_j} = \pi_j \Big[(1 + \log x_j) - (1 + \log \sum_l x_l \pi_l)\Big]$$

(здесь для простоты log — натуральный логарифм) и

$$\frac{\partial^2 f(x)}{\partial x_j \partial x_k} = \delta_{kj}\frac{\pi_j}{x_j} - \frac{\pi_j \pi_k}{\sum\limits_l \pi_l x_l}; \quad \sum_{j,k} c_j c_k \frac{\partial^2 f(x)}{\partial x_j \partial x_k} = \sum_j c_j^2 \frac{\pi_j}{x_j} - \frac{(\sum_j \pi_j c_j)^2}{\sum_l \pi_l x_l}.$$

Согласно неравенству Коши-Буняковского,

$$\sum_j \pi_j c_j \sum \pi_j \sqrt{x_j}\frac{c_j}{\sqrt{x_j}} \le \sum \pi_l x_l \sum \frac{\pi_l}{x_l} c_j^2,$$

что и доказывает выпуклость функции f, а значит, и шенноновской информации.

Задача 22. Максимум непрерывной выпуклой функции на компактном выпуклом множестве достигается в крайней точке этого множества.

Таким образом, надо исследовать крайние точки множества $\mathfrak{M}$.

Лемма 7. *Если наблюдаемая M' получена укрупнением исходов наблюдаемой M, то $\mathcal{J}\{M'\} \le \mathcal{J}\{M\}$.*

Доказательство. Достаточно показать, что если два исхода j_1, j_2 наблюдаемой M объединить в один, не трогая остальных (к таким операциям сводится последовательно любое укрупнение), то

$$p_{_M}(j_1|i)\Big[\log p_{_M}(j_1|i) - \log\sum_l p_{_M}(j_1|l)\pi_l\Big] +$$

$$+p_{_M}(j_2|i)\Big[\log p_{_M}(j_2|i) - \log\sum_l p_{_M}(j_2|l)\pi_l\Big] \ge$$

$$\ge \underbrace{\{p_{_M}(j_1|i) + p_{_M}(j_2|i)\}}_{\text{Отвечает одному исходу в } M'} \Big[\log\Big(p_{_M}(j_1|i) + p_{_M}(j_2|i)\Big) -$$

$$- \log\sum_l \pi_l\Big(p_{_M}(j_1|i) + p_{_M}(j_2|i)\Big)\Big]$$

Введем множитель $1/2$:

$$\frac{1}{2}p_{_M}(j_1|i)[\ldots] + \frac{1}{2}p_{_M}(j_2|i)[\ldots] \ge$$

$$\ge \frac{p_{_M}(j_1|i) + p_{_M}(j_2|i)}{2}\left\{\log\frac{[\ldots]}{2} - \log\sum\frac{[\ldots]}{2}\right\}.$$

и просуммируем по i. Теперь утверждение следует из выпуклости функции f.

Теорема 14[2]. *Пусть дан набор квантовых состояний $S_1, \ldots S_m$ с определенными вероятностями $\pi_1, \ldots, \pi_m$, тогда существует наблюдаемая M^0, для которой $\max\limits_{\mathfrak{M}} \mathcal{J}\{M\} = \mathcal{J}\{M^0\}$ и такая, что ее компоненты — линейно независимые операторы ранга 1, т.е. $M_j^0 = |\psi_j\rangle\langle\psi_j|_{j=1,\ldots,n}$, а число компонент $n \le d^2$, где $d = \dim\mathcal{H}$. Если все операторы S_j имеют вещественные матрицы в некотором базисе, то $n \le d(d+1)/2$.*

[2] E. B. Davies, "Information and quantum measurement," *IEEE Trans. Inform. Theory* **24**, no. 6, 596-599 (1978).

Доказательство. Пусть $\widetilde{M}^0$ — оптимальная наблюдаемая, $\widetilde{M}^0 = \{\widetilde{M}_1^0,..., \widetilde{M}_j^0,...\}$. Поскольку ее компоненты — эрмитовы операторы, согласно спектральной теореме каждый из них можно разложить по ортонормированному базису собственных векторов, оставляя только компоненты с положительными собственными числами:

$$0 \le X = \sum x_j |e_j\rangle\langle e_j| = \sum \sqrt{x_j}|e_j\rangle\langle e_j|\sqrt{x_j} = \sum |\psi_j\rangle\langle\psi_j|,$$

где $|\psi_j\rangle = \sqrt{x_j}|e_j\rangle$. Построим "разукрупненную"наблюдаемую $M^0 = \{|\psi_j\rangle\langle\psi_j|\}_{j=1,...,n}$ (можем считать все ψ_j различными после объединения одинаковых в одну компоненту.) Пользуясь леммой 7 об укрупнении, имеем $\mathcal{J}\{M^0\} \ge \mathcal{J}\{\widetilde{M}^0\}$.

Согласно лемме 6 и задаче 22, можно считать, что M^0 крайняя точка, имеющая компоненты $M_j^0 = |\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|$ $(j = 1,\dots,n)$. Отсюда по теореме 6 следует, что операторы $|\psi_j\rangle\langle\psi_j|$ линейно независимы. Но максимальное число линейно независимых эрмитовых операторов в d-мерном унитарном пространстве равно d^2 ($d(d+1)/2$ в вещественном случае).

Явное решение возможно в случаях, когда есть некоторая симметрия.

Пример 1. Рассмотрим простейший случай — два вектора на плоскости $S_j = |\psi_j\rangle\langle\psi_j|; j = 0, 1$.

Конфигурацию состояний полностью характеризует вещественный параметр $\varepsilon = |\langle\psi_0|\psi_1\rangle|$. Кроме того, имеется априорное распределение π_0, π_1. Согласно теореме, достаточно взять $n = 3$ ($d = 2$, вещественный случай.) Специальными рассуждениями можно показать, что на самом деле максимум достигается на ортонормированном базисе, оптимальном по максимуму правдоподобия (т.е. минимуму средней ошибки), так что фактически $n = 2$ (Левитин, 1994).

Интересен симметричный случай, когда $\pi_0 = \pi_1 = 1/2$. В этом случае

$$\max_M \mathcal{J}\{M\} = 1 - h\left(\frac{1+\sqrt{1-\varepsilon^2}}{2}\right) \tag{4.25}$$

и максимум информации достигается на базисе, расположенном симметрично по отношению к векторам ψ_0, ψ_1 (рис. 4.2), оптимальном по критерию максимального правдоподобия.

