МАТЕМАТИЧЕСКИЕ ОСНОВЫ КВАНТОВОЙ ИНФОРМАТИКИ

А. С. Холево

Москва – 2015

Оглавление

Предисловие

Квантовая информатика – быстро развивающаяся научная дисциплина, которая изучает общие закономерности передачи, хранения и преобразования информации в системах, подчиняющихся законам квантовой механики. Квантовая информатика использует математический аппарат матричного и операторного анализа, некоммутативной теории вероятностей и статистики для исследования потенциальных возможностей таких систем. Немаловажным попутным результатом является существенное прояснение логической структуры квантовой механики, ее оснований и соотношения с реальностью. Известно, что сознательное усвоение основ квантовой теории представляет собой трудности, требующие значительных интеллектуальных усилий, намного превосходящих те, которые требуются, скажем, для перехода от детерминистического к вероятностному описанию классических систем. Последний предполагает серьезную смену парадигмы, но не отменяет использование наглядных механических моделей, взятых из макроскопического мира, доступного непосредственному человеческому восприятию. Смена парадигмы при переходе от классического статистического описания к квантовой теории требует выработки особого рода интуиции, опирающейся на более абстрактные математические модели, поскольку механистические представления приводят к кричащим противоречиям с экспериментальными фактами микромира. В этом смысле, несколько перефразируя знаменитое высказывание Р. Фейнмана, "квантовую теорию понять нельзя, но к ней можно привыкнуть". [1].

Настоящие лекции предлагают путь к пониманию статистической структуры квантовой теории (а именно это представляет собой наибольшую познавательную трудность), доступный заинтересованному читателю, владеющему основными общематематическими дисциплинами, и не требующий глубоких познаний в физике. Это может быть специалист по компьютерным наукам или защите информации, желающий понять принципы квантовых вычислений или квантовой криптографии, математик, стремящийся неформально освоить парадигму квантовой теории и найти новые задачи и постановки вопросов, или физик, жаждущий найти свежий взгляд на, казалось бы, уже пройденные вещи. Изложение ведется на уровне конечномер-

[1] На самом деле, именно Фейнман сделал очень много для понимания квантовой теории: см. в частности его обсуждение спиновых систем в книге [9].

ных моделей: с одной стороны, почти все особенности квантового описания проявляются, причем наиболее выпукло, уже на этом уровне, с другой – большая часть результатов квантовой теории информации (математически совершенно нетривиальных) относится именно к таким моделям. Конечномерность и линейная алгебра в квантовой информатике играют ту же фундаментальную роль, что конечность и комбинаторика в классической – размерность квантовой системы определяет ее потенциальный информационный ресурс. Также существенным для наших целей является владение основами теории вероятностей и математической статистики, поскольку статистическое описание (конечных) классических систем служит для нас как отправной точкой, так и эталоном для сопоставления. С одной стороны, мы опираемся на глубокое структурное родство статистического описания классических и квантовых систем (фактически – на конечномерный вариант известного "алгебраического подхода" [8]), с другой – стараемся вычленить их наиболее базовые принципиальные различия.

В первой части курса мы опираемся на "стандартную статистическую модель квантовой системы". Этот термин, не вполне стандартный, означает обычную аксиоматику квантовой механики [2], [6], [8], рассматриваемую, однако, под углом ее сопоставления с классическим статистическим описанием. Последовательное развитие такой точки зрения привело во второй половине XX века к обобщенной статистической модели квантовой теории [11], [12], в которой наблюдаемые (вообще говоря, "нечеткие") описываются вероятностными операторно-значными мерами, а эволюции (вообще говоря, необратимые) – вполне положительными отображениями. Начальные сведения об этих понятиях, которые лежат в основе математических моделей квантовых информационных систем при наличии шумов и ошибок и имеют многочисленные приложения в квантовой информатике, излагаются во второй части курса (подробнее см. в монографиях [13], [7]).

Часть I

Глава 1

Стандартная статистическая модель квантовой системы

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 [6], [11]).

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 квантовой информатики (впрочем, в квантовой информатике большое внимание привлекают и "системы с непрерывными переменными", которые описываются бесконечномерными гильбертовыми 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, $d = |\Omega|$. (Статистическим) *состоянием* называется распределение вероятностей на Ω:

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

Случайная величина – это вещественная функция на Ω :

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

Математическое ожидание случайной величины X в состоянии P дается формулой

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

В реальности эти понятия и формулы используются следующим образом: состояние P описывает "статистический ансамбль", т.е. случайную выборку большого количества независимых, одинаково распределенных экземпляров системы, над каждым из которых соответствующим прибором производится измерение наблюдаемой величины X. При неограниченном увеличении количества экземпляров среднее значение результатов измерения по всему ансамблю приближается к теоретическому значению $\mathsf{M}_P X$.

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

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

где Tr – след матрицы.

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

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

Частным случаем является классическое состояние, представленное диагональной матрицей $S = P$. Вещественная *квантовая наблюдаемая* задается эрмитовой матрицей

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

Частным случаем является классическая случайная величина, представленная диагональной матрицей X. *Математическое ожидание* наблюдаемой X в состоянии S дается *статистическим постулатом* Борна-фон Неймана [6]:

$$\mathsf{M}_S X = \operatorname{Tr} SX. \tag{1.1}$$

Если матрицы S, X диагональны, мы формально возвращаемся к классическому описанию.

При таком подходе обнаруживается аналогия в статистическом описании классических и квантовых систем: сначала некоторым прибором приготавливается статистическое состояние (P или S), затем другим прибором производится измерение случайной величины или наблюдаемой X. Как приготовление, так и измерение могут нести в себе случайность, в результате чего исход измерения случаен, причем его математическое ожидание задается формулой (1.1). В любом случае, практически речь идет о большом большом числе независимых, одинаково организованных повторений опыта. В случае квантовых систем это обычно бывают эксперименты с пучками, состоящими из большого числа частиц, при которых статистика набирается одновременно.

При этом для каждой квантовой величины – состояния или наблюдаемой, представляемой эрмитовой матрицей – существует свой 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:

Теорема 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 приборов. Одновременное использование различных приборов, соответствующих измерениям разных наблюдаемых, может быть взаимно исключающим (несмотря на то, что они применяются к одинаково приготовленному микрообъекту). Такие измерения называются взаимно *дополнительными*. Аналогичные соображения относятся и к различным способам приготовления 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

Комплексные $d \times d$-матрицы можно рассматривать как линейные операторы, действующие в пространстве $\mathbb{C}^d$; в дальнейшем будет удобнее иметь

дело с операторами, действующими в конечномерном гильбертовом пространстве (хотя при этом мы и не выигрываем в общности, такой подход более геометричен и полезен с точки зрения бесконечномерных обобщений).

Пусть $\mathcal{H}$ - комплексное векторное пространство размерности $\dim \mathcal{H} = d < \infty$, со скалярным произведением $\langle\phi|\psi\rangle, \phi, \psi \in \mathcal{H}$ [5]; 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 [2]: вектор ψ из $\mathcal{H}$ (в случае $\mathbb{C}^d$ ему соответствует вектор-столбец) обычно будет обозначаться $|\psi\rangle$; соответственно, $\langle\psi|$ обозначает вектор сопряженного пространства (эрмитово сопряженную 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 $\{e_i\}_{i=1,\dots,d}$ – 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 *соотношению полноты*

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

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

$$\langle\phi|\psi\rangle = \sum_{i=1}^{d} \langle\phi|e_i\rangle\langle e_i|\psi\rangle = \sum_{i=1}^{d} \overline{\langle e_i|\phi\rangle}\langle e_i|\psi\rangle$$

показывает, что отображение

$$|\psi\rangle \longrightarrow \begin{bmatrix} \langle e_1|\psi\rangle \\ \vdots \\ \langle e_d|\psi\rangle \end{bmatrix}$$

является изометрическим изоморфизмом (взаимно-однозначным линейным отображением, сохраняющим скалярные произведения) гильбертовых пространств $\mathcal{H}$ и $\mathbb{C}^d$. При этом базис $\{e_i\}_{i=1,\dots,d}$ переходит в *вычислительный базис* в $\mathbb{C}^d$:

$$|e_1\rangle \longrightarrow \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots \quad , |e_d\rangle \longrightarrow \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}.$$

З а д а ч а 1. Воспользовавшись соотношением полноты для о.н.б. $\{e_i\}_{i=1,..}$, запишите матричное представление для операторов в $\mathcal{H}$, аналогичное представлению векторов (1.2). Покажите, что матрица оператора A в этом базисе есть $[\langle e_i|A|e_j\rangle]_{i,j=1,\ldots,d}$.

Фундаментальное отличие комплексного гильбертова пространства от вещественного (евклидова) пространства проявляется в наличии *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 квадратичной форме $\beta(\psi,\psi)$, $\mathcal{H}$ (в вещественном случае подобное восстановление возможно лишь для симметричных форм). Благодаря этому, например, для доказательства операторного равенства $A=B$ достаточно установить равенство соответствующих квадратичных форм $\langle\psi|A|\psi\rangle=\langle\psi|B|\psi\rangle, \psi\in\mathcal{H}$.

Необходимый минимум сведений об операторах в конечномерном гильбертовом пространстве приведен в следующем разделе.

1.3 Операторы в конечномерном гильбертовом пространстве

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

$$\langle\phi|A^*\psi\rangle=\overline{\langle\psi|A\phi\rangle}\quad\phi,\psi\in\mathcal{H}. \tag{1.5}$$

Если $[a_{jk}]$ матрица оператора A в ортонормированном базисе, то матрицей оператора A^* в том же базисе является $[\overline{a}_{kj}]$. 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.

Если

$$A|e\rangle = a|e\rangle, \quad |e\rangle \neq 0, \tag{1.6}$$

то a называется собственным значением оператора A, а $|e\rangle$ – соответствующим собственным вектором. Собственные значения находятся как корни характеристического уравнения

$$\det(A - aI) = 0,$$

после чего соответствующие собственные векторы находятся как ненулевые решения однородной системы уравнений (1.6).

Теорема 2. *Для любого эрмитова оператора A существует ортонормированный базис из собственных векторов, которым отвечают вещественные собственные значения a_i.*

Умножая уравнение (1.6), т.е. $A|e\rangle_j = a_j|e\rangle_j$ на $\langle e|_j$, суммируя по j и учитывая соотношение полноты (1.3), получаем *спектральное разложение* оператора A

$$A = \sum_{i=1}^{d} a_j |e_i\rangle\langle e_i|. \tag{1.7}$$

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

$$A = \sum_{a\in\mathrm{spec}(A)} aE_a. \tag{1.8}$$

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

$$E_a E_{a'} = \delta_{aa'} E_a, \qquad \sum_{a\in\mathrm{spec}(A)} E_a = I. \tag{1.9}$$

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 $a \in \mathrm{spec}(A)$.

Практически удобным условием положительности является *критерий Сильвестра*: пусть $[a_{jk}]_{j,k=1,\dots,d}$ матрица оператора A в каком-либо базисе; A положителен тогда и только тогда, когда

$$\det[a_{jk}]_{j,k=1,\dots,m} \geq 0; \quad m = 1, \dots, d.$$

З а д а ч а 2. 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

$$A = A_+ - A_-, \tag{1.10}$$

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

След оператора T определяется соотношением

$$\operatorname{Tr} T = \sum_{i=1}^{d} \langle e_i | T e_i \rangle, \tag{1.11}$$

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

З а д а ч а 3. Покажите, что это определение не зависит от выбора базиса и что

$$\operatorname{Tr} A^* = \overline{\operatorname{Tr} A}, \quad \operatorname{Tr} AB = \operatorname{Tr} BA. \tag{1.12}$$

Покажите, что

$$\operatorname{Tr} |\psi\rangle\langle\varphi| A = \langle\varphi|A\psi\rangle. \tag{1.13}$$

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

$$\operatorname{Tr} AB \geq 0 \tag{1.14}$$

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

1.4 Выпуклость

Подмножество $\mathfrak{S}$ вещественного линейного пространства называется *выпуклым*, если для любого конечного набора точек $\{S_j\} \subset \mathfrak{S}$ и любого распределения вероятностей $\{p_j\}$ *выпуклая комбинация* $S = \sum_j p_j S_j$ принадлежит $\mathfrak{S}$ (достаточно потребовать выполнения указанного условия только для наборов из двух точек, то есть чтобы множество $\mathfrak{S}$ вместе с любыми двумя точками содержало и соединяющий их отрезок). В выпуклых множествах особо важны *крайние точки*, не представимые в виде нетривиальной выпуклой комбинации других точек. Это эквивалентно утверждению, что из $S = pS_1 + (1-p)S_2$, $0 < p < 1$, следует $S = S_1 = S_2$ т. е. что ни один

отрезок в $\mathfrak{S}$ не содержит S в качестве своей внутренней точки. Мы обозначаем $\mathrm{extr}(\mathfrak{S})$ множество всех крайних точек выпуклого множества $\mathfrak{S}$. Имеет место следующий общий результат:

Теорема 3 (Каратеодори). *Пусть $\mathfrak{S}$ – компактное выпуклое подмножество $\mathbb{R}^n$, тогда любая точка $S \in \mathfrak{S}$ может быть представлена в виде выпуклой комбинации не более чем $n+1$ крайних точек $S_j \in extr(\mathfrak{S})$:*

$$S = \sum_{j=1}^{n+1} p_j S_j, \quad S_j \in \mathrm{extr}(\mathfrak{S}). \tag{1.15}$$

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

Вещественная функция $\mathcal{F}$, определенная на выпуклом подмножестве $\mathfrak{S}$ конечномерного линейного пространства называется *выпуклой* (*вогнутой*), если

$$\mathcal{F}(\sum_j p_j S_j) \leq (\geq) \sum_j p_j \mathcal{F}(S_j),$$

для любой выпуклой комбинации точек $S_j \in \mathfrak{S}$. Функция называется *аффинной*, если она как выпукла, так и вогнута, т. е.

$$\mathcal{F}(\sum_j p_j S_j) = \sum_j p_j \mathcal{F}(S_j).$$

Задача 4. Непрерывная выпуклая (в частности, аффинная) функция на компактном (ограниченном и замкнутом) выпуклом множестве $\mathfrak{S}$ достигает своего максимума в крайней точке этого множества.

1.5 Квантовые состояния

Состояние квантово-механической системы, представляющее собой статистический ансамбль большого количества независимых, одинаково приготовленных экземпляров системы (например, пучок частиц, вылетающих из ускорителя), описывается *оператором плотности* (матрицей плотности в фиксированном базисе), т.е. оператором S в $\mathcal{H}$, удовлетворяющим условиям $S \geq 0$, $\mathrm{Tr}\, S = 1$. 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. 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 4. *Крайние точки множества квантовых состояний $\mathcal{S}(\mathcal{H})$, называемые чистыми состояниями, суть (одномерные) проекторы, $S^2 = S$, и только они.*

Таким образом, оператор плотности чистого состояния имеет вид $S = |\psi\rangle\langle\psi|$, где $|\psi\rangle$ — единичный *вектор состояния*, определенный с точностью до фазового множителя, по модулю равного единице. Если $\{e_k\}$ — фиксированный ортонормированный базис, то $\psi(k) = \langle e_k|\psi\rangle$ и есть знаменитая пси-функция квантовой механики (в k-представлении).

Доказательство. Рассмотрим спектральное разложение эрмитова оператора S

$$S = \sum_{j=1}^{d} s_j |e_j\rangle\langle e_j|, \quad s_j \geq 0, \sum s_j = 1, \tag{1.16}$$

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:

$$pS_1(I - S_1) + (1-p)S_2(I - S_2) + p(1-p)(S_1 - S_2)^2 = S - S^2 = 0. \tag{1.17}$$

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. □

Задача 5. Доказать, что если $\dim \mathcal{H} = d$, то $\mathcal{S}(\mathcal{H})$ погружается в вещественное пространство размерности $n = d^2 - 1$.

Спектральное разложение (1.16) 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).

1.6 Наблюдаемые и функции от них. Распределение вероятностей квантовой наблюдаемой

Пусть – вещественная наблюдаемая, задаваемая эрмитовым оператором (матрицей) X. Рассмотрим ее спектральное разложение

$$X = \sum_j x_j |e_j\rangle\langle e_j|, \tag{1.18}$$

где x_j – собственные числа, $\{e_j\}$ – ортонормированный базис из собственных векторов. Среди собственных чисел могут быть совпадающие; обозначим

$$E_x = \sum_{j:x_j=x} |e_j\rangle\langle e_j|$$

ортогональный проектор на собственное подпространство оператора X, отвечающее собственному числу x. Множество различных собственных чисел $\mathcal{X}$ образует спектр оператора X, т.е. множество возможных значений вещественной наблюдаемой X. Спектральное разложение (1.18) можно переписать в виде

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

Семейство проекционных операторов $\{E_x\}_{x\in\mathcal{X}}$ называется спектральной мерой оператора X.

З а д а ч а 6. Покажите, что $E_x^2 = E_x$, $E_xE_y = 0$, для всех $x, y \in \mathcal{X}, x \neq y$. Используя этот факт, покажите, что

$$X^k = \sum_{x\in\mathcal{X}} x^k E_x, \quad k = 2, \ldots \tag{1.19}$$

Отсюда получаем, что для любой вещественной функции $f(x)$, определенной на спектре $\mathcal{X}$,

$$f(X) = \sum_j f(x_j)|e_j\rangle\langle e_j| = \sum_{x\in\mathcal{X}} f(x)E_x. \tag{1.20}$$

В самом деле, любую такую функцию можно приблизить многочленами на $\mathcal{X}$, а для многочленов (1.20) следует из (1.19). Математическое ожидание наблюдаемой $f(X)$ равно

$$\mathsf{M}_S f(X) = \operatorname{Tr} Sf(X) = \sum_{x\in\mathcal{X}} f(x) \operatorname{Tr} SE_x.$$

Таким образом, измерение наблюдаемой $f(X)$ можно трактовать как квантовое измерение X с последующим классическим вычислением функции $f(x)$ от результата измерения x.

Полагая $f(y) = \delta_{xy}$, получаем фундаментальную формулу

$$\mathsf{P}_S(X = x) = \operatorname{Tr} SE_x, \tag{1.21}$$

дающую *вероятность результата x при измерении наблюдаемой X в состоянии S*. Эта формула является наиболее полным выражением *статистического постулата* Борна-фон Неймана.

З а д а ч а 7. Докажите формулу для дисперсии наблюдаемой X

$$\mathsf{D}_S(X) = \operatorname{Tr} S(X - \mathsf{M}_S(X))^2 = \operatorname{Tr} SX^2 - (\mathsf{M}_S(X))^2.$$

З а д а ч а 8. Найдите математическое ожидание, дисперсию и распределение вероятностей наблюдаемой X в состоянии S, где

$$S = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad X = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Для любого оператора плотности S *мера чистоты* определяется соотношением

$$P(S) = \operatorname{Tr} S^2 = \sum_j s_j^2, \tag{1.22}$$

где s_j – собственные значения оператора S. *Энтропия* состояния S определяется соотношением

$$H(S) = -\operatorname{Tr} S \log S = -\sum_j s_j \log s_j, \tag{1.23}$$

с соглашением $0 \log 0 = 0$, а логарифм берется по фиксированному основанию $a > 1$. В теории информации удобно использовать двоичный логарифм ($a = 2$), при этом единицей измерения является *бит*. В статистической механике обычно используется натуральный логарифм, единицей измерения является *нат*, так что 1 нат=$\ln 2$ бит.

Задача 9. Покажите, что $1/d \le P(S) \le 1$, причем $P(S) = 1$ тогда и только тогда, когда состояние S – чистое. $P(S) = 1/d$ тогда и только тогда, когда состояние S – хаотическое.

Покажите, что $0 \le H(S) \le \log d$, причем $H(S) = 0$ тогда и только тогда, когда состояние S – чистое. $H(S) = \log d$ тогда и только тогда, когда состояние S – хаотическое.