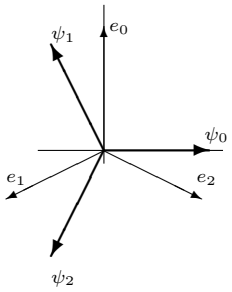

Рис. 4.4: Информационный оптимум для 3-х равноугольных состояний.

Пример 2. Случай трех равновероятных "равноугольных" чистых состояний (4.23) с углами $2\pi/3$ между направлениями спинов. Согласно теореме, $m \le d(d+1)/2 = 3$. Используя симметрию задачи, можно доказать, что информационно-оптимальная наблюдаемая имеет вид $M_k = \frac{2}{3}|e_k\rangle\langle e_k|_{k=0,1,2}$, где $e_k \perp \psi_k$ (см. рис. 4.2.3; задача 23). Таким образом, она не совпадает с наблюдаемой, оптимальной по максимуму правдоподобия, для которой $e_k = \psi_k$. Более того, можно показать, что последняя является наихудшей с точки зрения информационного

критерия. Максимум информации по всевозможным наблюдаемым: $\max_M \mathcal{J}\{M\} = \log(3/2) \approx 0.585$, тогда как $\max_{M\text{-стандартные}} \mathcal{J}(M) \approx 0.458$.

Пример 3. Пусть имеется n чистых состояний с линейно независимыми векторами. "Естественное" предположение, что существует информационно-оптимальная наблюдаемая с $m = n$ исходами, оказывается неверным. Это было показано в работе Шора[3], который рассмотрел конфигурацию из трех равноугольных векторов в трехмерном вещественном пространстве. Согласно теореме 7, число исходов информационно-оптимальной наблюдаемой ограничено величиной $d(d+1)/2 = 6$ и именно такое число исходов оказывается необходимым (хотя выигрыш по сравнению с тремя исходами настолько мал, что его трудно заметить при численной оптимизации).

Однако, как показал Дэвис[4], если состояния получены действием неприводимого представления некоторой группы симметрий, то существует ковариантная информационно-оптимальная наблюдаемая с числом исходов $m = n$. В случае трех равноугольных векторов на плоскости имеется вращательная симметрия, которая действует неприводимо над полем вещественных чисел (хотя, конечно, приводимо над полем комплексных чисел). Но в трех измерениях вращения вокруг оси приводимы даже над полем вещественных чисел, и упомянутый результат оказывается неприменим.

4.3 Сжатие квантовой информации

Выше уже было отмечено, что квантовая информация — это новый вид информации, который можно передавать, но нельзя размножать. Пусть имеется источник, производящий чистые состояния $|\psi_1\rangle, \ldots, |\psi_a\rangle$ с вероятностями $p_1, \ldots, p_a$ (аналог классического алфавита). Могут посылаться длинные последовательности букв (слова), т.е. каждое слово задается последовательностью $w = (x_1, \ldots, x_n)$, $x_j \in \{1, \ldots, a\}$.

Источник посылает сигнал $|\psi_w\rangle = |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle$ с вероятностью $p_w = p_{x_1} \cdot \cdots \cdot p_{x_n}$. *Кодирование* – это сопоставление чистому состоянию $|\psi_w\rangle\langle\psi_w|$ оператора плотности S_w в гильбертовом пространстве $\mathcal{H}_d \subset \mathcal{H}^{\otimes n}$. Проблема состоит в том, чтобы кодирующие состояния не слишком сильно отличались от исходных, и в то же время находились в подпространстве по возможности минимальной размерности. Точность воспроизведения исходных состояний кодирующими измеряется величиной

$$F_n = \sum_w p_w \langle\psi_w|S_w|\psi_w\rangle;$$

[3] P. W. Shor, "On the number of elements needed in a POVM attaining the accessible information," Arxiv:quant-ph/0009077.

[4] E. B. Davies, Information and quantum measurement, IEEE Trans. Inform. Theory **24** N6, pp. 596-599 1978.

чем ближе она к единице, тем точнее воспроизведение.

Для оператора плотности $S = \sum s_j |e_j\rangle\langle e_j|$ рассмотрим энтропию фон Неймана:

$$H(S) = -\sum_j s_j \log s_j = -\operatorname{Tr} S \log S. \tag{4.26}$$

Далее нам понадобятся элементарные свойства квантовой энтропии:

1) $0 \leq H(S) \leq \log d$, причем минимум достигается на чистых состояниях (и только на них), а максимум – на хаотическом состоянии $\overline{S} = I/d$.

2) $H(USU^*) = H(S)$, где U унитарный оператор (сохранение энтропии при обратимых преобразованиях).

3) $H(S_1 \otimes S_2) = H(S_1) + H(S_2)$ (аддитивность).

Следующий результат показывает, что, подобно энтропии Шеннона в классическом случае, квантовая энтропия определяет максимальную степень сжатия квантовых данных, т.е. количество квантовой информации.

Теорема 15[5]. *Обозначим* $\overline{S}_p = \sum_{x=1}^{a} p_x |\psi_x\rangle\langle\psi_x|$. *Тогда*

1) *Для любых* $\varepsilon, \delta > 0$ *и для достаточно больших* n *существует подпространство* $\mathcal{H}_d \subset \mathcal{H}^{\otimes n}$ *размерности* $d \leqslant 2^{n(H(\overline{S}_p)+\delta)}$ *и такие кодирующие состояния* S_w *в* $\mathcal{H}_d$, *что* $F_n > 1 - \varepsilon$;

2) *для любого подпространства* $\mathcal{H}_d$ *с* $d \leqslant 2^{n(H(\overline{S}_p)-\delta)}$ *и любого выбора* S_w *в* $\mathcal{H}_d$ *имеет место* $F_n < \varepsilon$ *для достаточно больших* n.

Замечание. Это утверждение раскрывает информационный смысл квантовой энтропии, подобно тому как идея сжатия данных раскрывала смысл классической энтропии. Для смеси чистых квантовых состояний

$$\sum_{x=1}^{a} p_x |\psi_x\rangle\langle\psi_x| = \overline{S}_p$$

энтропия оператора плотности $\overline{S}_p$ является мерой квантовой информации, содержащейся в ансамбле, поскольку $2^{nH(\overline{S}_p)}$ есть критическое значение размерности гильбертова пространства. (Напомним классический результат: пусть имеется источник, посылающий символы $1, \ldots, a$ с вероятностями $p_1, \ldots, p_a$, тогда количество слов, асимптотически безошибочно пересылаемых источником, есть $N \sim 2^{nH(p)}$, где $H(p) = -\sum_x p_x \log p_x$).

Доказательство.

1) В однобуквенном пространстве $\mathcal{H}$ рассмотрим спектральное разложение оператора

$$\overline{S}_p = \sum_j \lambda_j |e_j\rangle\langle e_j|. \tag{4.27}$$

[5]R. Jozsa, B. Schumacher, "A new proof of the quantum noiseless coding theorem," *J. Modern Optics* **41**, no. 12, 2343-2349 1994.

Пусть $J = (j_1, \dots, j_n)$, $\lambda_J = \lambda_{j_1} \dots \lambda_{j_n}$, $|e_J\rangle = |e_{j_1}\rangle \otimes \dots \otimes |e_{j_n}\rangle$, тогда спектральное разложение тензорной степени оператора $\overline{S}_p$ имеет вид

$$\overline{S}_p^{\otimes n} = \sum_J \lambda_J |e_J\rangle\langle e_J|.$$

Выделим в множестве всевозможных значений J подмножество

$$J_{n,\delta} = \left\{ J \,:\, 2^{-n(H(\overline{S}_p)+\delta)} < \lambda_J < 2^{-n(H(\overline{S}_p)-\delta)} \right\},$$

и обозначим E проектор на собственное подпространство, состоящее из векторов $|e_J\rangle$, $\lambda_J \in J_{n,\delta}$. Подпространство $E\mathcal{H}^{\otimes n}$ называется *типичным подпространством*. Оценим его размерность:

$$\dim E\mathcal{H}^{\otimes n} = \operatorname{Tr} E \leqslant \operatorname{Tr} \frac{\overline{S}_\pi}{2^{-n(H(\overline{S}_\pi)+\delta)}} \leqslant 2^{n(H(\overline{S}_\pi)+\delta)}. \tag{4.28}$$

Возьмем подпространство $\mathcal{H}_d = E\mathcal{H}^{\otimes n}$, а кодирование зададим правилом

$$S_w = \frac{E|\psi_w\rangle\langle\psi_w|E}{\langle\psi_w|E|\psi_w\rangle}.$$

Тогда точность воспроизведения

$$F_n = \sum_w p_w \langle\psi_w|S_w|\psi_w\rangle = \sum_w p_w \langle\psi_w|E|\psi_w\rangle$$

$$= \operatorname{Tr} E(\sum_w p_w |\psi_w\rangle\langle\psi_w|) = \operatorname{Tr} E S_p^{\otimes n} = \sum_{J\in J_{n,\delta}} \lambda_J. \tag{4.29}$$

Пусть $\lambda_J = \{\lambda_{j_1} \dots \lambda_{j_n}\}$ — классическое распределение вероятностей. Тогда сумма в правой части равна вероятности

$$\mathsf{P}\{2^{-n(H(\overline{S}_p)+\delta))} < \lambda_J < 2^{-n(H(\overline{S}_p)-\delta))}\} =$$

$$= \mathsf{P}\{H(\overline{S}_p) - \delta < -\frac{1}{n}\sum_{k=1}^n \log \lambda_{j_k} < H(\overline{S}_p) + \delta\}$$

$$= \mathsf{P}\{\left| -\frac{1}{n}\sum_{k=1}^n \log \lambda_{j_k} - H(\overline{S}_p) \right| < \delta\}, \tag{4.30}$$

где $\mathsf{E}\{-\log \lambda_{(\cdot)}\} = -\sum_{x=1}^a \lambda_x \log \lambda_x H(\overline{S}_p)$. Согласно закону больших чисел $F_n \longrightarrow 1$ при $n \to \infty$.

2) Пусть S_w произвольные операторы плотности в произвольном подпространстве $\mathcal{H}_d$ размерности d, и пусть P_d проектор на $\mathcal{H}_d$. Тогда $S_w \leqslant P_d$ и

$$F_n = \sum_w p_w \langle\psi_w|S_w|\psi_w\rangle \leqslant \operatorname{Tr} P_d \sum p_w |\psi_w\rangle\langle\psi_w| = \operatorname{Tr} P_d \overline{S}_p^{\otimes n}$$

Выберем теперь E как проектор на типичное подпространство, отвечающее $\epsilon/2$, $\delta/2$. Тогда правая часть оценивается как

$$\operatorname{Tr} \overline{S}_p^{\otimes n} E P_d + \operatorname{Tr} \overline{S}_p^{\otimes n} (1-E) P_d \leqslant \operatorname{Tr} P_d \| \overline{S}_p^{\otimes n} E \| + \operatorname{Tr} \overline{S}_p^{\otimes n} (1-E) \leqslant$$

$$\leqslant d 2^{-n(H(\overline{S}_p)-\delta/2)} + \frac{\varepsilon}{2} \leqslant 2^{-n\delta/2} + \frac{\varepsilon}{2} < \varepsilon \qquad (4.31)$$

для достаточно больших n.

4.4 Формулировка и обсуждение квантовой теоремы кодирования

Теорема Шеннона дает основу для введения такого понятия, как пропускная способность классического канала с шумом (максимальная скорость асимптотически безошибочной передачи информации через канал). Простейшая модель квантового канала предполагает, что есть классический параметр x, пробегающий (конечный) входной алфавит и отображение $x \to S_x$ в квантовые состояния на выходе канала. Например, двоичный оптический квантовый канал может быть реализован следующим образом: если $x = 0$, то поле излучения находится в вакуумном состоянии; если $x = 1$, то лазер генерирует когерентное состояние. Роль квантовой степени свободы может также играть поляризация или направление спина.

Теперь рассмотрим передачу слова — последовательности букв $w = \{x_1, \ldots, x_n\}$, которому сопоставляется состояние S_w :

$$w = \left.\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} \begin{matrix} \longrightarrow \\ \\ \\ \longrightarrow \end{matrix} \begin{matrix} S_{x_1} \\ \overset{\otimes}{\vdots} \\ \overset{\otimes}{S_{x_n}} \end{matrix} \right\} = S_w \ \text{ в } \ \mathcal{H}^{\otimes n} = \mathcal{H} \otimes \cdots \otimes \mathcal{H}$$

Предположение о том, что w кодируется в тензорное произведение состояний S_{x_j}, соответствует определению канала без памяти в классическом случае.