1.7 Двухуровневые системы. Квантовый бит

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, $\dim \mathcal{H} = 2$. Можно считать, что $\mathcal{H} = \mathbb{C}^2$. Будем использовать вычислительный базис: $|0\rangle = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$, $|1\rangle = \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. Удобно ввести базис Паули в вещественном пространстве эрмитовых 2×2-матриц:

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

Задача 10. Матрицы Паули подчиняются следующей таблице умножения

$$\sigma_x^2 = I, \quad \sigma_y^2 = I, \quad \sigma_z^2 = I,$$

$$\sigma_x\sigma_y = i\sigma_z, \quad \sigma_y\sigma_z = i\sigma_x, \quad \sigma_z\sigma_x = i\sigma_y. \tag{1.24}$$

Всякий оператор плотности $S \in \mathcal{S}(\mathcal{H})$ представляется как

$$S(\vec{a}) = \frac{1}{2}\begin{bmatrix} 1+a_z & a_x - ia_y \\ a_x + ia_y & 1-a_z \end{bmatrix} = \frac{1}{2}(I + \sigma(\vec{a})), \tag{1.25}$$

где $\vec{a} = (a_x, a_y, a_z)$ и $\sigma(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z$. Условие $\det S \ge 0$ накладывает следующее ограничение на *параметры Стокса* (a_x, a_y, a_z):

$$|\vec{a}|^2 \equiv a_x^2 + a_y^2 + a_z^2 \le 1.$$

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

Чистые состояния характеризуются условием $|\vec{a}| = 1$ 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

$$|\vec{a}\rangle = \begin{bmatrix} \cos(\theta/2)\, e^{-i\phi/2} \\ \sin(\theta/2)\, e^{i\phi/2} \end{bmatrix}. \tag{1.26}$$

Таким образом,

$$|\vec{a}\rangle\langle\vec{a}| = \frac{1}{2}(I + \sigma(\vec{a})), \tag{1.27}$$

где

$$\sigma(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z = \begin{bmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{bmatrix}, \quad |\vec{a}| = 1, \tag{1.28}$$

эрмитов унитарный оператор со свойствами $\sigma(\vec{a})^2 = I$, $\operatorname{Tr}\sigma(\vec{a}) = 0$.

В квантовых системах со спином $1/2$ состояние (1.27) описывает ансамбль (пучок частиц) со спином вдоль направления $\vec{a}$ [1]. *Хаотическим* является смешанное состояние с $a_x = a_y = a_z = 0$ (все направления спинов равновероятны), описываемое оператором плотности $S = I/2$.

Из (1.27) получается спектральное разложение эрмитова оператора

$$\sigma(\vec{a}) = |\vec{a}\rangle\langle\vec{a}| - |-\vec{a}\rangle\langle-\vec{a}|, \quad |\vec{a}| = 1. \tag{1.29}$$

Собственные векторы $|\pm\vec{a}\rangle$, отвечающие собственным значениям ±1, образуют о.н.б. Наблюдаемая (1.29), принимающая значения ±1, описывает *проекцию спина на направление* $\vec{a}$. Таким образом, спин – это векторный оператор с некоммутирующими компонентами $\sigma_x, \sigma_y, \sigma_z$. Эксперимент Штерна-Герлаха, описывающий приготовление состояний $S(\vec{a})$ и измерение наблюдаемых $\sigma(\vec{b})$ с помощью внешнего неоднородного магнитного поля, детально описан в лекциях Фейнмана [9].

З а д а ч а 11. Пользуясь соотношением (1.29), получите спектральное разложение оператора плотности (1.25):

$$S(\vec{a}) = \frac{1 + |\vec{a}|}{2}|\vec{a}'\rangle\langle\vec{a}'| + \frac{1 - |\vec{a}|}{2}|-\vec{a}'\rangle\langle-\vec{a}'|, \quad \vec{a}' = \vec{a}/|\vec{a}|.$$

Таким образом, собственные значения оператора плотности $S(\vec{a})$ суть $\frac{1\pm|\vec{a}|}{2}$. Отсюда, пользуясь определениями (1.22), (1.23), получаем меру чистоты и энтропию состояния $S(\vec{a})$:

$$P(S(\vec{a})) = \frac{1}{2}\left(1 + |\vec{a}|^2\right); \qquad H(S(\vec{a})) = h\left(\frac{1 + |\vec{a}|}{2}|\right),$$

где $h(p) = -p\log p - (1-p)\log(1-p)$ – энтропия случайного бита. Обе функции постоянны на концентрических сферах $|\vec{a}| = r$, причем мера чистоты является выпуклой функцией параметра $r, 0 \le r \le 1$, а энтропия – вогнутой функцией, принимающей минимальное значение 0 на границе шара и максимальное 1 в его центре.

З а д а ч а 12. Докажите формулу

$$\sigma(\vec{a}_1)\sigma(\vec{a}_2) = (\vec{a}_1 \cdot \vec{a}_2)I + i\sigma(\vec{a}_1 \times \vec{a}_2). \tag{1.30}$$

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

Задача 13. 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

$$\operatorname{Tr} S(\vec{a})\sigma(\vec{b}) = \vec{a} \cdot \vec{b} \tag{1.31}$$

Задача 14. Пользуясь спектральным разложением (1.29), покажите, что измерение наблюдаемой $\sigma(\vec{b})$, $|\vec{b}| = 1$ в состоянии $S(\vec{a})$, $|\vec{a}| \leq 1$ дает значения ± 1 с вероятностями

$$\mathsf{P}_{S(\vec{a})}(\sigma(\vec{b}) = \pm 1) = \frac{1}{2}\left(1 \pm \vec{a} \cdot \vec{b}\right). \tag{1.32}$$

1.8 Совместимые наблюдаемые

Коммутатором операторов X, Y называется оператор $[X, Y] = XY - YX$. Операторы X, Y коммутируют (перестановочны), если $[X, Y] = 0$.

Вещественные наблюдаемые X, Y называются *совместимыми*, если найдется вещественная наблюдаемая Z, такая что $X = f(Z), Y = g(Z)$, где f, g – вещественные функции. Другими словами, наблюдаемые X, Y могут быть измерены в одном эксперименте (измерения Z), с последующим пересчетом результатов измерения. Если X, Y совместимы, то однозначно определены наблюдаемые вида $h(X, Y)$, где h – произвольная функция двух переменных. При этом

$$\mathsf{M}_S h(X, Y) = \sum_{x,y} h(x, y)\mathsf{P}_S(X = x, Y = y), \tag{1.33}$$

где

$$\mathsf{P}_S(X = x, Y = y) = \operatorname{Tr} SE_x F_y = \operatorname{Tr} SF_y E_x \tag{1.34}$$

совместное распределение вероятностей наблюдаемых X, Y. Здесь $X = \sum_x xE_x, Y = \sum_y yF_y$ – спектральные разложения операторов X, Y.

Подобным образом можно определить совместное измерение и распределение вероятностей для любого конечного набора совместимых наблюдаемых.

Теорема 5. *Следующие утверждения эквивалентны:*

(i) X, Y совместимы;

(ii) $[X, Y] = 0$;

(iii) X, Y одновременно диагонализуемы, т.е. имеют общий базис из собственных векторов.

Доказательство.

(i) $\Rightarrow$ *(ii)*. Очевидно, что $[f(Z), g(Z)] = 0$ для многочленов f, g, а произвольные функции аппроксимируются многочленами.

(ii) $\Leftrightarrow$ *(iii)*. Это было доказано в теореме 1.

(iii) $\Rightarrow$ *(ii)*. Пусть $\{e_j\}$ – общий базис из собственных векторов операторов X, Y, и пусть $\{x_j, y_j\}$ – соответствующие собственные числа. Тогда

искомый оператор Z – оператор с теми же собственными векторами и собственными числами $\{j\}$. □

Задача 15. Покажите, что вещественная наблюдаемая, совместимая со всеми квантовыми наблюдаемыми, является постоянной величиной, то есть задается оператором, кратным единичному оператору.

Задача 16. Обобщите соотношения (1.33), (1.34) на случай n попарно совместимых наблюдаемых.

1.9 Соотношение неопределенностей

Пусть X, Y – два оператора. Тогда

$$XY = X \circ Y + \frac{1}{2}[X, Y],$$

где $X \circ Y = \frac{1}{2}(XY + YX)$ – симметризованное или *йорданово* произведение операторов X, Y.

Пусть S – некоторое состояние. Для произвольного набора $X = [X_1, \dots, X$ четких вещественных наблюдаемых положим $X_j^0 = X_j - I\mathsf{M}_S X_j$ и введем две вещественные матрицы: симметричную *матрицу ковариаций*

$$\mathsf{D}_S(X) = \left[\operatorname{Tr} SX_j^0 \circ X_k^0\right]_{j,k=1,\dots,n}, \tag{1.35}$$

и кососимметричную *коммутационную матрицу*

$$\mathsf{C}_S(X) = \left[i \operatorname{Tr} S[X_j, X_k]\right]_{j,k=1,\dots,n} = \left[i \operatorname{Tr} S[X_j^0, X_k^0]\right]_{j,k=1,\dots,n}. \tag{1.36}$$

Имеем

$$\mathsf{D}_S(X) \geq \frac{i}{2}\mathsf{C}_S(X) \tag{1.37}$$

в смысле неравенства между комплексными эрмитовыми матрицами. Действительно, эрмитова матрица

$$\mathsf{D}_S(X) - \frac{i}{2}\mathsf{C}_S(X) = \left[\operatorname{Tr} SX_j^0 X_k^0\right]_{j,k=1,\dots,n}$$

положительно определена, поскольку для произвольных $c_j \in \mathbb{C}$

$$\sum_{j,k=1}^{n} \bar{c}_j c_k \operatorname{Tr} SX_j^0 X_k^0 = \operatorname{Tr} SZ^* Z \geq 0,$$

где $Z = \sum_{j=1}^{n} c_j X_j$.

Для двух наблюдаемых $X_1 = X$ и $X_2 = Y$ неравенство (1.37) эквивалентно соотношению неопределенностей Шредингера - Робертсона

$$\mathsf{D}_S(X)\mathsf{D}_S(Y) \geq \{\mathsf{M}_S(X - I\mathsf{M}_S(X)) \circ (Y - I\mathsf{M}_S(Y))\}^2 + \frac{1}{4}|\mathsf{M}_S[X, Y]|^2, \tag{1.38}$$

где

$$\mathsf{D}_S(X) = \operatorname{Tr} S(X - I\mathsf{M}_S(X))^2 \tag{1.39}$$

– дисперсия вещественной наблюдаемой X в состоянии S. Если X, Y совместимые наблюдаемые, то величина

$$\mathsf{M}_S(X - I\mathsf{M}_S(X)) \circ (Y - I\mathsf{M}_S(Y)) \tag{1.40}$$

представляют собой *ковариацию* X, Y в состоянии S; в этом случае $[X, Y] = 0$ и (1.38) превращается в неравенство Коши-Буняковского для ковариации случайных величин. Если же X, Y несовместимы, то X, Y неизмеримы в одном эксперименте, и дисперсии $\mathsf{D}_S(X), \mathsf{D}_S(Y)$ в соотношении неопределенностей относятся к двум различным измерениям, произведенными над разными представителями одного статистического ансамбля.

З а д а ч а 17. Докажите некоммутативное неравенство Коши- Буняковского

$$|\operatorname{Tr} SX^*Y|^2 \leq \operatorname{Tr} SX^*X \ \operatorname{Tr} SY^*Y, \tag{1.41}$$

для произвольного состояния S и операторов X, Y в $\mathcal{H}$.

1.10 Апостериорное состояние. Последовательные измерения

Рассмотрим сначала классическую систему с фазовым пространством Ω, которая находится в (статистическом) состоянии P. Предположим, что производится измерение вещественной случайной величины X. Что можно сказать о состоянии системы после измерения, в результате которого получено значение x? Оно дается условным распределением вероятностей, при условии $X(\omega) = x$:

$$P_x = PE_x/P\{X = x\},$$

где через E_x обозначен индикатор подмножества $\{\omega : X(\omega) = x\} \subseteq \Omega$. Классическая формула Байеса

$$P = \sum_x P_x P\{X = x\} \tag{1.42}$$

означает, что исходный "статистический ансамбль" разбивается на подансамбли, соответствующие различным значениям x случайной величины X, которые имеют "веса" $P\{X = x\}$; если эти ансамбли вновь смешиваются, получается просто исходный ансамбль.

Перейдем к квантовым измерениям. Рассмотрим квантовую систему в пространстве $\mathcal{H}$, которая находится в состоянии S и пусть над ней производится измерение вещественной наблюдаемой $X = \sum_x xE_x$, где $E = \{E_x\}$ –спектральная мера оператора X. Предположим сначала, что все собственные числа X различны, так что $E_x = |e_x\rangle\langle e_x|$, где $\{|e_x\rangle\}$– базис из собственных векторов X. *Полное идеальное* квантовое измерение связывается

с ортонормированным базисом $|e_x\rangle$, векторы которого индексированы возможными исходами измерения x. Постулируется, что в результате такого измерения система переходит в состояния $|e_x\rangle\langle e_x|$ с вероятностями $\langle e_x|S|e_x\rangle$. Таким образом, статистический ансамбль после измерения разбивается на подансамбли, соответствующие различным исходам x, и в целом описывается состоянием

$$S' = \sum_x |e_x\rangle\langle e_x|S|e_x\rangle\langle e_x|. \tag{1.43}$$

В общем случае некоторые собственные числа X могут совпадать. Согласно *проекционному постулату фон Неймана - Людерса* [6], идеальное измерение наблюдаемой X дает значение x с вероятностью

$$p_x = \operatorname{Tr} SE_x = \operatorname{Tr} E_x S E_x,$$

при этом *апостериорное состояние*, т. е. состояние подансамбля, в котором был получен исход измерения x, равно

$$S_x = p_x^{-1} E_x S E_x, \quad \text{если} \quad p_x > 0.$$

Для состояния всего ансамбля после измерения имеет место квантовый аналог формулы Байеса

$$S' = \sum_x p_x S_x = \sum_x E_x S E_x. \tag{1.44}$$

Заметим, что в отличие от классической формулы Байеса (1.42), квантовое состояние (1.44) всего ансамбля после идеального измерения может отличаться от начального состояния S. Таким образом, идеальное квантовое измерение не сводится к простому наблюдению выходного символа и включает в себя воздействие, которое изменяет состояние системы, даже если исходы "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 в соответствии со значениями случайных величин.

Рассмотрим теперь последовательное измерение, при котором над системой, приготовленной в состоянии S, сначала измеряется наблюдаемая $X = \sum_x xE_x$, а затем $Y = \sum_y yF_y$. Используя формулу условной вероятности, а также определение апостериорного состояния, получаем распределение вероятностей

$$\begin{aligned} \mathsf{P}_S(X = x, Y = y) &= \mathsf{P}_S(Y = y|X = x)\mathsf{P}_S(X = x) = \operatorname{Tr} S_x F_y \operatorname{Tr} SE_x \\ &= \operatorname{Tr} E_x S_x E_x F_y = \operatorname{Tr} F_y E_x S_x E_x F_y. \end{aligned} \tag{1.45}$$

Заметим, что вообще говоря, $\mathsf{P}_S(Y = y, X = x) \neq \mathsf{P}_S(X = x, Y = y)$, т.е. совместное распределение зависит от порядка измерения. Однако если X, Y совместимы, то E_x, F_y коммутируют и здесь имеет место равенство для всех состояний S, при этом распределение (1.45) совпадает с (1.34).

Задача 18. Идеальное квантовое измерение удовлетворяет условию *воспроизводимости*: при повторном измерении наблюдаемой X исход с вероятностью 1 равен исходу первого измерения.

Задача 19. Обобщите соотношения (1.45) на случай n наблюдаемых.

Задача 20. Рассмотрите последовательное измерение наблюдаемых σ_x, σ в состоянии q-бита (1.25) и покажите, что

$$\mathsf{P}_{S(\vec{a})}(\sigma_x = 1, \sigma_z = 1) = \frac{1}{4}(1 + a_x), \quad \mathsf{P}_{S(\vec{a})}(\sigma_z = 1, \sigma_x = 1) = \frac{1}{4}(1 + a_z).$$

1.11 Обратимые эволюции

Обратимые эволюции классической системы описываются взаимно-однознач преобразованиями фазового пространства: $\omega' = T\omega$. При этом статистические состояния подвергаются аффинному взаимно-однозначному преобразованию $P = \{p_\omega\} \to P' = \{p_{\omega'}\}$.

По аналогии, в квантовом случае мы рассмотрим аффинные преобразования, которые переводят операторы плотности в операторы плотности $\Phi : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$,

$$\Phi\left[\sum_j p_j S_j\right] = \sum_j p_j \Phi[S_j]; \quad p_j \geq 0, \quad \sum_j p_j = 1; \quad S_j \in \mathcal{S}(\mathcal{H}).$$

Свойство аффинности имеет прямой статистический смысл: оно означает сохранение "весов" в смесях состояний.

Пример 1. Пусть U – унитарный оператор, тогда $\Phi[S] = USU^*$ является аффинным и взаимно-однозначным отображением множества квантовых состояний $\mathcal{S}(\mathcal{H})$ на себя, т. е. задает *обратимую* эволюцию.

Следующий результат, восходящий к Вигнеру, характеризует все обратимые квантовые эволюции.

Теорема 6. *Пусть* Φ *– аффинное взаимно-однозначное отображение выпуклого множества квантовых состояний* $\mathcal{S}(\mathcal{H})$ *на себя, тогда*

$$\Phi[S] = USU^*, \quad S \in \mathcal{S}(\mathcal{H}), \tag{1.46}$$

где U *- унитарный или антиунитарный оператор.*

Антиунитарный оператор U характеризуется свойствами:

1. $\|U\psi\| = \|\psi\|; \psi \in \mathcal{H}$;
2. $U(\sum c_j\psi_j) = \sum \overline{c}_j U\psi_j$.

Такой оператор всегда можно представить в виде $U = \widetilde{U}\Lambda$, где $\widetilde{U}$– унитарный оператор, а $\Lambda = \Lambda^*$– антиунитарный оператор комплексного сопряжения в некотором фиксированном базисе Соответствующая ему эволюция состояний задается транспонированием матрицы плотности в этом базисе

$$S^\top = \Lambda S \Lambda^*.$$

При обратимой эволюции чистые состояния переходят в чистые, при этом вектор исходного чистого состояния $|\psi\rangle$ преобразуется в $U|\psi\rangle$.

Пример 2. *Обратимые эволюции q-бита.* Всякое аффинное взаимнооднозначное отображение шара Блоха на себя оставляет инвариантной единичную сферу (множество крайних точек шара), откуда следует, что такое отображение является либо вращением в $\mathbb{R}^3$, либо комбинацией вращения и отражения относительно какой-либо плоскости, проходящей через начало координат. В координатах такое отображение задается вещественной ортогональной 3×3-матрицей O, причем в первом случае $\det O = 1$, а во втором – $\det O = -1$. Пусть $S(\vec{a})$ – произвольное состояние q-бита (1.25). Тогда

$$S(O\vec{a}) = US(\vec{a})U^*, \tag{1.47}$$

где U – в первом случае унитарный, а во втором – антиунитарный оператор. Формула (1.46) определяет проективное представление [5] группы ортогональных преобразований $O(3)$ в $\mathbb{C}^2$.

Если вектор $\vec{a}$ единичный, то $S(\vec{a})$ – чистое состояние (1.27), соответствующее направлению спина $\vec{a}$. При этом векторы чистых состояний в $\mathbb{C}^2$ преобразуются по формуле

$$|\psi(\vec{a})\rangle \to |\psi(O\vec{a})\rangle = U|\psi(\vec{a})\rangle.$$

Рассмотрим, например, вращение O вокруг оси z на угол φ. При этом единичный вектор $\vec{a}$ с углами Эйлера θ, ϕ преобразуется в вектор $O\vec{a}$ с углами $\theta, \phi+\varphi$, поэтому вектор состояния (1.26) в $\mathcal{H}$ переходит в вектор

$$|\psi(O\vec{a})\rangle = \begin{bmatrix} \cos\frac{\theta}{2}e^{-i(\phi+\varphi)/2} \\ \sin\frac{\theta}{2}e^{i(\phi+\varphi)/2} \end{bmatrix} = U|\psi(\vec{a})\rangle, \tag{1.48}$$

где $U = \mathrm{diag}[e^{-i\varphi/2}, e^{i\varphi/2}]$ – унитарная матрица.