На выходе производится измерение некоторой наблюдаемой $M = \{M_{\widehat{w}}^{(n)}\}$ в пространстве $\mathcal{H}^{\otimes n}$ (получив исход измерения $\widehat{w}$, считаем, что было послано $\widehat{w}$). В итоге приемник выдает ответ о принятом решении; таким образом, разложение единицы в пространстве $\mathcal{H}^{\otimes n}$ описывает статистику всей решающей процедуры, которая включает в себя физическое измерение и последующую классическую обработку его результатов. Выбор наблюдаемой M формально аналогичен выбору решающей процедуры в классическом случае, но как мы увидим, играет здесь гораздо более важную роль. После того, как M выбрана, мы получаем классический канал $p_M(y|x) = \operatorname{Tr} S_x M_y$ в однобуквенном случае, и $p_{M^{(n)}}(\widehat{w}|w) = \operatorname{Tr} S_w M_{\widehat{w}}^{(n)}$ – в n-буквенном.

Определим шенноновскую взаимную информацию между входом и выходом. Если есть априорное распределение вероятностей p на $\mathcal{X}$ и выбрана

процедура измерения M на выходе, то шенноновская информация между входом и выходом дается формулой

$$I_1(p,M)=\sum_x p_x\sum_y p_M(y|x)\Big[\log p_M(y|x)-\log\sum_z p_M(y|z)p_z\Big],$$

а максимальное количество информации, допустимое законами квантовой механики, равно

$$\max_{p,M} I_1(p,M)=C_1.$$

Аналогично, если для n-й степени канала задано априорное распределение $p^{(n)}$ на словах длины n и измерение $M^{(n)}$ в гильбертовом пространстве $\mathcal{H}^{\otimes n}$, то соответствующие информационные количества равны

$$\begin{aligned}&I_n(p,M)\\&=\sum_w p_w\sum_{\widehat{w}} p_{M^{(n)}}(\widehat{w}|w)\Big[\log p_{M^{(n)}}(\widehat{w}|w)-\log\sum_{w'} p_{M^{(n)}}(\widehat{w}|w')p_{w'}\Big],\\&\max_{p^{(n)},M^{(n)}} I_n(p^{(n)},M^{(n)})=C_n.\end{aligned}$$

Имеет место удивительный факт: если для классического канала без памяти всегда $C_n=nC_1$, то в квантовом случае уже для $d=2$ (двоичный канал) возможно строгое неравенство $C_n>nC_1$ (строгая супераддитивность классической информации в квантовом канале). Причина этого в том, что для n-й степени квантового канала существуют коллективные (сцепленные) наблюдаемые, которые ни в каком смысле не сводятся к разделимым наблюдаемым, даже с последующей классической обработкой результатов их измерений.

Можно сказать, что это есть двойственное проявление корреляций Эйнштейна–Подольского–Розена. Последние возникают, когда рассматривается сцепленное (т. е. неразделимое) состояние составной квантовой системы, а измерения разделимы. Строгая супераддитивность информации имеет место для разделимых состояний и обусловлена существованием коллективных (сцепленных) измерений.

Перейдем к формулировке теоремы кодирования, из которой, в частности, будет следовать свойство супераддитивности.

О п р е д е л е н и е. *Кодом* (W,M) длины n и размера N называется набор слов $W=\{w^{(1)},\dots,w^{(N)}\}$ вместе с разложением единицы $M=\{M_j\}$ в $\mathcal{H}^{\otimes n}$ с исходами $j=0,1,\dots,N$; исход 0 означает уклонение от принятия решения.

Средняя ошибка кода равна

$$\overline{P}_e(W,M)=\frac{1}{N}\sum_{j=1}^{N}[1-\underbrace{p_M(j|\,w^{(j)})}_{\substack{\text{вероятность}\\ \text{правильного}\\ \text{решения}}}\,]=\frac{1}{N}\sum_{j=1}^{N}[1-\operatorname{Tr}S_{w_j}M_j]$$

Обозначим $\min_{W,M} \overline{P}_e(W,M) = p_e(n,N)$ минимальную среднюю ошибку по всем кодам размера N, использующим слова длины n.

Обозначим

$$C_\chi = \max_p \left\{ H\left(\sum_x p_x S_x\right) - \sum_x p_x H(S_x) \right\},$$

где $H(S)$ – энтропия фон Неймана (4.26).

Т е о р е м а 16 (Квантовая теорема кодирования[6]). *При* $n \to \infty$

1) $p_e(n, 2^{nR}) \to 0$, *если* $R < C_\chi$ *(прямая теорема);*

2) $p_e(n, 2^{nR}) \not\to 0$, *если* $R > C_\chi$ *(слабое обращение);*

(сильное обращение: $p_e(n, 2^{nR}) \to 1,\ n \to \infty$*).*

Эта теорема оправдывает название *классическая пропускная способность* для величины C_χ. В самом деле, определим C_∞ как $\lim_n C_n/n$, где $C_n = \max I_n(p, M)$. Из классической теоремы кодирования (теорема 12) вытекает, что утверждение теоремы 16 выполняется с заменой C_χ на C_∞. Таким образом, утверждение теоремы 16 состоит в том, что $C_\infty = C_\chi$.

Если состояния $S_x = |\psi_x\rangle\langle\psi_x|$ чистые, то

$$C_\chi = \max_p H\left(\sum_x p_x |\psi_x\rangle\langle\psi_x|\right).$$

Из свойства 1) энтропии следует, что всегда $C_\chi \le \log d$. Таким образом, несмотря на то, что в унитарном пространстве имеется бесконечно много разных чистых состояний, это обстоятельство не может быть использовано для передачи неограниченного количества информации. Грубо говоря, чем гуще расположены векторы, тем труднее становится их различить. Верхняя граница и максимум информации достигаются, если выходные состояния являются ортогональными $|e_x\rangle\langle e_x|,\ x = 1, \ldots, d$, и $p_x = \frac{1}{d}$. Заметим, что такие выходные состояния, как правило, не могут быть получены на выходе реального канала связи. Замечательно, однако, что как показывает следующий пример, ортогональность выходных состояний не является необходимой для достижения пропускной способности идеального канала.

Рассмотрим конфигурацию (4.22) из трех равновероятных "равноугольных" векторов ψ_0, ψ_1, ψ_2. Тогда

$$\sum_{x=0}^{2} p_x |\psi_x\rangle\langle\psi_x| = \frac{1}{2} I$$

и, как следует из теоремы кодирования, пропускная способность такого канала имеет то же максимальное значение $C_\chi = 1$ бит, что и для ортогональных состояний. Заметим, что это достигается только благодаря использованию оптимального кода, включающего коллективное измерение.