Известно, что всякое вращение в $\mathbb{R}^3$ можно реализовать как поворот вокруг некоторой оси. Задача 21. Покажите, что вращению шара Блоха на угол φ вокруг оси $\vec{a}$ отвечает унитарный оператор

$$U = \exp\left[-\frac{i\varphi}{2}\sigma(\vec{a})\right] = \cos\frac{\varphi}{2}I - i\sin\frac{\varphi}{2}\sigma(\vec{a}). \tag{1.49}$$

Задача 22. Пусть O_γ – отражение в $\mathbb{R}^3$ относительно координатной оси $\gamma; \gamma = x, y, z$ (равносильное повороту на угол π относительно этой оси). Соответствующий унитарный оператор $U = \sigma_\gamma$.

Поскольку

$$\sigma_x|0\rangle = |1\rangle, \quad \sigma_x|1\rangle = |0\rangle,$$

операция σ_x называется "переворот бита". Операция σ_z

$$\sigma_z|0\rangle = |0\rangle, \quad \sigma_z|1\rangle = -|1\rangle,$$

называется "переворот фазы". Операция $\sigma_y = i\sigma_z\sigma_x$ является их комбинацией.

З а д а ч а 23. Отражению O_{xz} относительно плоскости xz с $\det O_{xz} = -1$ отвечает оператор $U = \Lambda$ комплексного сопряжения в базисе $|0\rangle, |1\rangle$. При этом отображение (1.47) сводится к транспонированию матрицы плотности $S(\vec{a})$.

Т е о р е м а 7. *Всякая непрерывная однопараметрическая группа обратимых квантовых эволюций $\Phi_t; t \in \mathbb{R}$, такая что $\Phi_0 = \mathrm{Id}$ (тождественное отображение), имеет вид*

$$\Phi_t[S] = \mathrm{e}^{-itH} S \mathrm{e}^{itH}, \quad S \in \mathcal{S}(\mathcal{H}), \tag{1.50}$$

где H - эрмитов оператор.

Если параметр t – время, то оператор H называется гамильтонианом или оператором энергии. Дифференцируя соотношение (1.50) по t, получаем уравнение Лиувилля для оператора плотности $S_t = \Phi_t[S]$:

$$\frac{dS_t}{dt} = i[S_t, H] \quad S_0 = S.$$

Для векторов чистых состояний $|\psi_t\rangle$ получается уравнение Шредингера

$$\frac{d|\psi_t\rangle}{dt} = -iH|\psi_t\rangle, \quad |\psi_0\rangle = |\psi\rangle.$$

Доказательство теоремы 6 см. например в [13]; доказательство теоремы 7 (см. например [11]) основывается на теореме Стоуна о непрерывных группах унитарных операторов. Последняя является операторным аналогом известного в анализе факта, что непрерывные решения функционального уравнения $f(t+s) = f(t)f(s)$ суть экспоненциальные функции.

Отметим, что обращение времени $t \to -t$ равносильно комплексному сопряжению в уравнении Шредингера или транспонированию в уравнении Лиувилля, рассматриваемым в базисе, в котором гамильтониан задается вещественной матрицей.

П р и м е р 3. *Прецессия спина в постоянном магнитном поле.* Гамильтониан спина во внешнем магнитном поле, направленном вдоль оси z имеет вид

$$H = \frac{1}{2}\omega\sigma_z,$$

где коэффициент ω пропорционален напряженности поля. Согласно (1.49), оператор унитарной эволюции

$$\mathrm{e}^{-itH} = \mathrm{e}^{-\frac{i}{2}\omega t\sigma_z}$$

соответствует повороту вокруг оси z на угол ωt. Таким образом, если первоначальное чистое состояние q-бита задавалось единичным вектором $\vec{a}$, то временная эволюция в шаре Блоха описывается вращением соответствующего единичного вектора вокруг оси z с угловой скоростью ω.

Любопытно отметить, что отправляясь от понятия обратимой эволюции конечной классической системы, мы пришли непрерывному семейству квантовых эволюций, не имеющему классического аналога.

Изучением и решением уравнения Шредингера, а также связанными с этим вопросами занимается квантовая механика, см. например, [2], [9], [8]. В связи с проблемами квантовой информатики представляет интерес задача *управления* для уравнения Шредингера: в ней гамильтониан зависит от управляемых параметров, траекторию которых на отрезке $[0, T]$ следует выбрать так, чтобы в момент T получить заданный унитарный оператор эволюции, либо получить эволюцию с заданными свойствами.

Глава 2

Составные квантовые системы

2.1 Классические и квантовые корреляции

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

$$S_1 = \left[\sum_j s_{(ij)(kj)}\right], \quad S_2 = \left[\sum_i s_{(ij)(il)}\right]. \tag{2.1}$$

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 24. 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

$$|00\rangle = |0\rangle_1 \otimes |0\rangle_2,\ |01\rangle = |0\rangle_1 \otimes |1\rangle_2,\ |10\rangle = |1\rangle_1 \otimes |0\rangle_2,\ |11\rangle = |1\rangle_1 \otimes |1\rangle_2. \tag{2.3}$$

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

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]:

Теорема 8 *[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.*

Из равенства ненулевых собственных значений и выражения для энтропии (1.23) вытекает

Следствие. Энтропии частичных состояний S_1 и S_2 равны между собой: $H(S_1) = H(S_2)$.

Доказательство теоремы. 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 9 *[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

[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.

$\mathcal{H}_1$. 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. □

Квантовая сцепленность отражает необычные свойства составных квантовых систем, которые описываются тензорным (а не декартовым, как в классической механике) произведением подсистем. Сцепленность возникает при квантовом взаимодействии подсистем. В силу принципа суперпозиции, пространство составной системы AB наряду с векторами вида $|\psi_A\rangle \otimes |\psi_B\rangle$ содержит и всевозможные их линейные комбинации $\sum_j |\psi_A^j\rangle \otimes |\psi_B^j\rangle$. Состояния составной системы, задаваемые векторами-произведениями, называются *разделимыми* или несцепленными, а все не сводящиеся к таковым – *сцепленными*. Смешанное состояние называется разделимым, если оно является смесью состояний-произведений. Сцепленность представляет собой чисто квантовое свойство, лишь отчасти родственное классической коррелированности, однако к ней не сводящееся (в физике говорят о корреляциях Эйнштейна-Подольского-Розена, поскольку эти авторы впервые обратили внимание на необычные свойства таких корреляций). Именно, наличие сцепленных состояний противоречит гипотезе о локальной теории со скрытыми параметрами, т. е. классического статистического описания квантовых систем, удовлетворяющего физическому требованию локальности (раздел 2.5).

Большой раздел квантовой теории информации посвящен количественной теории сцепленности, которая представляет собой своеобразную комбинаторную геометрию тензорных произведений гильбертовых пространств. В частности, показано, что мера сцепленности чистого состояния S_{AB} составной системы AB определяется однозначно как энтропия частичного следа $\mathrm{Tr}_B S_{AB}$, тогда как для смешанных состояний имеется целый ряд существенно различных характеристик, важнейшей из которых является *сцепленность формирования*

$$E_F(S_{AB}) = \min \sum_i p_i H(\mathrm{Tr}_B P_{\psi_i}),$$

где минимум берется по всевозможным ансамблям, представляющим состояние S_{AB}. Показано, что эта характеристика связана с количеством максимально сцепленных пар q-битов (т. н. e-битов), которое необходимо для создания состояния S_{AB} с использованием локальных операций (затрагивающих только A либо B) и обмена классической информацией между A и B.

Двойственным образом, в составных квантовых системах существуют сцепленные и несцепленные *наблюдаемые* (измерения). Если квантовые системы A и B находятся в несцепленном состоянии, то максимальные шенноновские количества информации о состоянии I_A, I_B, I_{AB} , получаемые, соответственно, из измерений над системами A, B и составной системой AB

удовлетворяют в общем случае соотношению $I_{AB} > I_A + I_B$. Этот неклассический феномен строгой *супераддитивности* информации обнаруживается и играет важную роль в теории пропускных способностей квантового канала связи (раздел 6.4).

2.4 Два q-бита

Ортонормированный базис в пространстве двух q-битов образован четырьмя векторами (2.3). Разложение произвольного вектора чистого состояния двух q-битов имеет вид

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle,$$

где

$$|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1.$$

Соответствующий оператор плотности

$$|\psi\rangle\langle\psi| = (a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle)\left(\bar{a}\langle 00| + \bar{b}\langle 01| + \bar{c}\langle 10| + \bar{d}\langle 11|\right).$$

Частичное состояние первого q-бита

$$S_1 = \mathrm{Tr}_2|\psi\rangle\langle\psi|$$

$$= \left(|a|^2 + |b|^2\right)|0\rangle\langle 0| + \left(a\bar{c} + b\bar{d}\right)|0\rangle\langle 1| + \left(\bar{a}c + \bar{b}d\right)|1\rangle\langle 0| + \left(|c|^2 + |d|^2\right)|1\rangle\langle 1|.$$

Собственные числа матрицы плотности находятся из характеристического уравнения

$$\det\left(S_1 - \lambda I\right) = \det\begin{bmatrix} |a|^2 + |b|^2 - \lambda & a\bar{c} + b\bar{d} \\ \bar{a}c + \bar{b}d & |c|^2 + |d|^2 - \lambda \end{bmatrix} = \lambda^2 - \lambda + C^2/4 = 0,$$

где $C = 2\left|ad - bc\right|$, так что $0 \le C \le 1$. Отсюда

$$\lambda_\pm = \frac{1 \pm \sqrt{1 - C^2}}{2}.$$

Величина C называется "согласованностью" (concurrence) и служит мерой сцепленности для чистого состояния двух q-битов. При $C = 0$ получается несцепленное состояние, $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, а при $C = 1$ – максимально сцепленное состояние, для которого $\lambda_\pm = \frac{1}{2}$.

2.5 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

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big[|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle\Big],$$

где векторы

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

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

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big[|01\rangle - |10\rangle\Big].$$

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.

Предположим, что частицы разлетаются на некоторое макроскопическое расстояние, при этом их спиновое состояние – синглет – сохраняется. Рассмотрим эксперимент, в котором в двух удаленных друг от друга лабораториях A и B над этими разлетевшимися частицами производятся одновременные измерения: наблюдаемой спина $\sigma(\vec{a})$ для одной частицы и $\sigma(\vec{b})$ для другой. Операторы $X = \sigma(\vec{a}) \otimes I, Y = I \otimes \sigma(\vec{b})$ коммутируют, следовательно соответствующие наблюдаемые совместимы и их ковариация дается выражением (1.40).

З а д а ч а 27. Используя выражения для матричных элементов,

$$\langle 0|\sigma(\vec{a})|0\rangle = a_z, \quad \langle 1|\sigma(\vec{a})|0\rangle = a_x + ia_y, \quad \langle 1|\sigma(\vec{a})|1\rangle = -a_z, \qquad (2.9)$$

вытекающие из (1.28), покажите, что в синглетном состоянии среднее значение и дисперсия каждой наблюдаемой равны

$$\mathsf{M}\sigma(\vec{a}) = 0, \quad \mathsf{D}\sigma(\vec{a}) = 1$$

а ковариация между спинами задается формулой

$$\langle\psi|\sigma(\vec{a}) \otimes \sigma(\vec{b})|\psi\rangle = -\vec{a} \cdot \vec{b}. \qquad (2.10)$$

Отсюда следует, что если $\vec{b} = \vec{a}$, то коэффициент корреляции равен -1, и следовательно между исходами a, b измерений имеется детерминированная связь: $a = -b$. Из формул (2.9) и соотношения $\sigma(\vec{a})^2 = I$ следует

$$\langle\psi|[\sigma(\vec{a}) \otimes I + I \otimes \sigma(\vec{a})]^2|\psi\rangle = 0,$$

откуда

$$[\sigma(\vec{a}) \otimes I + I \otimes \sigma(\vec{a})]|\psi\rangle = 0.$$

Если бы спины описывались классическими векторными случайными величинами, то это означало бы, что при измерении спина первой частицы в произвольно выбранном направлении $\vec{a}$ спин второй частицы "моментально" принимает противоположное значение.

Таким образом, 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 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. Ковариация (2.10) не может быть смоделирована никакой классической моделью составной системы, удовлетворяющей принципу локальности. Это доказывается с помощью 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*

$$|\mathsf{M}X_1Y_1 + \mathsf{M}X_1Y_2 + \mathsf{M}X_2Y_1 - \mathsf{M}X_2Y_2| \le 2, \tag{2.11}$$

где M *— соответствующее математическое ожидание.*

Доказательство получается усреднением элементарного неравенства

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

(задача 28). 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.

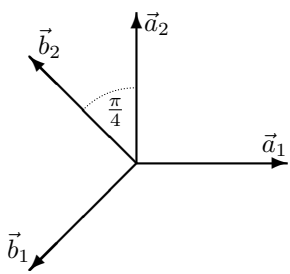


Рис. 2.1: Выбор векторов $\vec{a}_j$ и $\vec{b}_k$

При этом система приготавливается в одном и том же синглетном состоянии. Подстановка соответствующих значений корреляций из формулы (2.10) в левую часть формулы (2.11) 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.

2.6 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* $S = |\Psi\rangle_{AB}\langle\Psi|$,

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B\right) = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right). \qquad (2.12)$$

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.

З а д а ч а 29. Докажите утверждения 1-3.

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".

2.7 Корреляционные неравенства и операторные алгебры

Если бы четыре корреляции в (2.11) 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. Пусть X_j, Y_k $(j,k=1,2)$ вещественные четкие квантовые наблюдаемые, т.ч. $X_j^2 = I$, $Y_k^2 = I$, $X_jY_k = Y_kX_j$. Тогда для любого квантового состояния S имеет место неравенство Цирельсона

$$|\mathsf{M}_S X_1Y_1 + \mathsf{M}_S X_1Y_2 + \mathsf{M}_S X_2Y_1 - \mathsf{M}_S X_2Y_2| \le 2\sqrt{2}. \tag{2.13}$$

Доказательство вытекает из тождества (задача 30)

$$(X_1Y_1 + X_1Y_2 + X_2Y_1 - X_2Y_2)^2 = 4I - [X_1,X_2][Y_1,Y_2], \tag{2.14}$$

с учетом того, что $\|[X_1,X_2]\| \le 2$, так что норма выражения (2.14) не превосходит 8.

Для системы из двух q-битов AB равенство в (2.13) достигается для наблюдаемых

$$X_j = \sigma(\vec{a}_j)\otimes I_B, \quad Y_k = I_A \otimes \sigma(\vec{b}_k)$$

и состояния Белла (2.12).

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 (2.11))[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

Квантовые информационные протоколы

3.1 Квантовое состояние как информационный ресурс

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

$$P(y|x) = \langle e_y|S_x e_y\rangle. \tag{3.1}$$

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 5. Отметим здесь лишь один факт, имеющий принципиальное значение (см. (5.26)):

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 (3.1) не зависит от x, а значит количество передаваемой информации в самом деле равно нулю.

Хотя ЭПР-корреляции, т.е. квантовая сцепленность, 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, сцепленность квантового состояния составной системы выступает 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

$$|e_+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \quad |e_-\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right),$$

$$|h_+\rangle = \frac{1}{\sqrt{2}}\left(|10\rangle + |01\rangle\right), \quad |h_-\rangle = \frac{1}{\sqrt{2}}\left(|10\rangle - |01\rangle\right)$$

в системе из двух q-битов AB (мы используем вычислительный базис $|0\rangle, |1\rangle$ в пространстве одного q-бита) 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 Телепортация квантового состояния

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-корреляци 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 двухуровневы (q-битами).

1. Перед началом передачи система AB приготовляется в состоянии с вектором
$$\frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right).$$

2. C посылает A произвольное чистое состояние
$$|\psi\rangle = a|0\rangle + b|1\rangle.$$
Совокупность трех систем CAB описывается вектором состояния
$$(a|0\rangle + b|1\rangle) \otimes \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) = \frac{1}{\sqrt{2}}\left[a|000\rangle + b|100\rangle + a|011\rangle + b|111\rangle\right].$$

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 $|\psi\rangle$.

[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.

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

$$|00\rangle \to |00\rangle, \quad |01\rangle \to |01\rangle, \quad |10\rangle \to |11\rangle, \quad |11\rangle \to |10\rangle, \tag{3.2}$$

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

$$|0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \qquad |1\rangle \to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{3.3}$$

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

Начальное состояние всей системы CAB есть

$$\frac{1}{\sqrt{2}}\left(a|000\rangle + b|100\rangle + a|011\rangle + b|111\rangle\right).$$

После действия CNOT на CA получаем

$$\frac{1}{\sqrt{2}}\left(a|000\rangle + b|110\rangle + a|011\rangle + b|101\rangle\right).$$

Потом H действует на C

$$\frac{1}{2}\left[a(|000\rangle + |100\rangle) + b(|010\rangle - |110\rangle) + a(|011\rangle + |111\rangle) + b(|001\rangle - |101\rangle)\right].$$

Выделяя состояние системы CA, получаем

$$\frac{1}{2}\left[|00\rangle(a|0\rangle + b|1\rangle) + |01\rangle(a|1\rangle + b|0\rangle) + |10\rangle(a|0\rangle - b|1\rangle) + |11\rangle(a|1\rangle - b|0\rangle)\right].$$

Теперь производится измерение в системе CA, с вероятностями $1/4$ 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.)[2].

[2] http://en.wikipedia.org/wiki/Quantum teleportation.

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 сложение двоичных векторов по модулю 2 (логическая операция "XOR"). 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, растущее экспоненциально с n. Можно доказать, что и применение вероятностных алгоритмов, которые дают правильный ответ лишь с заданной вероятностью $p > 1/2$, требует не менее $O(2^{n/2})$ шагов. (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 $s(s-1)/2 \geq 2^n - 1$, откуда $s \geq O(2^{n/2})$.)

В теоретической информатике задачи, требующие для своего решения экспоненциальное число неких элементарных операций, принято считать "трудными", см. подробнее [4]. На практике решение такого рода задач уже при n порядка нескольких сотен может потребовать нереально большого времени даже при использовании современных суперкомпьютеров. С другой стороны, задачи, имеющие полиномиальную сложность, т.е. требующие числа шагов, растущего как степень n, обычно поддаются практическим вычислениям, и всякий новый алгоритм, обладающий таким свойством, представляет большой интерес.

Квантовый алгоритм нахождения периода 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., называемый вычислительным базисом. 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:

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

   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 двухкубитных 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 $|00\ldots\rangle$, 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 в вычислительном базисе, 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 двоичной "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*

$$\mathsf{P}\{y_1(\omega), \ldots, y_{n-1}(\omega) \ \} \geq e^{-2}.$$

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

$$\mathsf{P}\{y_k(\omega) \text{ линейно независим от } y_1(\omega), \ldots, y_{k-1}(\omega) | y_1(\omega), \ldots, y_{k-1}(\omega) \text{ линейно н}$$

$$= \frac{2^{n-1} - 2^{k-1}}{2^{n-1}} = 1 - \frac{1}{2^{n-k}}.$$

Тогда

$$\mathsf{P}\{y_1(\omega), \ldots, y_k(\omega) \text{ линейно независимы}\}$$

$$= \mathsf{P}\{y_k(\omega) \text{ линейно независим от } y_1(\omega), \ldots, y_{k-1}(\omega);$$

$$y_1(\omega), \ldots, y_{k-1}(\omega) \text{ линейно независимы}\}$$

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

Следовательно

$$\mathsf{P}\{y_1(\omega), \ldots, y_{n-1}(\omega) \text{ линейно независимы}\}$$

$$= \left(1 - \frac{1}{2^{n-1}}\right) \ldots \left(1 - \frac{1}{2}\right)$$

$$= 2^{\sum_{k=1}^{n-1} \log\left(1 - \frac{1}{2^k}\right)} \geq 2^{-2\sum_{k=1}^{n-1} \frac{1}{2^k}} \geq \frac{1}{4},$$

где использовано неравенство $\log(1-x) \geq -2x; 0 \leq x \leq 1/2$.

5) Повторяем всю процедуру m раз, где $(1-1/4)^m \leq \varepsilon$. 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 эффект 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 $O(n^2 \log n$ 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 [7, 3].

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

$$U_F|x\rangle = |x\rangle, x \neq x_0, \quad U_F|x_0\rangle = -|x_0\rangle. \tag{3.4}$$

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 H_n

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

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 $|\psi(\theta_0)\rangle$, 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

$$J|0\rangle = |0\rangle, \quad J|x\rangle = -|x\rangle \quad x \neq 0.$$

Таким образом,

$$J = 2|0\rangle\langle 0| - I$$

и

$$H_n J H_n = 2H_n|0\rangle\langle 0|H_n - I = 2|\psi(\theta_0)\rangle\langle\psi(\theta_0)| - I. \tag{3.5}$$

3. Операция U повторяется m раз, где $m = [(\pi/4)\sqrt{N}]$.

Лемма 2. *Для любого угла* θ

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

где

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

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

Доказательство. Используя (3.4), (3.5), получаем

$$\begin{aligned} U|x_0\rangle &= -H_n J H_n|x_0\rangle = -2|\psi(\theta_0)\rangle\langle\psi(\theta_0)|x_0\rangle + |x_0\rangle \\ &= \left(1 - \frac{2}{N}\right)|x_0\rangle - 2\frac{\sqrt{N-1}}{N}|\bar{x}_0\rangle \end{aligned}$$

и аналогично

$$U|\bar{x}_0\rangle = 2\frac{\sqrt{N-1}}{N}|x_0\rangle + \left(1 - \frac{2}{N}\right)|\bar{x}_0\rangle.$$

Таким образом, после применения оператора U $m = [(\pi/4)\sqrt{N}]$ раз

$$U^m|\psi(\theta_0)\rangle = |\psi(\theta_m)\rangle,$$

где

$$\theta_m = \theta_0 + m\varphi = \frac{1}{\sqrt{N}} + \frac{\pi\sqrt{N}}{4}\frac{2\sqrt{N-1}}{N} + O\left(\frac{1}{\sqrt{N}}\right) = \pi/2 + O\left(\frac{1}{\sqrt{N}}\right).$$

При этом конечное состояние $|\psi(\theta_m)\rangle, \theta_m = \theta_0 + m\varphi$ становится очень близким к искомому:

$$\||\psi(\theta_m)\rangle - |x_0\rangle\| = \sin(\theta_m - \pi/2) = O\left(\frac{1}{\sqrt{N}}\right),$$

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.

3.4.4 Замечания о моделировании унитарных операций

Предположим, что задан некоторый унитарный оператор на n q-битах. Желательно представить его в виде последовательности ("схемы") из некоторых основных элементарных операций, каждая из которых затрагивала бы минимальное число q-битов. Кроме того, желательно, чтобы схема содержала минимальное число элементов. Такого рода вопрос возникает в связи с оценкой сложности квантовых алгоритмов, а также при моделировании унттарной динамики квантовых систем. Его подробное рассмотрение дано в [7]. Здесь мы лишь кратко сформулируем основные выводы.

1. Произвольный унитарный оператор на n q-битах может быть реализован с помощью конечной схемы, состоящей только из одно-q-битных операций и (возможно, нескольких) двух-q-битных операций CNOT (контролируемое "нет") (3.2).

2. Произвольный унитарный оператор на n q-битах может быть с произвольной точностью реализован с помощью конечной схемы, состоящей только из (возможно, нескольких) применений однокубитной операции H (операция Адамара (3.3)), однокубитной операции

$$T: \quad |0\rangle \to |0\rangle, \qquad |1\rangle \to \mathrm{e}^{i\pi/4}|1\rangle,$$

и двух-q-битной операции CNOT. Известны и другие конечные универсальные наборы мало-q-битных операций.

3. Существуют унитарные операторы на n q-битах, для ε- аппроксимации которых требуется не менее порядка $2^n \log \frac{1}{\varepsilon} / \log n$ операций из любого фиксированного конечного набора. В этом смысле такие операции не могут быть реализованы эффективно. С другой стороны, унитарные динамики с "локальными" гамильтонианами, описываемыми "разреженными" матрицами, допускают эффективное моделирование [7].

3.5 Квантовые коды, исправляющие ошибки

3.5.1 Постановка вопроса

При передаче информации по каналу с шумом, а также при выполнении квантовых операций, желательно иметь код, который был бы устойчив относительно ошибок. В классическом случае принципиальная возможность такого кодирования при скоростях передачи, меньших пропускной способности, вытекает из теоремы Шеннона. Однако эта теорема не дает конструктивного способа построения помехоустойчивого кода, и практическому решению этой проблемы посвящена значительная часть исследований по теории информации.

Самый прямой способ застраховаться от ошибок состоит в повторении сообщений (что, конечно, снижает скорость передачи). Пусть в алфавите есть всего два символа $0, 1$. Предположим, что вероятность изменения одного бита в процессе передачи равна малой величине p, так что вероятность изменения двух битов p^2 — пренебрежимо малая величина. Рассмотрим код $0 \to 00$, $1 \to 11$. Хотя этот код и исправляет некоторые ошибки, он имеет существенный недостаток: например, в ситуации $00 \to 01$, $11 \to 01$ мы не можем сказать, какое сообщение было закодировано. Но от этого недостатка легко избавиться, если добавить еще один разряд: $0 \to 000$, $1 \to 111$. Такой код будет уже помехоустойчивым по отношению к любой ошибке в одном бите.

Прямолинейное обобщение этого рецепта на квантовый случай наталкивается на трудность – квантовую информацию невозможно размножить. Кроме того, по самому существу квантовой информации, при передаче через канал с шумом безошибочно должны приниматься не только базисные состояния, но и всевозможные их суперпозиции. На первый взгляд, такая задача кажется неразрешимой. Однако квантовый код, исправляющий ошибки, был построен независимо в работах Шора и Стина. Многие авторы сделали вклад в последующее развитие теории, фрагменты которой представлены в этом разделе, см. обзор в [7].

Следуя классической аналогии, рассмотрим сначала код

$$|0\rangle \to |000\rangle, |1\rangle \to |111\rangle. \tag{3.6}$$

Такой код исправляет "переворот бита", т. е. переход $|0\rangle \leftrightarrow |1\rangle$ в любом одном q-бите. Нас интересует произвольное состояние $a|0\rangle + b|1\rangle$, которое

при выбранном способе кодирования переходит в $a|\,000\rangle + b|\,111\rangle$. Пусть, например, произошла ошибка в первом q-бите:

$$a|\,0\rangle + b|\,1\rangle \to a|\,100\rangle + b|\,011\rangle,$$

Состояния $a|\,000\rangle + b|\,111\rangle, a|\,100\rangle + b|\,011\rangle$ ортогональны, следовательно их можно безошибочно различить.

Однако такой код не исправляет "переворот фазы" типа $|\,0\rangle \leftrightarrow |\,0\rangle, |\,1\rangle \leftrightarrow -|\,1\rangle$. В самом деле, в результате такой фазовой ошибки в одном бите получим $a|\,000\rangle - b|\,111\rangle$ вместо $a|\,000\rangle + b|\,111\rangle$, и эти состояния не ортогональны, т. е. безошибочно не различимы.

Теперь заметим, что преобразование Адамара

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad |\,0\rangle \to \frac{1}{\sqrt{2}}(|\,0\rangle + |\,1\rangle), \quad |\,1\rangle \to \frac{1}{\sqrt{2}}(|\,0\rangle - |\,1\rangle)$$

отображает переворот фазы в переворот бита и наоборот. Преобразуя соответствующим образом код (3.6), получим код

$$\begin{aligned} |\,0\rangle &\to \frac{1}{2^{3/2}}(|\,0\rangle + |\,1\rangle)(|\,0\rangle + |\,1\rangle)(|\,0\rangle + |\,1\rangle), \\ |\,1\rangle &\to \frac{1}{2^{3/2}}(|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle), \end{aligned} \tag{3.7}$$

который исправляет переворот фазы в любом одном q-бите, но не исправляет переворот бита.

Код Шора, который исправляет как переворот бита, так и переворот фазы в одном q-бите, получается комбинированием кодов (3.6), (3.7) и требует 9 q-битов

$$\begin{aligned} |\,0\rangle &\to \frac{1}{2^{3/2}}(|\,000\rangle + |\,111\rangle)(|\,000\rangle + |\,111\rangle)(|\,000\rangle + |\,111\rangle) \\ |\,1\rangle &\to \frac{1}{2^{3/2}}(|\,000\rangle - |\,111\rangle)(|\,000\rangle - |\,111\rangle)(|\,000\rangle - |\,111\rangle) \end{aligned} \tag{3.8}$$

Оказывается, что этот код исправляет не только битовую и фазовую, но любую ошибку, возникающую в результате применения *произвольного* канала в одном (любом) из q-битов. Прямая проверка этого факта является трудоемким упражнением, см. [7], n. 10.2.

3.5.2 Общая формулировка

Пусть S произвольное состояние в гильбертовом пространстве $\mathcal{M}$. *Кодом* называется изометрическое отображение $V : \mathcal{M} \to \mathcal{N}$, переводящее состояния S в закодированные состояния VSV^* в гильбертовом пространстве $\mathcal{N}$. На самом деле код можно задавать подпространством $\mathcal{L} = V\mathcal{M} \subset \mathcal{N}$, не вводя явно $\mathcal{M}, V$.

Система $\mathcal{N}$ может быть подвержена ошибкам, эффект которых описывается операторами класса $\mathcal{E} = \mathrm{Lin}(B_1, \ldots, B_p)$, где B_j – операторы *элементарных ошибок*.

Пример. Хранение квантовой информации в памяти квантового компьютера. Пусть $\mathcal{N} = \mathcal{H}_2^{\otimes n}$ – квантовый регистр, в котором предполагается хранить информацию из $\mathcal{M}$. Рассмотрим ошибки, при которых изменению может подвергнуться не более m q-битов регистра. Соответствующее множество $\mathcal{E}(n,m)$ состоит из линейных комбинаций операторов $V = V_1 \otimes \cdots \otimes V_n$, где количество $V_k \neq I$ не превышает m, причем ошибка в k-м q-бите V_k может быть произвольной. Операторами элементарных ошибок в каждом q-бите могут служить матрицы Паули, причем σ_x описывает переворот бита, σ_z переворот фазы, а $\sigma_y = i\sigma_x\sigma_z$ – их комбинацию. Вместе с единичным оператором I, который соответствует отсутствию ошибки, они образуют базис в алгебре наблюдаемых q-бита.

Пример Шора демонстрирует возможность исправления ошибок из $\mathcal{E}(n,1)$ если n достаточно велико (можно доказать, что наименьшее значение n для кода, исправляющего одну ошибку, равно 5). Возможность исправления только одной ошибки является, конечно, серьезным ограничением. Однако удалось показать, что существуют коды, исправляющие ошибки из $\mathcal{E}(n,m)$, где m может быть сколь угодно большим для достаточно больших размеров регистра n. Более того, была предложена принципиальная схема квантового компьютера, исправляющего ошибки не только в квантовой памяти, но и в самой схеме, исправляющей ошибки, при условии, что вероятность ошибки не превосходит некоторого порогового значения (fault-tolerant quantum computing) [3]. Оценки показывают, что квантовое устройство, способное решать задачи, недоступные классическому компьютеру, должно иметь регистр, насчитывающий 2000-3000 q-бит и порядка 10^{10} элементарных логических операторов, которые должны иметь уровень надежности порядка 10^{-13}. Более реалистический уровень надежности 10^{-4} достаточен при использовании исправления ошибок, влекущем увеличение вычислительных ресурсов в ≈ 100 раз[1]. В настоящее время имеются экспериментальные установки, позволяющие оперировать с состояниями 4-5 q-битов. Это свидетельствует о том, какой сложный и длинный путь предстоит пройти до практической реализации квантового компьютера. Однако не следует при этом забывать, что все многообразие современных средств обработки информации, в корне преобразившее мир, в котором мы живем, возникло в течение последнего столетия, а лежащие в его основе фундаментальные открытия были сделаны не ранее XIX-го века.

Нилл и Лафламм[3] дали общие необходимые и достаточные условия исправления ошибок, из которых нам понадобятся следующие:

1) для $\phi, \psi \in \mathcal{L}$, таких что $\langle\phi|\,\psi\rangle = 0$, имеет место $\langle\phi|\,B_i^* B_j\,|\psi\rangle = 0$, для всех $i, j = 1, \ldots p$.

2) для какого-либо ортонормированного базиса $\{|\,k\rangle\}$ в $\mathcal{L}$ выполняется

$$\begin{aligned} \langle k|\,B_i^* B_j\,|\,k\rangle &= \langle l|\,B_i^* B_j\,|\,l\rangle, \text{для всех } k, l, \\ \langle k|\,B_i^* B_j\,|\,l\rangle &= 0, \quad \text{для } k \neq l; \end{aligned}$$

[3]Ср. E. Knill, R. Laflamme, "Theory of quantum error-correcting codes," *Phys. Rev. A* **55**, 900-911 1997.

Смысл условия 1) состоит в том, что ошибки не нарушают ортогональности векторов состояний кода.

Задача 31. Докажите равносильность условий 1) и 2).

3.5.3 Аддитивные (симплектические) коды

Рассмотрим поле $B = \{0, 1\}$ с обычными бинарными операциями сложения и умножения. Заметим, что правила умножения для матриц Паули

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

могут быть записаны в форме канонических коммутационных соотношений Вейля на аддитивной группе B^2 из 4-х элементов $0 = (0,0), x = (1,0), y = (1,1), z = (0,1)$ с таблицей сложения

$+$	0	x	y	z
0	0	x	y	z
x	x	0	z	y
y	y	z	0	x
z	z	y	x	0

Именно, вводя кососимметричную функцию Δ с значениями

Δ	0	x	y	z
0	0	0	0	0
x	0	0	1	-1
y	0	-1	0	1
z	0	1	-1	0

имеем

$$\sigma_\gamma \sigma_{\gamma'} = (-i)^{\Delta(\gamma,\gamma')} \sigma_{\gamma+\gamma'} = (-1)^{\Delta(\gamma,\gamma')} \sigma_{\gamma'} \sigma_\gamma, \quad \gamma, \gamma' \in B^2. \tag{3.9}$$

Отметим, что B^2 может также рассматриваться как 2-мерное векторное пространство над полем B.

Задача 32. Покажите, что форма $\Delta(\gamma, \gamma')(\mathrm{mod}2)$ является билинейной и невырожденной на B^2.

Для системы из n q-битов положим $f = (\gamma_1, \ldots, \gamma_n) \in B^{2n}$ и введем эрмитовы операторы $\sigma(f) = \sigma^1_{\gamma_1} \otimes \cdots \otimes \sigma^n_{\gamma_n}$, удовлетворяющие каноническим коммутационным соотношениям

$$\sigma(f)\sigma(g) = i^{\Delta(f,g)} \sigma_{f+g} = (-1)^{\Delta(f,g)} \sigma(g)\sigma(f), \quad f, g \in B^{2n},$$

где $\Delta(f, g) = \Delta(\gamma_1, \gamma'_1) + \cdots + \Delta(\gamma_n, \gamma'_n)$ если $g = (\gamma'_1, \ldots, \gamma'_n)$. Таким образом, B^{2n} становится $2n-$мерным векторным пространством над полем B, снабженным невырожденной симплектической формой $\Delta(f, g)(\mathrm{mod}2)$. Заметим,

что операторы $\sigma(g), \sigma(f)$ коммутируют (антикоммутируют) тогда и только тогда, когда $\Delta(f,g) = 0(\mathrm{mod}2)$, (соответственно $\Delta(f,g) \neq 0(\mathrm{mod}2)$).

Пусть $g_1, \ldots, g_{n-k}$ линейно независимые векторы из B^{2n} такие, что $\Delta(g_i, g$ $0(\mathrm{mod}2)$ для всех i, j. Тогда операторы $\sigma(g_1), \ldots, \sigma(g_{n-k})$ коммутируют между собой. *Аддитивным кодом с проверочным операторами* $\sigma(g)_1, \ldots, \sigma(g)_{n-k}$ называется линейное подпространство

$$\mathcal{L} = \left\{\psi \in \mathcal{H}_2^{\otimes n} : \sigma(g_j)\psi = \psi; \quad j = 1, \ldots, n-k\right\}.$$

Легко видеть, что $\dim \mathcal{L} = 2^k$. Обозначим $G \quad (n-k)$-мерное подпространство пространства B^{2n}, порожденное векторами $g_1, \ldots, g_{n-k}$. Отметим, что $\Delta(f,g) = 0(\mathrm{mod}2)$, $f, g \in G$.

Пусть $\mathcal{E}$ класс ошибок, порождаемый элементарными ошибками $\sigma(f), \quad f$ E, где E – некоторое подмножество B^{2n}. В случае $\mathcal{E} = \mathcal{E}(n,m)$ имеем $E = E(n,m) = \left\{g \in B^{2n} : \mathrm{wt}(g) \leq m\right\}$, где вес $\mathrm{wt}(g)$ равен числу ненулевых компонент вектора g. Из канонических коммутационных соотношений следует, что для любого вектора $\psi \in \mathcal{L}$ и ошибки $\sigma(f)$, вектор $\sigma(f)\psi$ является собственным вектором проверочных операторов $\sigma(g_j)$ с собственными значениями $(-1)^{\Delta(f,g_j)}$. Совокупность этих значений образует *синдром ошибки*. Ошибки $\sigma(f_1), \sigma(f_2)$ *неразличимы*, если их синдромы совпадают, т. е. $\Delta(f_1, g_j) = \Delta(f_2, g_j)(\mathrm{mod}2)$, или

$$f_1 - f_2 \in G^\perp = \left\{f \in B^{2n} : \Delta(f,g)(\mathrm{mod}2) = 0, \quad g \in G\right\}.$$

Отметим, что в силу двоичной природы операций в $B^{2n}, \quad f_1 - f_2$ совпадает с $f_1 + f_2$. Ошибки *эквивалентны*, если $f_1 - f_2 \in G$. Поскольку $G \subset G^\perp$, эквивалентные ошибки неразличимы, но обратное, вообще говоря, неверно.

Теорема 10[4]. *Аддитивный код $\mathcal{L}$ исправляет ошибки класса $\mathcal{E}$, если для любых двух ошибок $\sigma(f_1), \sigma(f_2)$ выполняется либо*

1) *ошибки различимы, т. е. $f_1 - f_2 \notin G^\perp$, и в этом случае оператор $\sigma(f_1)\sigma(f_2)$ антикоммутирует по крайней мере с одним проверочным оператором;*

   либо

2) *ошибки эквивалентны, т. е. $f_1 - f_2 \in G$, и в этом случае оператор $\sigma(f_1)\sigma(f_2)$ пропорционален произведению проверочных операторов.*

Заметим, что условие теоремы может быть компактно записано как

$$(E - E) \cap (G^\perp \setminus G) = \emptyset.$$

Поскольку $E(n,m) - E(n,m) = E(n,2m)$, отсюда следует, что если $d = \min\left\{\mathrm{wt}(g) : g \in G^\perp \setminus G\right\}$, то данный код исправляет любые ошибки в $m = \left[\frac{d-1}{2}\right]$ из n q-битов.

[4] Ср. A. R. Calderbank, E. M. Rains, P. W. Shor, N. J. A. Sloane, "Quantum error correction and orthogonal geometry," *Phys. Rev. Lett.* **78**, 404-408 1997.

Доказательство. Проверим выполнение условия 1) исправления ошибок. Пусть ψ, φ – ортогональные векторы из $\mathcal{L}$, и $f_1, f_2 \in E$. Если ошибки различимы, то векторы $\sigma(f_1)\psi, \sigma(f_2)\varphi$ являются собственными векторами проверочных операторов с различными наборами собственных значений и, следовательно, они ортогональны. Если они неразличимы, то по условию, должно выполняться $f_1 - f_2 \in G$, и в этом случае, используя канонические коммутационные соотношения, получаем

$$\langle\sigma(f_1)\psi|\sigma(f_2)\varphi\rangle = i^{\Delta(f_1,f_2)}\langle\psi|\sigma(f_1 - f_2)\varphi\rangle = i^{\Delta(f_1,f_2)}\langle\psi|\varphi\rangle = 0,$$

поскольку $\sigma(f_1 - f_2)$ является произведением проверочных операторов.

Процедура исправления ошибок состоит из двух этапов: сначала производится измерение проверочных операторов, в результате чего находится синдром ошибки; после этого ошибка определятся с точностью до эквивалентности; применяя оператор ошибки, получаем исходное состояние.