[6] A.S. Holevo, The capacity of quantum channel with general signal states. IEEE Trans. Inform. Theory, 1998, v. 44, N1, 269-273; Arxiv quant-ph/9611023, 1996, а также B. Schumacher, M. D. Westmoreland, *Sending classical information via noisy quantum channel*, Phys. Rev. A. **56**, 131-138 1997.

З а д а ч а 24. Величина

$$C_1 = 1 - h\left(\frac{1+\sqrt{3}/2}{2}\right) \approx 0,645$$

достигается для не-равномерного распределения $p_0 = p_1 = 1/2$, $p_2 = 0$ и соответствующего оптимального измерения для двух равновероятных состояний (см. пример 1 в разделе 4.3).

З а д а ч а 25. Рассмотрим двоичный квантовый канал с чистыми состояниями ψ_0, ψ_1. Докажите, что

$$C_\chi = h\left(\frac{1-\varepsilon}{2}\right),$$

а максимум информации по измерениям и априорным распределениям за один шаг

$$\max_{p,M} I_1(p, M) = C_1$$

дается формулой (4.25). В пределе слабого сигнала ($\varepsilon \to 1$) выигрыш от использования сцепленности на выходе канала $C_\chi/C_1 \sim -\log(1-\varepsilon)$ стремится к бесконечности.

Эти примеры иллюстрируют феномен *супераддитивности* классической информации в квантовом канале связи. Именно, из неравенства $C_1 < C_\chi$ следует свойство супераддитивности $nC_1 < C_n$, для достаточно больших n.

4.5 Квантовая граница классической информации и доказательство обратной теоремы

Т е о р е м а 17 (Квантовая граница классической информации). *Для любого распределения p и любой наблюдаемой M*

$$I_1(p, M) \le H\left(\sum p_x S_x\right) - \sum p_x H(S_x), \tag{4.32}$$

причем имеет место строгое неравенство, если среди операторов $p_i S_i$ есть некоммутирующие.

З а д а ч а 26. Если все операторы $p_i S_i$ коммутируют, то равенство достигается для наблюдаемой $M = \{|e_k\rangle\langle e_k|\}$, где $\{e_k\}$ – о.н.б. из общих собственных векторов операторов $p_i S_i$.

Первое, прямое доказательство этой теоремы[7], опирающееся на исследование свойств выпуклости квантовой энтропии, достаточно сложно, поэтому мы приведем здесь лишь его схему. Впоследствии Линдбладом было

[7] А. С. Холево, Некоторые оценки для количества информации, передаваемого квантовым каналом связи. Пробл. передачи информ., 1973, т.9, N3, 3-11.

установлено общее свойство монотонности относительной энтропии (доказательство которого, однако, не менее сложно, но более формально), из которого вытекает и неравенство (4.32), см. часть II.

Доказательство (схема). Прежде всего докажем теорему в случае двух состояний S_0, S_1. Обозначим $S_t = (1-t)S_0 + tS_1$,

$$\chi(t) = H(S_t) - (1-t)H(S_0) - tH(S_1), \quad t \in [0,1]. \tag{4.33}$$

Пусть $M = \{M_y\}$ – произвольная наблюдаемая, $P_t(y) = \operatorname{Tr} S_t M_y = (1-t)P_0(y) + tP_1(y)$ – ее распределение в состоянии S_t и

$$J_M(t) = I_1(p, M),$$

где $p = \{1-t, t\}$. Заметим, что

$$\chi(0) = \chi(1) = 0; \quad I_M(0) = I_M(1) = 0.$$

Мы докажем, что функция $\chi(t)$ "более вогнута", чем $I_M(t)$:

$$\chi(t)'' \le I_M(t)'', \quad t \in [0,1]. \tag{4.34}$$

Отсюда, очевидно, следует

$$\chi(t) \ge I_M(t), \quad t \in [0,1]. \tag{4.35}$$

Положим $D = S_1 - S_0$ и пусть

$$S_t = \sum_k s_k E_k$$

– спектральное разложение оператора S_t. Доказательство следующей леммы[8] опирается на интегральную формулу Коши для матричных функций.

Лемма 8.

$$\chi''(t) = -\sum_{k,j} (\operatorname{Tr} E_k D E_j D) f(s_k, s_j), \quad t > 0, \tag{4.36}$$

где

$$f(a,b) = \frac{\log a - \log b}{a-b}, \quad a \ne b; \quad f(a,a) = a^{-1}. \tag{4.37}$$

Используя элементарное неравенство

$$f(a,b) \ge \frac{2}{a+b}, \quad 0 < a, b \le 1,$$

в котором равенство достигается тогда и только тогда, когда $a = b$, получаем

$$\chi''(t) \le -\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j}, \tag{4.38}$$

[8] ibid.

причем равенство достигается тогда и только тогда, когда $\operatorname{Tr} E_k D E_j D = 0$ для $k \neq j$. Но последнее эквивалентно тому, что $[D, S_t] = 0$, т.е. $[S_0, S_1] = 0$, в силу тождества

$$\operatorname{Tr}[D, S_t]^*[D, S_t] = \sum_{k,j} (s_k - s_j)^2 \operatorname{Tr} E_k D E_j D.$$

Задача 27. Покажите, что оператор

$$L_t = \sum_{k,j} E_k D E_j \frac{2}{s_k + s_j}$$

является решением уравнения

$$S_t \circ L_t \equiv \frac{1}{2}[S_t L_t + L_t S_t] = D,$$

причем

$$\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j} = \operatorname{Tr} D L_t = \operatorname{Tr} S_t L_t^2. \tag{4.39}$$

Оператор L_t является некоммутативным аналогом логарифмической производной семейства S_t, а (4.39)– аналогом информационного количества Фишера в математической статистике.

Из (4.38), (4.39) вытекает, что

$$\chi''(t) \leq -\operatorname{Tr} S_t L_t^2, \tag{4.40}$$

причем равенство достигается тогда и только тогда, когда $[S_0, S_1] = 0$. В частности $\chi''(t) \leq 0$, так что $\chi(t)$ – вогнутая функция на отрезке $[0, 1]$.

Пусть теперь $M = \{M_y\}$ – произвольная наблюдаемая, $P_t(y) = \operatorname{Tr} S_t M_y = (1-t)P_0(y) + tP_1(y)$ – ее распределение в состоянии S_t. Положим также $D(y) = P_1(y) - P_0(y) = \operatorname{Tr} D M_y$. Применяя полученные результаты к диагональной матрице $\operatorname{diag}[P_t(y)]$ в роли состояния S_t и учитывая коммутативность диагональных матриц, получаем вместо (4.40)

$$I''_M(t) = -\sum_y \frac{D(y)^2}{P_t(y)}. \tag{4.41}$$

Доказательство неравенства (4.34). Имеем

$$\begin{aligned} D(y) &= \operatorname{Tr} \sqrt{M_y} S_t \circ L_t \sqrt{M_y} \\ &= \Re \operatorname{Tr} \sqrt{M_y} S_t L_t \sqrt{M_y} \\ &= \Re \operatorname{Tr} \sqrt{M_y} \sqrt{S_t} \sqrt{S_t} L_t \sqrt{M_y} \\ &= \Re \operatorname{Tr} A^* B, \end{aligned}$$

где $A = \sqrt{S_t}\sqrt{M_y}$, $B = \sqrt{S_t} L_t \sqrt{M_y}$. В силу неравенства

$$|\operatorname{Tr} A^* B|^2 \leq \operatorname{Tr} A^* A \operatorname{Tr} B^* B,$$

получаем $D(y) \le \operatorname{Tr} S_t M_y \cdot \operatorname{Tr} L_t S_t L_t M_y$. Подставляя в (4.41), имеем

$$I''_M(t) \ge -\sum_y \operatorname{Tr} L_t S_t L_t M_y = \operatorname{Tr} S_t L_t^2 \ge \chi''(t),$$

причем при $[S_0, S_1] \ne 0$ имеет место строгое неравенство.

Это доказывает утверждение теоремы для случая двух состояний. Случай нескольких состояний $S_x; x = 0, 1, \ldots, k$, сводится к случаю двух состояний путем представления их выпуклой комбинации с распределением $p = \{p_x; x = 0, 1, \ldots, k\}$ в виде последовательности попарных выпуклых комбинаций[9]. □

Теперь докажем *слабое обращение теоремы кодирования*, используя классическое неравенство Фано и квантовую границу информации. Возьмем $N = 2^{NR}$, $R > C_\chi$, и рассмотрим произвольный набор кодовых слов $W = \{w^{(1)}, ..., w^{(N)}\}$ на входе, а на выходе — произвольное разложение единицы $M = \{M_j; j = 0, 1, \ldots, N\}$. Рассмотрим классическую случайную величину X со значениями $1, \ldots, N$ (номер посланного слова), которые имеют равные вероятности $1/N$. На выходе после измерения получим классическую случайную величину Y со значениями $0, 1, \ldots, N$. Взаимная информация равна $I(X;Y) = H(X) - H(X|Y)$, где $H(X) = \log N = nR$ – энтропия равномерного распределения. Условная энтропия оценивается с помощью неравенства Фано: $H(X|\,Y) \le 1 + \mathsf{P}(X \ne Y) \log N$. Таким образом, $\max I(X,Y) \ge nR(1 - p_e(n, 2^{nR})) - 1$ (это повторение доказательства слабого обращения классической теоремы Шеннона).

Из (4.32) вытекает неравенство

$$I(X;Y) \le \max_p \left[H\left(\sum_w p_w S_w \right) - \sum_w p_w H(S_w) \right) \right] = C_\chi^{(n)}.$$

Лемма 9. *Последовательность $C_\chi^{(n)}$ аддитивна: $C_\chi^{(n)} = nC_\chi$.*

Доказательство. Достаточно рассмотреть случай $n = 2$, когда $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, $i \to S_i^1$, $j \to S_j^2$. Надо доказать, что

$$\max_{p_{ij}} \left[H\left(\sum_{ij} p_{ij} S_i^1 \otimes S_j^2 \right) - \sum_{ij} p_{ij} H(S_i^1 \otimes S_j^2) \right] = \max_{p_i^1}[\,] + \max_{p_j^2}[\,].$$

Очевидно, что $\max\limits_{p_{ij}} \ge \max\limits_{p_{ij} = p_i p_j}$, тогда в силу свойства аддитивности квантовой энтропии

$$H\left(\sum_i p_i^1 S_i^1 \otimes \sum_j p_j^2 S_j^2 \right) = H\left(\sum_i p_i^1 S_i^1 \right) + H\left(\sum_j p_j^2 S_j^2 \right),$$

[9] ibid.

откуда $C_\chi^{(n)} \geq nC_\chi$.

Обратное неравенство вытекает из свойства субаддитивности квантовой энтропии (см. п. 7.2), из которого вытекает

$$H\left(\sum p_{ij} S_i^1 \otimes S_j^2\right) \leq H\left(\sum p_i^1 S_i^1\right) + H\left(\sum p_i^2 S_i^2\right),$$

где $p_i^1 = \sum\limits_j p_{ij}$, $p_j^2 = \sum\limits_i p_{ij}$ – маргинальные распределения.□

Окончательно, $nC_\chi \geq nR[1 - p_e(n, 2^{nR})] - 1$, т.е. $p_e(n, 2^{nR}) \geq 1 - C_\chi/R - 1/nR$, и если $R > C_\chi$, то не может быть $p_e(n, 2^{nR}) \to 0$ при $n \to \infty$. Это завершает доказательство слабого обращения. □

4.6 Доказательство прямой теоремы для канала с чистыми состояниями

Доказательство прямого утверждения теоремы кодирования дадим в простейшем случае чистых состояний $S_x = |\psi_x\rangle\langle\psi_x|$[10], когда

$$C_\chi = \max_p H\left(\sum_x p_x S_x\right).$$

Доказательство нетривиально уже в этом случае, тогда как классический аналог этой проблемы тривиален, поскольку чистые состояния с необходимостью ортогональны. *Доказательство.* Пусть $R < C_\chi$. Докажем, что $p_e(n, 2^{nR}) \to 0$. Рассмотрим среднюю вероятность ошибки кода

$$\frac{1}{N}\sum_{j=1}^{N}\Big(1 - \langle\psi_{w^{(j)}}|M_j\psi_{w^{(j)}}\rangle\Big) = \overline{P}_e(W, M),$$

которая зависит от выбора слов W и наблюдаемой M. Для ее минимизации желательно выбрать M_j как можно ближе к $|\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}|$, при этом размерность подпространства, в котором действуют M_j, должна быть по возможности минимальной.