Задача 33. Убедитесь, что код Шора является аддитивным кодом с проверочными операторами

$$\begin{array}{lcccccccccc}
g_1 & = (z & z & 0 & 0 & 0 & 0 & 0 & 0 & 0) \\
g_2 & = (0 & z & z & 0 & 0 & 0 & 0 & 0 & 0) \\
g_3 & = (0 & 0 & 0 & z & z & 0 & 0 & 0 & 0) \\
g_4 & = (0 & 0 & 0 & 0 & z & z & 0 & 0 & 0) \\
g_5 & = (0 & 0 & 0 & 0 & 0 & 0 & z & z & 0) \\
g_6 & = (0 & 0 & 0 & 0 & 0 & 0 & 0 & z & z) \\
g_7 & = (x & x & x & x & x & x & 0 & 0 & 0) \\
g_8 & = (0 & 0 & 0 & x & x & x & x & x & x)
\end{array}$$

Первые шесть операторов обнаруживают переворот бита, а последние два – переворот фазы в любом из блоков. Проверьте выполнение условий исправления ошибок для кода Шора (3.8) и операторов элементарных ошибок, задаваемыми матрицами Паули в произвольном q-бите.

3.6 Квантовая криптография: протоколы распределения секретного ключа

Предполагается, что есть два удаленных друг от друга участника A и B, которым нужен общий двоичный ключ, т.е. двоичная последовательность $\kappa = (\kappa_1, \ldots, \kappa_m)$ длины m. Этот ключ участники хотели бы использовать для кодирования и декодирования своих сообщений, также представляющих собой двоичные последовательности длины m, путем побуквенного применения операции XOR: $y_k = x_k \oplus \kappa_k, x_k = y_k \oplus \kappa_k$:

$x_k \backslash \kappa_k$	0	1
0	0	1
1	1	0

Поскольку ключ должен быть секретным, важной проблемой является нахождение способа его передачи (распределения) обоим участникам, при котором ключ не может быть перехвачен или поврежден. Квантовая криптография предлагает такие способы (протоколы), надежность которых может быть в принципе сколь угодно высока и обеспечивается закономерностями квантовой информатики, такими как невозможность клонирования квантового состояния и дополнительность между квантовым измерением и возмущением состояния. Все излагаемые ниже протоколы используют тот факт, что состояния $|0\rangle, |1\rangle$ возмущаются при измерении в базисе $\{|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$ и, наоборот, состояния $|+\rangle, |-\rangle$ возмущаются при измерении в базисе $\{|0\rangle, |1\rangle\}$.

3.6.1 Протокол $BB84$

Дадим пошаговое описание протокола, предложенного Беннетом и Брассаром в 1984г.

1. Участник A генерирует две случайные двоичные последовательности $a = (a_1, \ldots, a_N)$, $b = (b_1, \ldots, b_N)$ длины $N = (4 + \delta)n$, $n \gg 1$: (предполагается, что биты a_k, b_l независимы и имеют распределение $\{\frac{1}{2}, \frac{1}{2}\}$).

2. Участник A создает чистое состояние с вектором

$$|\psi\rangle = \otimes_{k=1}^{N} |\psi_{a_k b_k}\rangle,$$

где

$$|\psi_{00}\rangle = |0\rangle, \qquad |\psi_{10}\rangle = |1\rangle,$$
$$|\psi_{01}\rangle = |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \qquad |\psi_{11}\rangle = |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \quad (3.10)$$

Таким образом, $b_k = 0$ означает выбор одного из состояний базиса $\{|0\rangle, |1\rangle\}$, а $b_k = 1$ – выбор одного из состояний базиса $\{|+\rangle, |-\rangle\}$. Значение $a_k = 0$ соответствует первому состоянию базиса, $a_k = 1$ – второму состоянию.

3. Участник A посылает это состояние участнику B по открытому квантовому каналу. Если канал идеален, т.е. шум либо постороннее вмешательство отсутствуют, то B получает состояние $|\psi\rangle\langle\psi|$, в противном случае – возмущенное состояние $\mathcal{E}(|\psi\rangle\langle\psi|)$, где $\mathcal{E}$ обозначает действие канала.

Участник B генерирует случайную последовательность $b' = (b'_1, \ldots, b'_N)$ и на k-м шаге производит измерение в базисе $\{|0\rangle, |1\rangle\}$, если $b'_k = 0$, или в базисе $\{|+\rangle, |-\rangle\}$, если $b'_k = 1$. Результаты его измерений образуют двоичную последовательность $a' = (a'_1, \ldots, a'_N)$.

4. Участник A посылает последовательность b участнику B по открытому классическому каналу, который сравнивает b с b' и сообщает A номера тех битов, для которых $b_k = b'_k$. Если канал идеален, то для этих битов с вероятностью 1 должно выполняться $a_k = a'_k$, поскольку измерение проводилось в базисе, содержащем посланное состояние. При этом с высокой вероятностью количество таких битов должно быть $\approx N/2 \approx 2n$ (см. ниже). Эти совпадающие биты последовательностей a и a' участники A и B оставляют себе в качестве просеянного ключа. Если же количество таких битов существенно отличается от $2n$, то это говорит о возможности постороннего вмешательства, поэтому A и B прекращают этот раунд протокола и пытаются повторить все сначала.

5. A и B выполняют тесты, чтобы определить величину возмущения из-за возможного шума или подслушивания. Для этого A наугад выбирает случайные n битов просеянного ключа и сообщает их B по открытому каналу. Если количество несовпадающих битов превосходит некоторый заранее выбранный порог, то A и B решают, что было вмешательство и также прекращают этот раунд протокола. Вероятность получить $\leq \delta n$ ошибок в этих n контрольных битах и при этом $\geq (\delta+\varepsilon)n$ ошибок в остальных битах имеет порядок $\exp[-O(\varepsilon^2 n)]$ (см. Упражнение 12.27 в [7]).

6. A и B выполняют согласование информации и усиление конфиденциальности по оставшимся $\approx n$ битам, чтобы получить m битов совместного секретного ключа. Эти последние операции, как и выбор порогового значения t, относятся к области классической информатики; их детальное описание можно найти, например, в [7].

Покажем, что при малых δ

$$\mathsf{P}\left\{(\#k : b_k = b'_k) \geq 2n\right\} \geq 1 - \exp\left[-\frac{1}{8}n\left(\delta^2 + o(\delta^2)\right)\right]. \qquad (3.11)$$

Вводя бит несовпадения $\nu_k = b_k \oplus b'_k$, имеем $\mathsf{P}\{\nu_k = 1\} = \frac{1}{2}$. Интересующая нас вероятность равна

$$\mathsf{P}\left\{\sum_{k=1}^{(4+\delta)n} \nu_k \geq 2n\right\} = 1 - \mathsf{P}\left\{\sum_{k=1}^{(4+\delta)n} \nu_k < 2n\right\},$$

при этом

$$\mathsf{P}\left\{\sum_{k=1}^{(4+\delta)n} \nu_k < 2n\right\} = \mathsf{P}\left\{\frac{1}{(4+\delta)n}\sum_{k=1}^{(4+\delta)n}\left(\nu_k - \frac{1}{2}\right) < -\varepsilon\right\},$$

где $\varepsilon = \frac{1}{2} - \frac{2n}{(4+\delta)n} = \frac{1}{8}\left(\delta^2 + o(\delta^2)\right)$. Интересующая нас оценка (3.11) вытекает тогда из *неравенства Хеффдинга*: если ν_k – последовательность независимых двоичных случайных величин с распределением $\mathsf{P}\{\nu_k = 1\} = p$,

то

$$\mathsf{P}\left\{\sum_{k=1}^{N}\nu_k \geq (p+\varepsilon)N\right\}, \mathsf{P}\left\{\sum_{k=1}^{N}\nu_k \leq (p-\varepsilon)N\right\} \leq \exp\left[-2N\varepsilon^2\right].$$

Мы докажем более грубое, но достаточное для наших целей *неравенство Чернова* (при $p=\frac{1}{2}$):

$$\mathsf{P}\left\{\frac{1}{N}\sum_{k=1}^{N}\left(\nu_k-\frac{1}{2}\right)\geq\varepsilon\right\}=\mathsf{P}\left\{\frac{1}{N}\sum_{k=1}^{N}\left(\nu_k-\frac{1}{2}\right)\leq-\varepsilon\right\}\leq\exp\left[-2N(\varepsilon^2+o(\varepsilon^2\right. \tag{3.12}$$

Доказательство. Пусть $s>0$, тогда

$$\mathsf{P}\left\{\frac{1}{N}\sum_{k=1}^{N}\left(\nu_k-\frac{1}{2}\right)\geq\varepsilon\right\}=\mathsf{P}\left\{\mathrm{e}^{s\sum_{k=1}^{N}\left(\nu_k-\frac{1}{2}\right)}\geq\mathrm{e}^{sN\varepsilon}\right\}\leq\mathrm{e}^{-sN\varepsilon}\mathsf{M}\mathrm{e}^{s\sum_{k=1}^{N}\left(\nu_k-\frac{1}{2}\right)}$$

$$=\mathrm{e}^{-sN\varepsilon}\left[\mathsf{M}\mathrm{e}^{s\left(\nu_k-\frac{1}{2}\right)}\right]^N=\mathrm{e}^{-sN\varepsilon}\left[\cosh\frac{s}{2}\right]^N=\mathrm{e}^{-N(s\varepsilon-\ln\cosh\frac{s}{2})}.$$

При малых значениях s

$$\ln\cosh\frac{s}{2}=\frac{s^2}{8}+o(s^2).$$

С другой стороны,

$$\max_{s>0}(s\varepsilon-\frac{s^2}{8})=2\varepsilon^2,$$

причем максимум достигается для $s=4\varepsilon$. Отсюда получаем (3.12).

3.6.2 Протокол $B92$

Этот протокол, предложенный Беннетом в 1992г., является усовершенствованием предыдущего.

1. Участник A генерирует случайную двоичную последовательность $a=(a_1,\ldots,a_N)$ и на k-м шаге создает чистое состояние с вектором $|\psi_0\rangle=|0\rangle$, если $a_k=0$ или $|\psi_1\rangle=|+\rangle=\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$, если $a_k=1$.

2. Участник A посылает состояние, задаваемое вектором

   $$|\psi\rangle=\otimes_{k=1}^{N}|\psi_{a_k}\rangle,$$

   участнику B по открытому квантовому каналу.

   Участник B генерирует случайную последовательность $a'=(a'_1,\ldots,a'_N$ и на k-м шаге производит измерение в базисе $\{|0\rangle,|1\rangle\}$, если $a'_k=0$, или в базисе $\{|+\rangle,|-\rangle\}$, если $a'_k=1$. Результаты его измерений образуют двоичную последовательность $b=(b_1,\ldots,b_N)$. Заметим, что $a_k=a'_k$ влечет $b_k=0$, поэтому из $b_k=1$ следует $a_k\neq a'_k$, т.е. $a_k=1-a'_k$.

3. Участник B посылает последовательность b участнику A по открытому классическому каналу. A (соответственно B) оставляет только те биты последовательности a (соответственно a'), для которых $b_k = 1$. Просеянные таким образом биты $a_k = 1 - a'_k$ и составляют общий секретный ключ.

4. Остальные шаги протокола совпадают с $BB84$.

3.6.3 Протокол $E91$

Протокол, предложенный Экертом в 1991г., предполагает распределение сцепленного состояния между участниками A и B.

Участники A и B получают "половинки" сцепленного состояния

$$|\psi\rangle = \otimes_{k=1}^{N} |\psi_k\rangle$$

где

$$|\psi_k\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right) = \frac{1}{\sqrt{2}} \left(|++\rangle + |--\rangle\right),$$

где первый q-бит посылается участнику A, а второй – B.

Задача 34. Проверьте это равенство, используя определения (3.10).

Из этого равенства следует, что на если на k-м шаге A и B проводят измерения в одинаковых базисах $\{|0\rangle, |1\rangle\}$ либо $\{|+\rangle, |-\rangle\}$, они получают совпадающие результаты, $\mathsf{P}\{a_k = a'_k\} = 1$.

A (соответственно B) генерирует случайную двоичную последовательность $b = (b_1, \ldots, b_N)$ (соответственно $b' = (b'_1, \ldots, b'_N)$). На k-м шаге A проводит измерение в базисе $\{|0\rangle, |1\rangle\}$, если $b_k = 0$, или в базисе $\{|+\rangle, |-\rangle\}$, если $b_k = 1$ (соответственно B использует b'_k) и получает результаты $a = (a_1, \ldots, a_N)$ (соответственно $a' = (a'_1, \ldots, a'_N)$). Соответственно B генерирует последовательность $b' = (b'_1, \ldots, b'_N)$, измеряет в базисе, выбранном по b'_k и получает результаты $a' = (a'_1, \ldots, a'_N)$.

Используя открытый классический канал, A и B сравнивают b и b' и оставляют в качестве просеянного ключа только те биты a (соответственно a'), для которых $b_k = b'_k$. В этом протоколе биты $a_k = a'_k$ не просто отбираются, но *создаются* в ходе измерений.

Описанные выше протоколы реализованы в эксперименте; более того, для первых двух созданы коммерческие образцы оборудования [7].

3.7 Нобелевская премия по физике 2012 г.

Нобелевскую премию по физике 2012 года получили Дэвид Уайнленд (Национальный Институт Стандартов и Технологий США (NIST)) и Серж Арош (Коллеж де Франс и Высшая Нормальная Школа, Франция) за новаторские экспериментальные методы, которые позволяют измерять и манипулировать индивидуальными квантовыми системами.

Д. Уайнленд получил международное признание благодаря исследованиям ионных ловушек, в которых отдельные электрически заряженные атомы удерживаются с помощью лазерного охлаждения при температуре, близкой к абсолютному нулю. Его достижения включают создание сцепленных состояний сначала двух, а потом и четырех ионов, что демонстрирует принципиальную возможность квантовых вычислений (в сообщении Нобелевского комитета говорится о перспективе создания квантового компьютера); демонстрацию квантовой телепортации состояний массивных частиц (одновременно с группой Цайлингера); создание атомных часов, в сотни раз превосходящих по точности существующие стандарты времени. Следует отметить, что эти эксперименты основываются на эпохальных открытиях теоретиков, предложивших эффективные квантовые алгоритмы и протоколы передачи информации.

В парижской лаборатории под руководством С. Арош физики работают с микроволновыми фотонами, которые удерживаются в полости, образованной миниатюрными, почти идеально отражающими сверхпроводящими зеркалами. Для измерений и управления состояниями фотонов используются сверхмассивные ридберговские атомы, которые с хорошим приближением можно считать макрочастицами. В результате их взаимодействия с фотонами возникают невозможные с точки зрения классической физики суперпозиции макроскопических состояний, в свое время гротескно описанные Шредингером на примере суперпозиции живого либо мертвого кота. Изощренная техника эксперимента позволяет в реальном времени отслеживать процесс декогерентизации – перехода подобных “раздвоенных” состояний в одно из классических.

В настоящее время создание квантовой сцепленности, распределенной в пространстве на макроскопические расстояния, остается трудной экспериментальной задачей, для которой известны “штучные” решения, подобные описанным выше. Широкое применение квантовых информационных технологий предполагает научно-техническую революцию, масштабы которой сейчас даже трудно представить[5].

[5] J. Preskill, Quantum computing and the entanglement frontier, arXiv:1203.5813.

Часть II

Глава 4

Квантовые измерения и разложения единицы

4.1 Статистический анализ понятия "наблюдаемая"

В первой части подчеркивалось, что во всяком физическом эксперименте присутствуют две основные стадии: приготовление состояния S и измерение M (наблюдаемых величин). Даже если приготовляется чистое квантовое состояние, в котором нет классической стохастичности, результат измерения все равно представляет собой случайную величину, распределение которой $\mu_S^M(x)$ зависит от приготовления ансамбля S и от измерительного прибора M. Если измеряется вещественная наблюдаемая X, то распределение $\mu_S^M(x)$ дается статистическим постулатом Борна-фон Неймана (1.21).

Заметим, что при этом смешивание ансамблей приводит к смешиванию распределений с теми же весами, т.е. если $S = \sum\limits_j p_j S_j$, то

$$\mu_S^M(x) = \sum_j p_j \mu_{S_j}^M(x). \tag{4.1}$$

Другими словами, вероятности исходов измерения являются *аффинными* функциями состояния. Это естественное и на первый взгляд слабое ограничение оказывается достаточным для того, чтобы, опираясь на выпуклую структуру множества квантовых состояний, логически вывести математическое описание (обобщенных) квантовых наблюдаемых и соответствующее обобщение "статистического постулата". В дальнейшем удобно считать, что множество исходов измерения $\mathcal{X}$ – произвольное конечное множество (не обязательно подмножество вещественных чисел).

Теорема 11. *Пусть $S \to \mu_S$ отображение квантовых состояний в вероятностные распределения на некотором конечном множестве исходов $\mathcal{X}$. Если отображение аффинно, то есть обладает свойством (4.1),*

то существует единственное разложение единицы[1] *в* $\mathcal{H}$*, т.е. семейство эрмитовых операторов* $\{M_x; x \in \mathcal{X}\}$ *такое, что*

$$M_x \geq 0, \quad \sum_{x\in\mathcal{X}} M_x = I \tag{4.2}$$

для которого

$$\mu_S(x) = \operatorname{Tr} SM_x. \tag{4.3}$$

Обратно, для всякого разложения единицы в $\mathcal{H}$ *соотношение (4.3) определяет аффинное отображение* $S \to \mu_S$ *квантовых состояний в вероятностные распределения на* $\mathcal{X}$*.*

Обратное утверждение почти очевидно: неотрицательность чисел $\mu_S(x)$ вытекает из первого условия в (4.2) и (1.14), вероятностная нормировка – из второго условия в (4.2), а аффинность – из линейности следа в (4.3). Доказательство прямого утверждения дано в [11].

Разложение единицы называется *ортогональным*, если

$$M_x^2 = M_x, \quad M_xM_y = 0, \quad x \neq y, \quad x, y \in \mathcal{X}.$$

З а д а ч а 35. Второе условие является следствием первого. Таким образом, ортогональное разложение единицы характеризуется свойством: все операторы M_x — проекторы.

Основываясь на этих понятиях, можно было бы назвать *обобщенной квантовой наблюдаемой* со значениями в $\mathcal{X}$ разложение единицы $M = \{M_x; x \in \mathcal{X}\}$ в гильбертовом пространстве системы $\mathcal{H}$. Распределение вероятностей такой наблюдаемой в состоянии S дается обобщением (4.3) статистического постулата Борна-фон Неймана. Спектральное разложение

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

где $\mathcal{X} \subset \mathbb{R}$, задает взаимно-однозначное соответствие между эрмитовыми операторами X (вещественными наблюдаемыми в стандартной статистической модели квантовой механики, см. ч. I) и ортогональными разложениями единицы $E = \{E_x\}$ в пространстве системы $\mathcal{H}$.

Чтобы уточнить используемую терминологию, а также пояснить 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" $\Omega = \{\omega\}$. Диагональный эрмитов оператор $X = \sum\limits_\omega x_\omega |\omega\rangle\langle\omega|$

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

может быть записан в виде

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

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

$$\sum_x M(x|\omega) = 1, \quad \omega \in \Omega, \tag{4.4}$$

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 ортогональных разложений единицы, удовлетворяющих условию $M_x^2 = M_x$ и соответствующих обычным случайным величинам, эти вероятности принимают только значения 0 или 1.

В современных текстах разложения единицы называются просто наблюдаемыми, тогда как ортогональные разложения единицы – *четкими* наблюдаемыми. В дальнейшем нам будет удобно придерживаться именно такой терминологии.

4.2 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 обобщенных 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 12. *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 (1.17)

$$pM_x^1(I - M_x^1) + (1-p)M_x^2(I - M_x^2) + p(1-p)(M_x^1 - M_x^2)^2 = 0. \qquad (4.5)$$

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 (4.4) 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.

4.3 Переполненные системы векторов

Система векторов $\{|\psi_j\rangle; j = 1, \ldots, n\} \subset \mathcal{H}$ называется *переполненной*, если

$$\sum_{j=1}^{n} |\psi_j\rangle\langle\psi_j| = I.$$

Другими словами

$$\sum_{j=1}^{n} |\langle\psi|\psi_j\rangle|^2 = \langle\psi|\psi\rangle, \quad \psi \in \mathcal{H}.$$

С необходимостью $n \geq d$, так как в противном случае обязательно найдется $\psi \neq 0$, ортогональный всем ψ_j, что невозможно ввиду предыдущего равенства.