С этой целью произведем сжатие квантовых данных, следуя процедуре, описанной в разделе 4.3. Рассмотрим оператор плотности

$$\sum_{x=1}^{a} p_x|\psi_x\rangle\langle\psi_x| = \overline{S}_p,$$

в котором распределение p выбрано так, что оно максимизирует энтропию, т.е. $H\left(\sum_x p_x S_x\right) = C_\chi$. Фиксируем ε, положим $2\delta = C_\chi - R > 0$ и обозначим E проектор на типичное подпространство $\mathcal{H}^{n,\delta} = E\mathcal{H}^{\otimes n}$ оператора плотности $\overline{S}_p^{\otimes n}$.

[10] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, W. Wootters, "Classical information capacity of a quantum channel," *Phys. Rev. A* **54**, no. 3, 1869-1876 1996.

Положим

$$|\widetilde{\psi}_{w^{(j)}}\rangle = E|\psi_{w^{(j)}}\rangle \in \mathcal{H}^{n,\delta} \tag{4.42}$$

и пусть $G = \sum\limits_{j=1}^{N} |\widetilde{\psi}_{w^{(j)}}\rangle\langle\widetilde{\psi}_{w^{(j)}}|$ — оператор Грама системы (4.42). Оператор Грама всегда можно обратить на подпространстве $\mathcal{H}^{n,\delta}$. Обозначим $G^{-1/2}$ корень из обобщенного обратного к G, равного 0 на ортогональном дополнении к $\mathcal{H}^{n,\delta}$. Для данного набора кодовых слов W введем наблюдаемую M, обобщающую "square-root measurement" (1.18):

$$M_0 = I - E, \quad M_j = G^{-1/2}|\widetilde{\psi}_{w^{(j)}}\rangle\langle\widetilde{\psi}_{w^{(j)}}|G^{-1/2}, \quad j = 1, \ldots, N.$$

Тогда средняя ошибка кода (W, M) равна

$$\overline{P}_e(W; M) = \frac{1}{N}\sum_{j=1}^{N}(1 - |\langle\widetilde{\psi}_{w^{(j)}}|G^{-1/2}|\widetilde{\psi}_{w^{(j)}}\rangle|^2).$$

Используя неравенство $1 - \alpha^2 = (1 - \alpha)(1 + \alpha) \leq 2(1 - \alpha)$, получим

$$\overline{P}_e(W; M) \leq \frac{2}{N}\sum_{j=1}^{N}(1 - \langle\widetilde{\psi}_{w^{(j)}}|G^{-1/2}|\widetilde{\psi}_{w^{(j)}}\rangle)$$

$$= \frac{1}{N}\sum_{j=1}^{N}(2\operatorname{Tr} S_{w^{(j)}} - 2\operatorname{Tr} S_{w^{(j)}}G^{-1/2}). \tag{4.43}$$

Получим теперь удобную оценку для $G^{-1/2}$. Имеем

$$-2x^{-1/2} \leq -3 + x, \quad x \geq 0.$$

Отсюда следует, что

$$-2G^{-1/2} \leq -3E + G.$$

Подставляя в (4.43), получаем

$$\overline{P}_e(W; M) \leq \frac{1}{N}\sum_{j=1}^{N}(2\operatorname{Tr} S_{w^{(j)}} - 3\operatorname{Tr} S_{w^{(j)}}E + \operatorname{Tr} S_{w^{(j)}}G).$$

Учитывая, что

$$G = E\sum_{j=1}^{N} S_{w^{(j)}}E$$

и

$$\operatorname{Tr} S_{w^{(j)}}E = \operatorname{Tr} ES_{w^{(j)}}E \geq \operatorname{Tr}[ES_{w^{(j)}}E]^2,$$

получаем окончательную оценку

$$\overline{P}_e(W; M) \leq \frac{1}{N}\sum_{j=1}^{N}[2\operatorname{Tr} S_{w^{(j)}}(I - E) + \sum_{k\neq j}\operatorname{Tr} ES_{w^{(j)}}ES_{w^{(k)}}E]. \tag{4.44}$$

Теперь применим метод случайных кодов. Надо доказать существование кода, для которого вероятность ошибки стремится к нулю. Идея состоит в том, чтобы рассмотреть случайное распределение на всевозможных словах, тогда минимальная ошибка оценивается сверху средним по ансамблю случайных слов.

Пусть слова $w^{(1)}, \ldots, w^{(N)}$ — независимы, и каждое из них имеет распределение

$$P(w = (i_1, \ldots, i_n)) = p_{i_1} \cdot \ldots \cdot p_{i_n}$$

(буквы берутся также независимо с одинаковым распределением p на алфавите). Усреднение по такому случайному ансамблю слов будет обозначаться $\ll\gg$. Отметим, что

$$\ll S_{w^{(j)}} \gg = \ll |\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}| \gg = \overline{S}_p^{\otimes n}.$$

Тогда, используя одинаковую распределенность, а также независимость слов $w^{(j)}, w^{(k)}$, получаем

$$\ll \overline{P}_e(W, M) \gg \leq 2 \operatorname{Tr} \overline{S}_p^{\otimes n}(I - E) + (N - 1) \operatorname{Tr}(E \overline{S}_p^{\otimes n} E)^2$$

$$\leq 2 \operatorname{Tr} \overline{S}_p^{\otimes n}(I - E) + (N - 1) \|\overline{S}_p^{\otimes n} E\|.$$

Согласно свойствам типичного подпространства, для достаточно больших n

$$\operatorname{Tr} \overline{S}_p^{\otimes n}(I - E) \leq \epsilon, \quad \|\overline{S}_p^{\otimes n} E\| \leq 2^{-n(C_\chi - \delta)}.$$

Так как $N = 2^{nR} \leq 2^{n(C_\chi - 2\delta)}$, и абсолютный минимум не превосходит среднего по ансамблю, то

$$p_e(n, 2^{nR}) \leq \ll \overline{P}_e(W, M) \gg \leq 2\epsilon + 2^{-n\delta} \leq 3\epsilon$$

для достаточно больших n. Итак, $p_e(n, 2^{nR}) \to 0$ при $R < C_\chi$. $\square$

ПРИЛОЖЕНИЕ

Операторы в конечномерном гильбертовом пространстве

Если A – оператор в $\mathcal{H}$, то A^* обозначает оператор, *сопряженный* к A , который определяется равенством

$$\langle\phi|A^*\psi\rangle = \langle A\phi|\psi\rangle \quad \phi,\psi \in \mathcal{H}. \tag{4.45}$$

Оператор A называется *эрмитовым*, если $A = A^*$. (Ортогональным) *проектором* называется эрмитов оператор P, такой, что $P^2 = P$. Областью значений проектора P является подпространство

$$\mathcal{L} = \{\psi : P|\psi\rangle = |\psi\rangle\}.$$

Если $\|\psi\| = 1$, то оператор $|\psi\rangle\langle\psi|$ является проектором на единичный вектор $|\psi\rangle$. Более обще, для любой ортонормированной системы $\{e_i\}_{i\in I}$, оператор $\sum_{i\in I}|e_i\rangle\langle e_i| = P$ является проектором на подпространство, порожденное системой $\{e_i\}_{i\in I}$.