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

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

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

Задача 36. Система $\{|\psi_j\rangle\}$ является переполненной тогда и только тогда, когда
1) система полна, т.е. $\{|\psi_j\rangle; j = 1, \ldots, n\}^\perp = \{0\}$;
2) матрица $P = [\langle\psi_j|\psi_k\rangle]_{j,k=1,\ldots,n}$ идемпотентна, т.е. $P = P^2$.

Пусть $\{|\phi_j\rangle\}$ — произвольная полная (не обязательно ортонормированная) система векторов. Тогда в силу 1) ее *оператор Грама*

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

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

Покажем, что *всякая* переполненная система в подпространстве гильбертова пространства возникает при проецировании ортонормированного базиса пространства на это подпространство. Рассмотрим отображение

$$V: \quad |\psi\rangle \to [\langle\psi_j|\psi\rangle]_{j=1,\ldots,n} = \begin{bmatrix} \langle\psi_1|\psi\rangle \\ \vdots \\ \langle\psi_n|\psi\rangle \end{bmatrix},$$

которое является изометрическим вложением пространства $\mathcal{H}$ в $\widetilde{\mathcal{H}} = \mathbb{C}^n$. В самом деле, для любого вектора $|\psi\rangle \in \mathcal{H}$

$$\|V\psi\|^2 = \sum_{j=1}^n \langle\psi|\psi_j\rangle\langle\psi_j|\psi\rangle = \|\psi\|^2.$$

Это отображение позволяет отождествить $\mathcal{H}$ с подпространством $V\mathcal{H} \subseteq \widetilde{\mathcal{H}}$. Исходной переполненной системе $\{|\psi_k\rangle\}$ соответствует система $\{V|\psi_k\rangle\} \subseteq V\mathcal{H}$. Матрица $P = [\langle\psi_j|\psi_k\rangle]_{j,k=1,\ldots,n}$ задает проекцию пространства $\widetilde{\mathcal{H}} = \mathbb{C}^n$ на $V\mathcal{H}$. Рассмотрим вычислительный ортонормированный базис в $\mathbb{C}^n$:

$$|e_k\rangle = [\delta_{jk}]_{j=1,\ldots,n}; \quad k = 1, \ldots, n.$$

Проецируя векторы этого базиса на $V\mathcal{H}$, получаем переполненную систему

$$V|\psi_k\rangle = P|e_k\rangle, \quad k = 1, \ldots, n,$$

изометричную исходной.

В дальнейшем будет доказана теорема, принадлежащая М.А. Наймарку, которая обобщает этот результат на произвольное разложение единицы.

Очевидно, что с каждой переполненной системой связано разложение единицы, т.е. наблюдаемая

$$M_j = |\psi_j\rangle\langle\psi_j|. \tag{4.6}$$

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

$$M_j = G^{-1/2}|\phi_j\rangle\langle\phi_j|G^{-1/2} \tag{4.7}$$

задает наблюдаемую, в определенном смысле "измеряющую" состояния $|\phi_j\rangle\langle\phi$

Т е о р е м а 13. *Наблюдаемая (4.6) является экстремальной тогда и только тогда, когда операторы M_x линейно независимы.*

Доказательство. Пусть M – крайняя точка, и предположим, что

$$\sum_x c_x|\psi_x\rangle\langle\psi_x| = 0. \tag{4.8}$$

Взяв достаточно малое $\epsilon > 0$, определим

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

Тогда $M^{\pm} = \{M_x^{\pm}\}$ 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 (4.8) 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

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

откуда

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

[2] 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*.

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. □

4.4 Переполненные системы для q-бита

Теорема 14. *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*

$$\frac{2}{m}\sum_{j=1}^m |\vec{a}_j\rangle\langle\vec{a}_j| = I. \tag{4.9}$$

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 (4.9), непосредственно следует из (1.27). Также используя это соотношение получаем, что линейная зависимость операторов $|\vec{a}_j\rangle\langle\vec{a}_j|$, равносильная, в силу теоремы 13, 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. $|\vec{a}_1\rangle = |0\rangle, |\vec{a}_2\rangle = |1\rangle$ и ортогональное разложение единицы

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

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

$$\sqrt{\frac{2}{3}}\,|\vec{a}_1\rangle = \sqrt{\frac{2}{3}}\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \sqrt{\frac{2}{3}}\,|\vec{a}_{2,3}\rangle = \sqrt{\frac{2}{3}}\begin{bmatrix} 1/2 \\ \pm\sqrt{3}/2 \end{bmatrix}$$

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 $\mathcal{H}$

$$\sqrt{\frac{1}{2}}\,|\vec{a}_1\rangle = \sqrt{\frac{1}{2}}\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \sqrt{\frac{1}{2}}\,|\vec{a}_2\rangle = \sqrt{\frac{1}{2}}\begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix},$$

$$\sqrt{\frac{1}{2}}\,|\vec{a}_{3,4}\rangle = \sqrt{\frac{1}{2}}\begin{bmatrix} 1/\sqrt{3}\,(-1/2 \mp i\sqrt{3}/2) \\ \sqrt{2}/\sqrt{3}\,(-1/2 \pm i\sqrt{3}/2) \end{bmatrix}$$

и неортогональное разложение единицы

$$M_1 = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_2 = \frac{1}{2}\begin{bmatrix} 1/3 & \sqrt{2}/3 \\ \sqrt{2}/3 & 2/3 \end{bmatrix},$$

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

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

4.5 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 (которое имеет вещественную размерность 4). Таким образом, вероятности

$$\mu_S(j) = \operatorname{Tr} SM_j$$

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

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

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

$$\operatorname{Tr} M_j M_k = c, \quad j \neq k,$$

причем константа оказывается равной $[d^2(d+1)]^{-1}$. В самом деле,

$$d = \operatorname{Tr} I^2 = \sum_{j,k=1}^{d^2} \operatorname{Tr} M_j M_k = \sum_{j=1}^{d^2} \operatorname{Tr} M_j^2 + \sum_{j \neq k} \operatorname{Tr} M_j M_k = 1 + d^2(d^2-1)c.$$

Существование SIC-POVM показано аналитически, либо численно, для $d \leq 67$. Имеется гипотеза, что они существуют во всех размерностях[3].

Задача 37. Покажите, что любое состояние S восстанавливается по формуле

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

В силу информационной полноты, для этого достаточно проверить, что $\operatorname{Tr} S M_k = \operatorname{Tr} S' M_k$, где S' – оператор в правой части равенства.

Восстановление состояния по статистике измерений (одного или целого ряда) называют *томографией* квантового состояния. Например, формула (1.25) 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:

$$|\langle e_j | h_k \rangle|^2 = \text{const} \tag{4.10}$$

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

4.6 Теорема Наймарка

Геометрический смысл неортогональных разложений единицы проясняет следующая теорема.

Теорема 15. *Пусть $\{M_x\}_{x \in \mathcal{X}}$ – разложение единицы в гильбертовом пространстве $\mathcal{H}$, $\dim \mathcal{H} = d$, $|\mathcal{X}| = n$. Существует гильбертово пространство $\widetilde{\mathcal{H}}$, $\dim \widetilde{\mathcal{H}} \leq n \cdot d$, изометрический оператор $V : \mathcal{H} \to \widetilde{\mathcal{H}}$ и ортогональное разложение единицы $\{E_x\}$ в $\widetilde{\mathcal{H}}$, такие, что*

$$M_x = V^* E_x V.$$

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

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 $\mathcal{H}$:

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

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

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

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

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

Задача 38. 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 $\mathcal{H}_n$

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

где $\mathbf{0}$ обозначает нулевые компоненты вектора. При этом $\langle\psi|V^*E_yV|\psi'\rangle = \langle\psi|M_y|\psi'\rangle$. □

Эта теорема обобщает утверждение п. 4.3 о том, что всякая переполненная система является проекцией ортонормированного базиса. Далее с ее помощью будет прояснен статистический смысл нечетких наблюдаемых.

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

Следствие. *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*

$$M_j = \mathrm{Tr}_{\mathcal{H}_0}(I \otimes |\psi_0\rangle\langle\psi_0|)E_j. \tag{4.11}$$

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

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

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 $\psi \in \mathcal{H}$:

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

т.е. соотношение (4.11). □

Итак, всякую наблюдаемую можно реализовать в виде четкой 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 внешнего генератора случайности при принятии решений, хотя и может оказаться полезным приемом (например, в теории игр), никогда не увеличивает информации о состоянии наблюдаемой системы. Далее мы покажем, что в квантовой статистике это уже не так: парадоксальным образом, квантовая рандомизация позволяет извлекать больше информации о наблюдаемой системе, нежели содержится в четких наблюдаемых, не использующих вспомогательной системы.

4.7 Оптимальное различение квантовых состояний

4.7.1 Постановка задачи

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

Предположим, что a priori квантовая система может находиться в одном из состояний S_j, $j = 1, \ldots, n$. Над системой можно производить произвольное измерение. Требуется найти *оптимальное* измерение, позволяющую наилучшим образом выяснить, в каком из этих состояний действительно находится данная 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 значение j имеет априорную вероятность π_j (например, в случае равновероятных сигналов $\pi_j = 1/n$.) Тогда средняя вероятность правильного решения

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

и задача состоит в ее максимизации $\mathcal{P}\{M\}$.

4.7.2 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 16. *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

$$\max\{\mathcal{P}\{M\} : M \in \mathfrak{M}_n\} = \min\{\operatorname{Tr}\Lambda : \Lambda \geq W_k, k = 1, \ldots, n\}. \tag{4.12}$$

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 39. 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 $A_1 B = A_2 B$.

Теперь докажем необходимость.

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

$$\operatorname{Tr}\sum_k W_k X_k^2 - \operatorname{Tr}\Lambda(\sum_k X_k^2 - I), \tag{4.13}$$

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 (4.13) 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 40. Доказать, что операторный множитель Лагранжа Λ является единственным решением двойственной задачи в правой части (4.12).

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. *Различение классических состояний.* Пусть операторы $W_k = \pi_k S_k$ коммутируют, тогда существует 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. теорему 12).

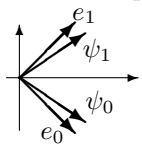


Рис. 4.1: Различение двух чистых состояний.

П р и м е р 2. *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 четкие наблюдаемые характеризуются условием $M_0^2 = M_0$, которое в точности соответствует крайним точкам "некоммутативного отрезка" $\{0 \leq M_0 \leq I\}$ (задача 41). Таким образом, для различения двух состояний достаточно четких 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. [10], гл. IV). Пусть для простоты $\pi_0 = \pi_1 = 1/2$, тогда оптимальный базис расположен симметрично по отношению к $|\psi_0\rangle, |\psi_1\rangle$ (рис. 4.1) и

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

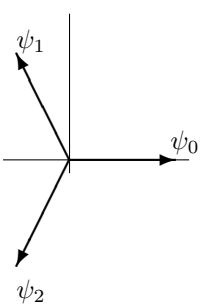


Рис. 4.2: *Векторы трех состояний*

Пример 3. На плоскости (рассматриваемой как вещественное подпространство двумерного унитарного пространства) рассмотрим "равноугольную" конфигурацию трех векторов (рис. 4.2)

$$|\psi_j\rangle = \begin{bmatrix} \cos\frac{2j\pi}{3} \\ \sin\frac{2j\pi}{3} \end{bmatrix}, \quad j = 0, 1, 2. \tag{4.14}$$

Соответствующие операторы плотности $S_j = |\psi_j\rangle\langle\psi_j|$, описывают состояния двухуровневой системы, например, плоскополяризованного фотона или частицы со спином 1/2, см., например, [7].

Имеем

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

$$= \frac{1}{2}\left(I + \begin{bmatrix} \cos\frac{4j\pi}{3} & \sin\frac{4j\pi}{3} \\ \sin\frac{4j\pi}{3} & -\cos\frac{4j\pi}{3} \end{bmatrix}\right). \tag{4.15}$$

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 четким 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. Таким образом, задача сводится к оптимальному различению только двух состояний S_0, S_1. Подставляя решение из примера 2, получаем

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

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 существуют крайние точки множества наблюдаемых $\mathfrak{M}_3$ (среди которых и находится наиболее информативная экстремальная наблюдаемая), которые не описываются ортогональными разложениями единицы.

Глава 5

Классически-квантовые каналы связи

5.1 Основные понятия классической теории информации

5.1.1 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

$$H(X) = -\sum_{x \in \mathcal{X}} p_x \log p_x, \tag{5.1}$$

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

Задача 42. $0 \leq H(X) \leq \log |\mathcal{X}|$, 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 17. *Если* $X_1, \ldots, X_n, \ldots$ *независимые и одинаково распределенные случайные величины с распределением* $p = \{p_x\}$, *то*

$$-\frac{1}{n}\sum_{i=1}^{n}\log p_{x_i} \longrightarrow H(X) \qquad \text{по вероятности.} \tag{5.2}$$

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

$$\mathsf{P}\left\{\left| -\frac{1}{n}\sum_{i=1}^{n}\log p_{x_i} - H(X)\right| < \delta\right\} > 1-\epsilon. \tag{5.3}$$

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

$$p_w = p_{x_1} \cdot \ldots \cdot p_{x_n} = 2^{-n\left(-\frac{1}{n}\sum_{i=1}^{n}\log p_{x_i}\right)} \tag{5.4}$$

мы теперь можем использовать соотношение (5.3) чтобы ввести понятие *типичного слова*: слово w, имеющее вероятность p_w, называется δ-типичным, если

$$2^{-n(H(X)+\delta)} < p_w < 2^{-n(H(X)-\delta)}. \tag{5.5}$$

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 43).

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.

5.1.2 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*:

$$I(X;Y) = H(X) - H(X|Y), \tag{5.6}$$

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 (5.6), мы видим, что оно симметрично по X и Y, и поэтому может быть также названо *взаимной информацией*

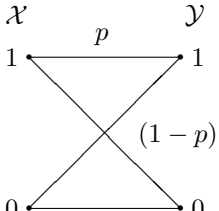
$$I(X;Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X), \tag{5.7}$$

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 44). Отсюда также вытекает свойство субаддитивности энтропии: $H(XY) \leq H(X) + H(Y)$. Далее, $I(X;Y) = 0$ (т.е. $H(XY) = H(X) + H(Y)$) тогда и только тогда, когда X и Y независимые случайные величины: $p_{x,y} = p_x \cdot p_y$.

Шенноновская пропускная способность определяется как

$$C = \max_{\{p_x\}} I(X;Y), \tag{5.8}$$

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



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*

$$h(p) = -p \log p - (1-p) \log(1-p), \tag{5.9}$$

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

$$C = 1 - h(p), \tag{5.10}$$

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

Если посылается последовательность букв $x_1, x_2, dots$, и $p(y|x)$ действует независимо на каждую посланную букву, то такой составной канал называется *каналом без памяти*. Применяя *блочное кодирование* для канала без памяти, когда канал используется для посылки n букв, имеем

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

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: $C_{n+m} \geq C_n + C_m$. На самом деле имеет место аддитивность:

Лемма 3.

$C_n = nC_n$.

Доказательство. Покажем, что

$$I(X^n; Y^n) \leq \sum_{i=1}^{n} I(X_i; Y_i). \tag{5.11}$$

Имеет место *цепное правило* для условной энтропии:

$$H(X_1, \ldots, X_n) = \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_{i-1}), \tag{5.12}$$

которое легко доказать по индукции, используя формулу:

$$H(X, Y) = H(X) + H(Y|X). \tag{5.13}$$

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,

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

Беря максимум выражения (5.11), получаем аддитивность $C_n = nC$. □

Определение. *Кодом* (W,V) *размера* N для канала $p(y|x)$ называется совокупность N слов $w^{(1)},\dots,w^{(N)}$ длины n вместе с разбиением множества $\mathcal{Y}^n$ на N непересекающихся подмножеств $V^{(0)},V^{(1)},\dots,V^{(N)} \subset \mathcal{Y}^n$. Подмножества $V^{(1)},\dots,V^{(N)}$ интерпретируются как области принятия решения: если на выходе принято значение $y^n \in V^{(j)}; j = 1,\dots,N$, 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

$$P_e(W,V) = \max_{1\le j\le N} \left(1 - p(V^{(j)}|w^{(j)})\right), \tag{5.14}$$

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

$$\overline{P}_e(W,V) = \frac{1}{N}\sum_{i=1}^{N} \left(1 - p(V^{(j)}|w^{(j)})\right) \le P_e(W,V), \tag{5.15}$$

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

Лемма 4. *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. □

Теорема 18 (Теорема кодирования для канала без памяти). *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. Теорема кодирования раскрывает, таким образом, операциональный смысл шенноновской пропускной способности как максимальной скорости асимптотически безошибочной передачи информации через данный канал связи без памяти.

Доказательство слабого обращения. Лемма 5 (Неравенство Фано). *Пусть X, Y случайные величины и $\widehat{X} = \widehat{X}(Y)$ оценка случайной величины X с вероятностью ошибки $P_e = P(\widehat{X}(Y) \neq X)$, тогда*

$$H(X|Y) \leq h(p_e) + p_e \log(|\mathcal{X}| - 1) \leq 1 + p_e \log |\mathcal{X}|. \tag{5.16}$$

Доказательство. Пусть E индикатор ошибки оценивания,

$$E = \begin{cases} 0, & \text{если } \widehat{X}(Y) = X \\ 1, & \text{в противном случае.} \end{cases} \tag{5.17}$$

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

$$H(E|X, Y) = H(E, X|Y) - H(X|Y) = 0, \tag{5.18}$$

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. □

Теперь 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 $N = 2^{nR}$, оптимизируя по W, V и деля на n, получаем

$$\frac{C}{R} \geq (1 - p_e(n, 2^{nR})) - \frac{1}{nR},$$

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

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

□

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

$$N \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}. \tag{5.19}$$

Таким образом, $N \approx 2^{nC}$. Конечно, это рассуждение в высшей степени эвристично; строгое доказательство, реализующее эту идею, можно найти, например, в [13], [14].

5.2 Сжатие квантовой информации

Выше уже было отмечено, что квантовая информация — это новый вид информации, который можно передавать, но нельзя размножать. Пусть имеется *квантовый 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 $p_w = p_{x_1} \cdot \ldots \cdot p_{x_n}$. *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

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

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

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:

$$H(S) = -\sum_j s_j \log s_j = -\operatorname{Tr} S \log S. \tag{5.20}$$

Далее нам понадобятся элементарные свойства квантовой энтропии, которые вытекают из определения и соответствующих свойств классической энтропии Шеннона (Задача 45.)

:

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 19[2]. *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.

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

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

$$\overline{S}_p = \sum_j \lambda_j |e_j\rangle\langle e_j|. \tag{5.21}$$

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:

$$\dim E\mathcal{H}^{\otimes n} = \operatorname{Tr} E \leqslant \operatorname{Tr} \frac{\overline{S}_\pi}{2^{-n(H(\overline{S}_\pi)+\delta)}} \leqslant 2^{n(H(\overline{S}_\pi)+\delta)}. \tag{5.22}$$

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

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

$$= \operatorname{Tr} E\Big(\sum_w p_w |\psi_w\rangle\langle\psi_w|\Big) = \operatorname{Tr} E S_p^{\otimes n} = \sum_{J\in J_{n,\delta}} \lambda_J. \tag{5.23}$$

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

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

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

$$= \mathsf{P}\Big\{\Big|-\frac{1}{n}\sum_{k=1}^n \log\lambda_{j_k} - H(\overline{S}_p)\Big| < \delta\Big\}, \tag{5.24}$$

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

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

$$\leqslant d 2^{-n(H(\overline{S}_p)-\delta/2)} + \frac{\varepsilon}{2} \leqslant 2^{-n\delta/2} + \frac{\varepsilon}{2} < \varepsilon \qquad (5.25)$$

для достаточно больших n.