Унитарным называется оператор U , такой что $U^*U = I$; в конечномерном случае это равенство влечет $UU^* = I$. *Частичной изометрией* называется оператор U такой, что $U^*U = P$ является проектором; в этом случае $UU^* = Q$ также есть проектор. Оператор U отображает область значений P на область значений Q *изометрично*, то есть сохраняя скалярное произведение и нормы векторов.

Т е о р е м а 18*[Спектральное разложение]Для любого эрмитова оператора A существует ортонормированный базис из собственных векторов, которым отвечают вещественные собственные значения a_i, так что*

$$A = \sum_{i=1}^{d} a_i|e_i\rangle\langle e_i|. \tag{4.46}$$

Другая полезная форма спектрального разложения получается, если рассмотреть *различные* собственные значения $\{a\}$ и соответствующие им спектральные проекторы

$$E_a = \sum_{i:a_i=a} |e_i\rangle\langle e_i|.$$

Набор различных собственных значений $\mathrm{spec}(A) = \{a\}$ называется *спектром* оператора A. В этих обозначениях

$$A = \sum_{a \in \mathrm{spec}(A)} aE_a. \tag{4.47}$$

Такое представление единственно с точностью до порядка перечисления собственных значений. Набор проекторов $\{E_a\}$ образует *ортогональное разложение единицы*:

$$E_a E_{a'} = \delta_{aa'} E_a, \qquad \sum_{a \in \mathrm{spec}(A)} E_a = I. \tag{4.48}$$

Гильбертово пространство $\mathcal{H}$ разлагается в прямую ортогональную сумму областей значений проекторов $\{E_a\}$, на которых A действует как умножение на число a.

Эрмитов оператор A называется *положительным*, $A \geq 0$, если $\langle\psi|A\psi\rangle \geq 0$ для любого $\psi \in \mathcal{H}$. Собственные значения положительного оператора неотрицательны: $a \geq 0$ для $a \in \mathrm{spec}(A)$.

Оператор является положительным тогда и только тогда, когда он может быть представлен в виде $A = B^*B$ для некоторого оператора B. Для любого положительного оператора A существует единственный положительный квадратный корень $C = \sqrt{A} = A^{1/2}$, такой что $C^2 = A$.

Для любого эрмитова оператора A имеет место разложение

$$A = A_+ - A_-, \tag{4.49}$$

где $A_+ = \sum_{a>0} aE_a$, $A_- = -\sum_{a<0} aE_a$ – положительные операторы, называемые *положительной* и *отрицательной* частями оператора A.

Теорема 19*[Полярное разложение] Любой оператор A в $\mathcal{H}$ может быть представлен в виде*

$$A = U|A| = |A^*|U, \tag{4.50}$$

где $|A| = \sqrt{A^*A}$ – положительный оператор, а U – унитарный оператор.

Носителем $\mathrm{supp}\ A$ положительного оператора A называется его собственное подпространство, соответствующее положительным собственным значениям. Унитарный оператор в полярном разложении определяется единственным образом только на $\mathrm{supp}\ |A|$.

В вещественном гильбертовом пространстве эрмитовы операторы заменяются на симметричные, унитарные – на ортогональные, причем определения формально остаются теми же. Полярное разложение также имеет место, причем $|A|$ - симметричный положительный, а U – ортогональный оператор.

След оператора T определяется соотношением

$$\mathrm{Tr}\, T = \sum_{i=1}^{d} \langle e_i|Te_i\rangle, \tag{4.51}$$

где $\{e_i\}$ – произвольный ортонормированный базис.

З а д а ч а 28. Покажите, что это определение не зависит от выбора базиса и что

$$\operatorname{Tr} A^* = \overline{\operatorname{Tr} A}, \quad \operatorname{Tr} AB = \operatorname{Tr} BA. \tag{4.52}$$

Покажите, что

$$\operatorname{Tr} |\psi\rangle\langle\varphi| A = \langle\varphi|A\psi\rangle. \tag{4.53}$$

Покажите, что для $A, B \geq 0$ выполнено

$$\operatorname{Tr} AB \geq 0 \tag{4.54}$$

и равенство нулю имеет место тогда и только тогда, когда $AB = 0$.

Литература

[1] П.А.М. Дирак, Принципы квантовой механики. Наука, 1970.

[2] А. Ю. Китаев, Квантовые вычисления: алгоритмы и исправление ошибок *УМН* т. 52, N6, 53-112, 1997.

[3] А. Китаев, А. Шень, М. Вялый, Классические и квантовые вычисления. МЦНМО 1999.

[4] Дж. фон Нейман, Математические основы квантовой механики. Наука 1964.

[5] М. А. Нильсен, И. Чанг, Квантовые вычисления и квантовая информация, пер. с англ., М.: Мир, 2006.

[6] Л. Д. Фаддеев, О. Я. Якубовский, Лекции по квантовой механике для студентов-математиков. М.-Ижевск: РХД 2001.

[7] Р. Фейнман, Р. Лейтон, М. Сэндс, Фейнмановские лекции по физике. 8. Квантовая механика. Мир, 1986.

[8] К. Хелстром, Квантовая теория проверки гипотез и оценивания. М.: Мир, 1978.

[9] А. С. Холево, Вероятностные и статистические аспекты квантовой теории. 2-е изд., М.-Ижевск: ИКИ, 2003.

[10] А. С. Холево, Квантовые системы, каналы, информация. М.: МЦНМО 2010. Ч. I, II.

[11] С.И. Чечета, Введение в дискретную теорию информации и кодирования, М.: МЦНМО 2011.