5.3 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). Напомним, что классический канал задается переходной вероятностью $p(y|x)$ из входного алфавита $\mathcal{X}$ в выходной алфавит $\mathcal{Y}$. Эквивалентно, его можно рассматривать как отображение $x \to P_x$, переводящее буквы входного алфавита x в распределения вероятностей $P_x(y) = p(y|x)$ на выходном алфавите $\mathcal{Y}$. Распределение вероятностей P_x задает классическое статистическое состояние на выходном алфавите, которое описывает результат воздействия шума на классический сигнал x.

Это естественно приводит к модели *классически-квантового канала* как отображения $x \to S_x$ входного алфавита $\mathcal{X}$ в квантовые состояния S_x в гильбертовом пространстве, описывающем выход канала. Например, двоичный оптический классически-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 на выходе канала может также играть поляризация или направление спина.

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

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

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

На выходе канала "приемник" производит 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 $\widehat{w}$; 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

$$I_n(p, M)$$

$$= \sum_w p_w \sum_{\widehat{w}} p_{M^{(n)}}(\widehat{w}|w) \Big[\log p_{M^{(n)}}(\widehat{w}|w) - \log \sum_{w'} p_{M^{(n)}}(\widehat{w}|w') p_{w'} \Big],$$

$$\max_{p^{(n)}, M^{(n)}} I_n(p^{(n)}, M^{(n)}) = C_n.$$

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.

Определение. *Кодом* (W, M) длины n и размера N называется набор слов $W = \{w^{(1)}, \ldots, w^{(N)}\}$ вместе с разложением единицы $M = \{M_j\}$ в $\mathcal{H}^{\otimes n}$ с исходами $j = 0, 1, \ldots, N$; 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 $\min\limits_{W,M} \overline{P}_e(W, M) = p_e(n, N)$ 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 (5.20).

Теорема 20 (Квантовая теорема кодирования[3]). *При* $n \to \infty$

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

2) $p_e(n, 2^{nR}) \nrightarrow 0$, *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

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

[3] 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.

Из свойства 1) энтропии следует, что всегда

$$C_\chi \leq \log d. \tag{5.26}$$

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 асимптотического достижения пропускной способности идеального канала.

Рассмотрим конфигурацию (4.14) 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. С другой стороны, можно показать[4], что 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 S_0, S_1 (см. пример 1 в разделе 4.3).

5.4 Квантовая граница классической информации и доказательство обратной теоремы

Теорема 21 (Квантовая граница классической информации). *Для любого распределения p и любой наблюдаемой M*

$$I_1(p, M) \leq H\left(\sum p_x S_x\right) - \sum p_x H(S_x), \tag{5.27}$$

[4] M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, O. Hirota, "Accessible information and optimal strategies for real symmetric quantum sources", Phys. Rev. A **59**, 3325, 1999.

причем имеет место строгое неравенство, если среди операторов p_iS_i есть некоммутирующие.

Задача 46. Если все операторы p_iS_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_iS_i.

Приведем здесь первое, прямое доказательство этой теоремы[5], опирающееся на исследование свойства выпуклости квантовой энтропии. Впоследствии было установлено более общее свойство монотонности относительной энтропии (доказательство которого, однако, не менее сложно, но более формально), из которого вытекает и неравенство (5.27), см. часть II.

Отметим также, что очевидным следствием неравенства (5.27) является вогнутость квантовой энтропии как функции операторов плотности:

$$H\left(\sum p_xS_x\right) - \sum p_xH(S_x) \geq 0. \tag{5.28}$$

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

$$\chi(t) = H(S_t) - (1-t)H(S_0) - tH(S_1), \quad t \in [0,1]. \tag{5.29}$$

Пусть $M = \{M_y\}$ – произвольная наблюдаемая, $P_t(y) = \operatorname{Tr} S_tM_y = (1-t)P_0(y) + tP_1(y)$ – 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 $I_M(t)$:

$$\chi(t)'' \leq I_M(t)'', \quad t \in [0,1]. \tag{5.30}$$

Отсюда, очевидно, следует

$$\chi(t) \geq I_M(t), \quad t \in [0,1]. \tag{5.31}$$

Положим $D = S_1 - S_0$ и пусть

$$S_t = \sum_k s_kE_k$$

– спектральное разложение оператора S_t. Доказательство следующей леммы[6] опирается на интегральную формулу Коши для матричных функций.

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

[6] ibid.

Лемма 6.

$$\chi''(t) = -\sum_{k,j} (\operatorname{Tr} E_k D E_j D) f(s_k, s_j), \quad t > 0, \tag{5.32}$$

где

$$f(a,b) = \frac{\log a - \log b}{a - b}, \quad a \neq b; \quad f(a,a) = a^{-1}. \tag{5.33}$$

Используя элементарное неравенство

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

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

$$\chi''(t) \leq -\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j}, \tag{5.34}$$

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 47. 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

$$\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j} = \operatorname{Tr} D L_t = \operatorname{Tr} S_t L_t^2. \tag{5.35}$$

Оператор L_t является некоммутативным аналогом логарифмической производной семейства S_t, а (5.35)– аналогом информационного количества Фишера в математической статистике.

Из (5.34), (5.35) вытекает, что

$$\chi''(t) \leq -\operatorname{Tr} S_t L_t^2, \tag{5.36}$$

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

$D(y) = P_1(y) - P_0(y) = \operatorname{Tr} DM_y$. Применяя полученные результаты к диагональной матрице $\operatorname{diag}[P_t(y)]$ в роли состояния S_t и учитывая коммутативность диагональных матриц, получаем вместо (5.36)

$$I''_M(t) = -\sum_y \frac{D(y)^2}{P_t(y)}. \tag{5.37}$$

Доказательство неравенства (5.30). 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 некоммутативного неравенства Коши- Буняковского 1.41 получаем $D(y) \le \operatorname{Tr} S_t M_y \cdot \operatorname{Tr} L_t S_t L_t M_y$. Подставляя в (5.37), 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[7]. □

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 (5.27) вытекает неравенство

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

[7] ibid.

Лемма 7. *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 $\max\limits_{p_{ij}} \geq \max\limits_{p_{ij}=p_i p_j}$, тогда в силу свойства аддитивности квантовой энтропии

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

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. □

5.5 Доказательство прямой теоремы для канала с чистыми состояниями

Доказательство прямого утверждения теоремы кодирования дадим в наиболее простом случае чистых состояний $S_x = |\psi_x\rangle\langle\psi_x|$[8], когда

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

Доказательство непросто уже в этом случае, тогда как классический аналог этой проблемы тривиален, поскольку чистые состояния с необходимостью ортогональны. Доказательство в общем случае см. в [13].

[8] 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.

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 5.2. 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 $\overline{S}_p^{\otimes n}$.

Положим

$$|\widetilde{\psi}_{w^{(j)}}\rangle = E|\psi_{w^{(j)}}\rangle \in \mathcal{H}^{n,\delta} \tag{5.38}$$

и пусть $G = \sum\limits_{j=1}^{N} |\widetilde{\psi}_{w^{(j)}}\rangle\langle\widetilde{\psi}_{w^{(j)}}|$ — оператор Грама системы (5.38). 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" (4.7):

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

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

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

$$= \frac{1}{N}\sum_{j=1}^{N}(2\operatorname{Tr} S_{w^{(j)}} - 2\operatorname{Tr} S_{w^{(j)}}G^{-1/2}). \tag{5.39}$$

Получим теперь удобную оценку для $G^{-1/2}$, асимптотически точную при $G \simeq E$. Имеем

$$-2x^{-1/2} \le -3 + x, \quad x \ge 0,$$

(причем линейная функция в правой части является касательной к левой части при $x = 1$). Отсюда следует, что

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

Подставляя в (5.39), получаем

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

Учитывая, что

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

и

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

получаем окончательную оценку

$$\overline{P}_e(W; M) \le \frac{1}{N} \sum_{j=1}^{N} [2 \operatorname{Tr} S_{w^{(j)}} (I - E) + \sum_{k \ne j} \operatorname{Tr} E S_{w^{(j)}} E S_{w^{(k)}} E]. \qquad (5.40)$$

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

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

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

Согласно свойствам типичного подпространства, для достаточно больших n

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

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

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

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

Глава 6

КВАНТОВЫЕ КАНАЛЫ

6.1 Вполне положительные отображения

В этом разделе понятие квантового канала будет рассмотрено с общей точки зрения. Это позволит исследовать более реалистичные модели, для которых возникают и более интересные вопросы.

Напомним, что классический канал задается переходной вероятностью $p(y|x)$ из входного алфавита $\mathcal{X}$ в выходной алфавит $\mathcal{Y}$. Эквивалентным образом, он задается линейным отображением

$$p'_y = \sum_x p(y|x)p_x, \tag{6.1}$$

которое переводит распределения вероятностей (классические состояния) $P = \{p_x\}$ на входном алфавите в распределения вероятностей $P' = \{p'_y\}$ на выходном алфавите.

Каковы минимальные требования к теоретическому описанию квантового канала связи? Всякий канал должен преобразовывать состояния на входе в состояния на выходе. Необходимое требование для согласованности со статистической интерпретацией состоит в том, чтобы смеси входных состояний переходили в аналогичные смеси выходных состояний, т.е. канал должен задаваться аффинным отображением:

$$\Phi[\sum_j p_j S_j] = \sum_j p_j \Phi[S_j], \quad p_j \geq 0, \quad \sum_j p_j = 1. \tag{6.2}$$

Пусть Φ – отображение алгебры операторов $\mathcal{B}(\mathcal{H})$ на входе в алгебру операторов $\mathcal{B}(\mathcal{H}')$ на выходе, обладающее следующими свойствами:

1) $\Phi[\sum c_j T_j] = \sum c_j \Phi[T_j]$ (*линейность*);
2) $T \geq 0 \Rightarrow \Phi[T] \geq 0$ (*положительность*);
3) $\mathrm{Tr}\,\Phi[T] = \mathrm{Tr}\,T$ (*сохранение следа*).

Тогда его ограничение на множество $\mathcal{S}(\mathcal{H})$ переводит состояния в состояния и обладает свойством (6.2), более того, таким образом можно получить всякое аффинное отображение множества состояний [13].

Однако для содержательного определения квантового канала требуется следующее ключевое свойство, усиливающее положительность:

Определение. Линейное отображение Φ алгебры $\mathcal{B}(\mathcal{H})$ называется *вполне положительным*, если выполняется условие

Для любой блочной положительно определенной матрицы

$$[X_{ij}]_{i,j=1,\dots,m} \geq 0,$$

выполняется

$$[\Phi^*[X_{ij}]]_{i,j=1,\dots,m} \geq 0.$$

Дадим другую формулировку, используя изоморфизм (см. п. 2.2)

$$\mathcal{H} \otimes \mathbb{C}^m \approx \underbrace{\mathcal{H} \oplus \dots \oplus \mathcal{H}}_{m}, \quad X \approx [X_{ij}]_{i,j=1,\dots,m},$$

где X_{ij} — операторы в $\mathcal{H}$.

Тогда условие полной положительности можно переформулировать в виде:

Для любого $m = 1, 2, \dots$ отображение $\Phi \otimes \mathrm{Id}_m$, где Id_m — тождественное отображение алгебры всех $m \times m$-матриц, положительно.

Здесь мы используем естественное определение тензорного произведения отображений:

$$(\Phi^{(1)} \otimes \Phi^{(2)})(X^{(1)} \otimes X^{(2)}) = \Phi^{(1)}(X^{(1)}) \otimes \Phi^{(2)}(X^{(2)})$$

на элементах $X^{(1)} \otimes X^{(2)}$, порождающих алгебру $\mathcal{B}(\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)})$.

Определение. *Квантовым каналом* называется линейное, вполне положительное, сохраняющее след отображение Φ из $\mathcal{B}(\mathcal{H})$ в $\mathcal{B}(\mathcal{H}')$.

Задача 48. Докажите, что *последовательное* применение каналов $\Phi^{(2)} \circ \Phi^{(1)}$, где $\circ$ означает композицию отображений, определяет канал.

Свойство полной положительности означает, что *параллельное* использование канала Φ и тождественного канала Id_m также задает канал. Отсюда следует, что тензорное произведение каналов является каналом, поскольку его можно представить как последовательное применение двух каналов вида $\Phi \otimes \mathrm{Id}$:

$$\Phi^{(1)} \otimes \Phi^{(2)} = [\mathrm{Id}^{(1)} \otimes \Phi^{(2)}] \circ [\Phi^{(1)} \otimes \mathrm{Id}^{(2)}].$$

В теории квантовых вычислений каналы представляют собой неидеальные вентили (элементарные операции, подверженные случайным искажениям). В принципе любая квантовая схема представляет собой комбинацию последовательных и параллельных соединений элементарных вентилей и, как таковая, сама определяет квантовый канал.

Пример 1. *Эволюция открытой квантовой системы, взаимодействующей с окружением.* Пусть $\mathcal{H}$ — гильбертово пространство системы, $\mathcal{H}_0$ – пространство окружения. Эволюция составной системы является обратимой и описывается унитарным оператором U.

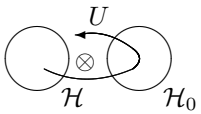


Рис. 6.1: Открытая квантовая система

В начальный момент система и окружение находятся в состоянии $S \otimes S_0$, затем взаимодействие "подкручивает" это состояние. Усредняя по окружению, получаем необратимую эволюцию самой системы

$$\Phi[S] = \mathrm{Tr}_{\mathcal{H}_0} U(S \otimes S_0)U^*. \tag{6.3}$$

Задача 49. Докажите, что отображение (6.3) является квантовым каналом. Указание:

$$\Phi^{(1)} \otimes \mathrm{Id}^{(2)}[S_{12}] = \mathrm{Tr}_{\mathcal{H}_0}(U^{(1)} \otimes I^{(2)})(S_{12} \otimes S_0)(U^{(1)} \otimes I^{(2)})^*.$$

Задача 50. Не все положительные отображения вполне положительны. Фиксируем базис в $\mathcal{H}$, тогда всякий оператор X задается матрицей $[x_{ij}]$. Рассмотрим преобразование транспонирования $\Phi[X] = X^\top$, которое совпадает с комплексным сопряжением на эрмитовых матрицах. Докажите, что условие 1) нарушается уже при $m = 2$.

Вспоминая, что физический смысл транспонирования – обращение времени (п. 1.11), мы видим, что нарушение условия полной положительности в данном случае можно интерпретировать так, что в одной системе происходит обращение времени, тогда как в другой — нет, таким образом получается нефизическое преобразование составной системы.

Рассмотрим два частных случая, позволяющих установить связь определения квантового канала с теми "полуклассическими" каналами, которые рассматривались в ч. II.

Определение. *Канал* Φ *называется классически-квантовым* (c-q), *если*

$$\Phi[S] = \sum_j S_j \langle e_j|S|e_j\rangle, \tag{6.4}$$

где S_j *— состояния в* $\mathcal{B}(\mathcal{H}')$, e_j *— о.н.б. в* $\mathcal{H}$.

Если S — состояние на входе, то $\langle e_j|S|e_j\rangle = p_j$ – распределение вероятностей на наборе состояний S_j и $\Phi[S] = \sum\limits_j p_j S_j$. Если $p_j = \delta_{kj}$, то получим на выходе состояние S_k, а в общем случае — смесь. Такой канал переводит классическое состояние $[p_i]$ в квантовое состояние (можно рассматривать это также как отображение диагональных матриц в произвольные матрицы плотности. Фактически, именно такие каналы рассматривались в разделе 5.3, где для них была доказана теорема кодирования.

Определение. *Канал* Φ *называется квантово-классическим* (q-c), если

$$\Phi[S] = \sum_j (\mathrm{Tr}\, SM_j)|e_j\rangle\langle e_j| \tag{6.5}$$

где $\{e_j\}$ — о.н.б. в $\mathcal{B}(\mathcal{H}')$, $\{M_j\}$ — наблюдаемая в $\mathcal{B}(\mathcal{H})$. При произвольном входном квантовом состоянии получаем классический выход (диагональную матрицу плотности). Это определение устанавливает связь между каналами и квантовыми измерениями.

Задача 51. Докажите полную положительность отображений (6.4), (6.5).

6.2 Квантовые каналы и открытые системы

Свойство полной положительности было введено Стайнспрингом (в более общем контексте операторных алгебр), который доказал теорему, обобщающую теорему Наймарка. В нашем случае теорема Стайнспринга сводится к следующему:

Теорема 22 (Представление Крауса). *Линейное отображение* $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ *является вполне положительным и сохраняющим след тогда и только тогда, когда*

$$\Phi[S] = \sum_{j=1}^{D} V_j S V_j^*, \tag{6.6}$$

где $D \le dd'$, *а* V_j *операторы из* $\mathcal{H}$ *в* $\mathcal{H}'$, *такие что*

$$\sum_{j=1}^{D} V_j^* V_j = I. \tag{6.7}$$

Доказательство. Рассмотрим "максимально сцепленный" вектор

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

в пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2$, где $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. В силу того, что отображение Φ вполне положительно, оператор

$$S_{12} = \left(\Phi \otimes \mathrm{Id}^2\right)\left[|\psi_{12}\rangle\langle\psi_{12}|\right] = \sum_{j,k=1}^{d} \Phi[|e_j^1\rangle\langle e_k^1|] \otimes |e_j^2\rangle\langle e_k^2|$$

является положительным в $\mathcal{H}' \otimes \mathcal{H}$. Используя его спектральное разложение, получаем

$$S_{12} = \sum_{l=1}^{D} |\Psi_l\rangle\langle\Psi_l|,$$

где $D \le dd'$, а векторы $|\Psi_l\rangle \in \mathcal{H}' \otimes \mathcal{H}$ являются (ненулевыми) собственными векторами этого оператора.

Определим линейные операторы $V_l : \mathcal{H} \to \mathcal{H}'$ соотношением

$$\langle\psi'|V_l|e_j^1\rangle = \langle\psi' \otimes e_j^2|\Psi_l\rangle, \quad \psi' \in \mathcal{H}'.$$

Тогда, используя матричное представление оператора в $\mathcal{H}$,

$$X = \sum_{j,k=1}^{d} |e_j^1\rangle\langle e_j^1|X|e_k^1\rangle\langle e_k^1|,$$

получаем

$$\langle\psi'|\Phi[X]|\psi'\rangle = \sum_{j,k=1}^{d} \langle e_j^1|X|e_k^1\rangle\langle\psi'|\Phi[|e_j^1\rangle\langle e_k^1|]|\psi'\rangle = \sum_{j,k=1}^{d} \langle e_j^1|X|e_k^1\rangle\langle\psi'\otimes e_j^2|S_{12}|\psi'\otimes$$

$$= \sum_{j,k=1}^{d} \langle e_j^1|X|e_k^1\rangle \sum_{l=1}^{D}\langle\psi'\otimes e_j^2|\Psi_l\rangle\langle\Psi_l|\psi'\otimes e_k^2\rangle = \sum_{j,k=1}^{d} \langle e_j^1|X|e_k^1\rangle \sum_{l=1}^{D}\langle\psi'|V_l|e_j^1\rangle\langle e_k^1|V_l^*|$$

то есть

$$\Phi[X] = \sum_{l=1}^{D} V_l X V_l^*.$$

Используя сохранение следа, получаем

$$\mathrm{Tr}X = \mathrm{Tr}\Phi[X] = \mathrm{Tr}X \sum_{l=1}^{D} V_l^* V_l$$

для любого оператора X, откуда следует (6.7). □

Следствие. *Всякий канал Φ, действующий в алгебре $\mathcal{B}(\mathcal{H})$, можно продолжить до унитарной эволюции (6.3) составной системы в $\mathcal{H}\otimes\mathcal{H}_0$, где вторая система ("окружение") находится в чистом состоянии $|\psi_0\rangle\langle\psi_0|$.*

Доказательство. Введем пространство окружения $\mathcal{H}_0 = \mathbb{C}^D$ и рассмотрим тензорное произведение

$$\mathcal{K} = \mathcal{H}\otimes\mathcal{H}_0 \approx \underbrace{\mathcal{H}\oplus\ldots\oplus\mathcal{H}}_{D},$$

описывающее систему+окружение. Введем оператор $V : \mathcal{H}\to\mathcal{K}$, действующий по правилу

$$V = \begin{bmatrix} V_1 \\ \vdots \\ V_D \end{bmatrix}.$$

Из (6.7) следует, что $V^*V = I$, т.е. V – изометрический оператор. Но это равносильно тому, что столбцы матрицы V образуют ортонормированную

систему. Дополняя эту систему до ортонормированного базиса в пространстве $\mathcal{K}$ (что всегда возможно)), получаем унитарный оператор, действующий в этом пространстве:

$$U = \begin{bmatrix} V_1 & \dots & \dots \\ \vdots & \vdots & \vdots \\ V_D & \dots & \dots \end{bmatrix}.$$

Рассмотрим единичный вектор в $\mathcal{H}_0 = \mathbb{C}^D$:

$$|\psi_0\rangle = \begin{bmatrix} 1 \\ \mathbf{0} \\ 0 \end{bmatrix},$$

где $\mathbf{0}$ обозначает вектор с нулевыми компонентами, тогда начальное состояние окружения

$$|\psi_0\rangle\langle\psi_0| = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ \mathbf{0} & \ddots & \mathbf{0} \\ 0 & \mathbf{0} & 0 \end{bmatrix},$$

так что начальное состояние системы+окружения есть

$$S \otimes |\psi_0\rangle\langle\psi_0| = \begin{bmatrix} S & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & 0 \end{bmatrix}.$$

Оператор унитарной эволюции преобразует его в

$$U\left(S \otimes |\psi_0\rangle\langle\psi_0|\right)U^* = \begin{bmatrix} V_1 S V_1^* & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & V_D S V_D^* \end{bmatrix}.$$

Беря частичный след по окружению, получаем

$$\mathrm{Tr}_0 U\left(S \otimes |\psi_0\rangle\langle\psi_0|\right)U^* = \sum_{l=1}^{D} V_l X V_l^* = \Phi[X].$$

□

Пример 3. *Деполяризующий канал.* Пусть $\mathcal{H}' = \mathcal{H}$, $\dim \mathcal{H} = d$. Рассмотрим отображение

$$\Phi[\mathcal{S}] = (1-p)\mathcal{S} + p\,\frac{I}{d}\,\mathrm{Tr}\,\mathcal{S}. \tag{6.8}$$

По определению, оно представляет собой смесь идеального канала Id и полностью деполяризующего канала, который отображает любое состояние в хаотическое $\overline{S} = I/d$.

Задача 52. Докажите полную положительность и найдите представление Крауса для отображения Φ.

Задача 53. Покажите, что деполяризующий канал может быть охарактеризован свойством унитарной ковариантности: $\Phi[U\mathcal{S}U^*] = U\Phi[\mathcal{S}]U^*$ для произвольного унитарного оператора U в $\mathcal{H}$.

Квантовый канал Φ называется *бистохастическим*, если он переводит хаотическое состояние в хаотическое, т. е. единицу в единицу. Таким образом, как Φ, так и Φ^* являются каналами в обеих картинах. Деполяризующий канал является бистохастическим.

Отметим, что для классического канала (6.1) аналогичное свойство выполняется, если $\sum_x p(y|x) \equiv 1$, что совпадает с обычным определением бистохастичности.

6.3 Q-битные каналы

Всякий канал Φ в пространстве q-бита $\mathbb{C}^2$ определяет аффинное отображение в пространстве $\mathbb{R}^3$

$$\vec{a} \to T\vec{a} + \vec{b},$$

где T – некоторая вещественная 3×3-матрица, $\vec{b} = [b_\gamma]_{\gamma=x,y,z}$– вектор в $\mathbb{R}^3$, при котором единичный шар отображается в подмножество себя, и которое удовлетворяют некоторым дополнительным ограничениям, вытекающим из условия полной положительности. При этом имеем

$$\Phi[S(\vec{a})] = S(T\vec{a} + \vec{b}). \tag{6.9}$$

Образом шара Блоха при этом отображении является некий эллипсоид, лежащий внутри шара.

Можно доказать[1], что с помощью вращений в $\mathbb{R}^3$, которым соответствуют унитарные эволюции в $\mathbb{C}^2$, канал Φ можно привести к виду (6.9), в котором матрица T диагональна:

$$T = \mathrm{diag}[\lambda_\gamma]_{\gamma=x,y,z}.$$

Разумеется, полная положительность налагает нетривиальные ограничения на параметры λ_γ, b_γ[2]. Наиболее прозрачен случай, когда $\vec{b} = 0$, т. е. отображение представляет собой сжатие единичного шара вдоль осей x, y, z с коэффициентами $|\lambda_x|, |\lambda_y|, |\lambda_z|$ (сочетающееся с отражением в случае отрицательных коэффициентов). Это имеет место тогда и только тогда, когда канал Φ бистохастичен.

Задача 54. Покажите, что в этом случае

$$\Phi[S] = \sum_{\gamma=0,x,y,z} \mu_\gamma \sigma_\gamma S \sigma_\gamma, \tag{6.10}$$

[1] M. B. Ruskai, S. Szarek, E. Werner, "A characterization of completely-positive trace-preserving maps on $\mathcal{M}_2$," quant-ph/0005004.

[2] *ibid.*

где

$$\mu_0 = \frac{1}{4}\left(1+\lambda_x+\lambda_y+\lambda_z\right), \quad \mu_x = \frac{1}{4}\left(1+\lambda_x-\lambda_y-\lambda_z\right), \tag{6.11}$$
$$\mu_y = \frac{1}{4}\left(1-\lambda_x+\lambda_y-\lambda_z\right), \quad \mu_z = \frac{1}{4}\left(1-\lambda_x-\lambda_y+\lambda_z\right),$$

и что неотрицательность этих чисел необходима и достаточна для полной положительности отображения Ф.

Пример 1. *Канал с ошибкой.* Рассмотрим канал, в котором с вероятностью $1-p$ состояние q-бита передается без помех, а с вероятностью p происходит ошибка – "переворот бита", см. п. 1.11. Уравнение канала

$$\Phi[S] = (1-p)S + p\sigma_x S\sigma_x.$$

Из уравнений (6.11) находим $\lambda_x = 1, \lambda_y = \lambda_z = 1-2p$, что соответствует равномерному сжатию шара Блоха в $|1-2p|$ раз в плоскости yz. Аналогично, уравнение

$$\Phi[S] = (1-p)S + p\sigma_y S\sigma_y,$$

описывающее случайный "переворот фазы", соответствует сжатию шара Блоха в $|1-2p|$ раз в плоскости xz.

Пример 2. *Деполяризующий канал.* Имеет место тождество

$$\sigma_0 S\sigma_0 + \sigma_x S\sigma_x + \sigma_y S\sigma_y + \sigma_z S\sigma_z = (2\,\mathrm{Tr}\,S)I,$$

откуда следует, что в случае $d=2$ деполяризующий канал (6.8) допускает представление

$$\Phi[S] = \left(1-\frac{3p}{4}\right)S + \frac{p}{4}\left(\sigma_x S\sigma_x + \sigma_y S\sigma_y + \sigma_z S\sigma_z\right).$$

Из уравнений (6.11) находим $\lambda_x = \lambda_y = \lambda_z = 1-p$, что соответствует равномерному сжатию шара Блоха в $1-p$ раз.

Пример 3. *Канал с затуханием амплитуды.* Канал задается представлением Крауса

$$\Phi[S] = V_0 S V_0^* + V_1 S V_1^*,$$

где

$$V_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{bmatrix}, \quad V_1 = \begin{bmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{bmatrix}.$$

Такое уравнение возникает при описании случайного перехода q-бита с уровня $|1\rangle$ на уровень $|0\rangle$ (матрица V_1), сопровождающегося уменьшением амплитуды состояния $|1\rangle$ в $\sqrt{1-p}$ раз (матрица V_0) [7].

Преобразование матрицы плотности имеет вид (6.9), где

$$T = \begin{bmatrix} 1-p & 0 & 0 \\ 0 & \sqrt{1-p} & 0 \\ 0 & 0 & \sqrt{1-p} \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} p \\ 0 \\ 0 \end{bmatrix},$$

что соответствует сжатию шара Блоха и перемещению образовавшегося эллипсоида вдоль оси z к северному полюсу, при котором эллипсоид касается единичной сферы (в точке северного полюса).

6.4 ПРОПУСКНЫЕ СПОСОБНОСТИ КВАНТОВОГО КАНАЛА

6.4.1 Передача информации по квантовому каналу

В этом разделе мы дадим беглый обзор теории пропускных способностей квантовых каналов связи, которая является развитием классической шенноновской теории. Мы ограничимся формулировками основных теорем кодирования, доказательства которых выходят за рамки настоящего курса. При желании их можно найти, например, в монографии [13].

В теории информации центральную роль играет понятие канала связи и его пропускной способности, дающей предельную скорость безошибочной передачи. Математический подход придает этим понятиям универсальную значимость: например, память компьютера (классического или квантового) может рассматриваться как канал из прошлого в будущее, а пропускная способность дает количественное выражение для предельной емкости памяти при исправлении ошибок. Важность рассмотрения квантовых каналов связи обусловливается тем, что всякий физический канал в конечном счете является квантовым, и такой подход позволяет учесть фундаментальные квантово-механические закономерности. Существенно, что в квантовом случае понятие пропускной способности разветвляется, порождая целый спектр информационных характеристик канала, зависящих от вида передаваемой информации (квантовой или классической), а также от дополнительных ресурсов, используемых при передаче.

Теоремы кодирования дают явные выражения для пропускных способностей через энтропийные параметры канала. Одним из главных достижений квантовой теории информации является открытие целого набора важнейших энтропийных характеристик.

6.4.2 Классическая пропускная способность квантового канала

В общем случае как выход, так и вход канала являются квантовыми; такой канал представляет собой открытую квантовую систему, взаимодействующую с окружением, которое привносит помехи в передаваемое состояние. Рассмотрим (вообще говоря, необратимую) эволюцию открытой системы, взаимодействующей с окружением. Обозначим $\mathcal{H}$ гильбертово пространство системы, $\mathcal{H}_E$ – пространство окружения, и пусть S_E – начальное чистое состояние окружения. Предположим, что обратимая эволюция, описывающая взаимодействие системы с окружением, задается унитарным оператором U. Тогда эволюция системы дается формулой

$$\Phi[S] = \mathrm{Tr}_{\mathcal{H}_E} U\,(S \otimes S_E)\,U^*. \tag{6.12}$$

С точки зрение теории информации канал связи вполне определяется отображением $S \to \Phi[S]$, переводящим состояния на входе в состояния на

выходе. Отображение Φ дает сжатое статистическое описание результата взаимодействия системы на входе с ее окружением (шумом). Например, деполяризующий канал (6.8) описывает смесь идеального канала и полностью деполяризующего канала, который переводит любое состояние в хаотическое $\bar{S} = \frac{I}{d}$.

При передаче классической информации (т. е. сообщения $w = (x_1, \ldots, x_n)$) по квантовому каналу связи она записывается в квантовом состоянии посредством задания значений параметров прибора, приготавливающего состояние S_w. Приемник производит квантовое измерение над состоянием на выходе канала связи, результатом которого являются значения $w' = (y_1, \ldots, y_n)$. Процесс передачи классической информации описывается диаграммой

$$w \overset{\text{кодирование}}{\longrightarrow} S_w \overset{\text{канал}}{\longrightarrow} S'_w \overset{\text{декодирование}}{\longrightarrow} w' \tag{6.13}$$

Применение квантовой теоремы кодирования (теорема 20 в разделе 5.3) дает следующее выражение для классической пропускной способности канала Φ

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} C_\chi(\Phi^{\otimes n}), \tag{6.14}$$

где

$$C_\chi(\Phi) = \max_{p_i, S_i} \left\{ H\left(\sum_i p_i \Phi\left[S_i\right]\right) - \sum_i p_i H\left(\Phi\left[S_i\right]\right) \right\} \tag{6.15}$$

квантовый аналог шенноновской формулы (5.8). Если эта величина обладает свойством *аддитивности*, $C_\chi(\Phi^{\otimes n}) = nC_\chi(\Phi)$, то $C(\Phi) = \bar{C}(\Phi)$.

Аддитивность величины $C_\chi(\Phi)$ означает, что использование сцепленных кодовых состояний не позволяет увеличить количество передаваемой классической информации. Долгое время оставался открытым вопрос, существуют ли вообще каналы, не обладающие таким свойством аддитивности. Лишь недавно удалось показать[3], что такие каналы существуют, по крайней мере в очень высоких размерностях, хотя конструктивного примера до сих пор нет.

Для q-битных бистохастических каналов, а также для деполяризующего канала, доказательство аддитивности дано в работах[4].

З а д а ч а 55. Для деполяризующего канала (6.8)

$$C_\chi(\Phi) = \log d + \left(1 - p\frac{d-1}{d}\right) \log \left(1 - p\frac{d-1}{d}\right) + p\frac{d-1}{d} \log \frac{p}{d}. \tag{6.16}$$

Покажите, что эта величина достигается для ансамбля равновероятных чистых состояний, отвечающих ортонормированному базису в $\mathcal{H}$.

Вычислим величину $C_\chi(\Phi)$ для произвольного q-битного бистохастического канала.

[3] M. B. Hastings, 'A counterexample to additivity of minimum output entropy, Nature Physics, **5** 2009, 255 - 257.

[4] C. King, Additivity for unital qubit channels, J. Math. Phys., **43** (2002), 4641-4653. C. King, The capacity of the quantum depolarizing channel, IEEE Trans. Inform. Theory, **49** (2003), 221-229.

Лемма 8. *Пусть* Φ *бистохастический канал в* $\mathbb{C}^2$, *тогда*

$$C_\chi(\Phi) = 1 - \min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S)). \tag{6.17}$$

Доказательство. Неравенство $\leq$ в (6.17) вытекает из того, что для любого канала

$$C_\chi(\Phi) \leq \log \dim\mathcal{H} - \min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S)), \tag{6.18}$$

так что остается доказать неравенство $\geq$. Поскольку энтропия – вогнутая функция состояния, минимум достигается на чистом состоянии S. Беря равновероятно чистые состояния $S_0 = S, S_1 = I - S$, получаем

$$C_\chi(\Phi) \geq H(\frac{1}{2}\Phi(I)) - \frac{1}{2}[H(\Phi(S)) + H(\Phi(I-S))].$$

Поскольку канал бистохастический, правая часть равна $H(\frac{1}{2}I) - \frac{1}{2}[H(\Phi(S)) + H(I - \Phi(S))]$, а в силу двумерности пространства, это равно правой части в (6.17).

При вычислении $\min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S))$, в силу унитарной инвариантности энтропии достаточно рассмотреть случай Φ вида (6.10). Отметим также, что собственные значения оператора плотности (1.25) в $\mathcal{H}_2$ равны $(1 \pm |\vec{a}|)/2$, а значит, энтропия равна

$$H(S) = h\left(\frac{1-|\vec{a}|}{2}\right). \tag{6.19}$$

Поскольку единичный шар отображается каналом Λ в эллипсоид с полуосями $|\lambda_\gamma|, \gamma = x, y, z$, минимум энтропии достигается на конце самой длинной полуоси, соответствующем оператору плотности с собственными значениями $(1 \pm \max_\gamma |\lambda_\gamma|)/2$. Отсюда получаем

$$C_\chi(\Phi) = 1 - h\left(\frac{1-\max_\gamma|\lambda_\gamma|}{2}\right). \tag{6.20}$$

6.4.3 Выигрыш от сцепленности между входом и выходом

Неклассический феномен супераддитивности информации в квантовом канале связи имеет в своей основе сцепленные кодирования и декодирования. Еще более впечатляющий выигрыш приносит введение сцепленности между входом и выходом как дополнительного информационного ресурса. Классическая пропускная способность канала Φ может быть увеличена путем использования сцепленности между входом и выходом канала, при том, что наличие одной только сцепленности не позволяет передавать информацию. Здесь, как и в ряде других случаев, сцепленность играет роль "катализатора", выявляющего скрытые ресурсы квантовой системы. Если Φ – идеальный канал, т. е. канал без шума, то выигрыш в пропускной способности, доставляемый так называемым сверхплотным кодированием, двукратен. Чем

более канал отличается от идеального, тем выигрыш больше, и в пределе каналов с очень большим шумом может быть сколь угодно большим. Для *классической пропускной способности с использованием сцепленного состояния* имеется простая формула[5]

$$C_{ea}(\Phi) = \max_S I(S, \Phi), \tag{6.21}$$

где $I(S, \Phi)$ *квантовая взаимная информация* между передатчиком и приемником, задаваемая соотношением

$$I(S, \Phi) = \{H(S) + H(\Phi(S)) - H(S; \Phi)\}. \tag{6.22}$$

Здесь $H(S), H(\Phi[S])$, – энтропии, соответственно, входного и выходного состояний, а $H(S; \Phi)$ – так называемая *обменная энтропия*, равная приращению энтропии окружения. Поскольку начальное состояние окружения S_E предполагается чистым, то $H(S; \Phi) = H(S'_E)$, где S'_E – конечное состояние окружения

$$S'_E = \mathrm{Tr}_{\mathcal{H}} U(S \otimes S_E)U^*.$$

Например, для деполяризующего канала (6.8)

$$C_{ea}(\Phi) = \log d^2 + \left(1 - p\frac{d^2-1}{d^2}\right)\log\left(1 - p\frac{d^2-1}{d^2}\right) + p\frac{d^2-1}{d^2}\log\frac{p}{d^2}. \tag{6.23}$$

Сравнивая это с величиной $C(\Phi)$, которая дается формулой (6.16), мы видим, что выигрыш $\frac{C_{ea}(\Phi)}{C_\chi(\Phi)} > 1$, причем $\frac{C_{ea}(\Phi)}{C_\chi(\Phi)} \to d+1$ в пределе сильного шума $p \to 1$.

6.4.4 Квантовая пропускная способность

Преобразование квантового состояния $S \to \Phi[S]$ можно рассматривать как передачу квантовой информации. Теория предсказывает возможность нетривиального способа передачи, при котором носитель состояния физически не пересылается, а передается лишь некоторая классическая информация (*телепортация* квантового состояния, раздел 3.3.). При этом необходимым дополнительным ресурсом также является сцепленность между входом и выходом канала связи. Свести передачу произвольного квантового состояния только к передаче классической информации, не используя дополнительного квантового ресурса, невозможно: поскольку классическая информация копируема, это означало бы возможность копирования и квантовой информации.

Открытие *квантовых кодов, исправляющих ошибки* (раздел 3.5) поставило вопрос об асимптотически (при $n \to \infty$) безошибочной передаче квантовых сообщений каналом $\Phi^{\otimes n}$. При этом *квантовая пропускная способность* $Q(\Phi)$ определяется как максимальное количество передаваемой квантовой информации и связана с размерностью подпространства векторов

[5] C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal, Entanglement-assisted capacity and the reverse Shannon theorem, IEEE Trans. Inform. Theory; e-print quant-ph/0106052.

входного пространства ($\approx 2^{nQ(\Phi)}$), отвечающие которым состояния передаются асимптотически безошибочно. Для нее имеется выражение через *когерентную информацию* $I_c(S,\Phi) = H(\Phi(S)) - H(S;\Phi)$, а именно

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{S^{(n)}} I_c(S^{(n)}, \Phi^{\otimes n}). \tag{6.24}$$

Доказательство[6] основано на глубокой аналогии между квантовым каналом и классическим каналом с перехватом, причем в квантовом случае роль перехватчика информации играет окружение рассматриваемой системы. Величина $Q(\Phi)$ тесно связана с криптографическими характеристиками канала, такими как пропускная способность для секретной передачи классической информации и скорость распределения случайного ключа.

Аналитическое выражение для квантовой пропускной способности деполяризующего канала до сих пор неизвестно, хотя имеются достаточно близкие нижние и верхние оценки.

6.4.5 Многообразие пропускных способностей

В классической теории информации хорошо известно, что обратная связь не увеличивает пропускную способность канала и шенноновская пропускная способность остается основной характеристикой. В квантовом случае аналогичный факт установлен для $C_{ea}(\Phi)$, а относительно $Q(\Phi)$ известно следующее: квантовая пропускная способность не может быть увеличена с помощью дополнительного классического канала от входа к выходу, сколь ни велика была бы его пропускная способность. Однако она может увеличиться, если есть возможность передачи классической информации в обратном направлении. Такая передача позволяет создать максимальную сцепленность между входом и выходом, которая может быть использована для телепортации квантового состояния. Даже канал с нулевой квантовой пропускной способностью, дополненный классической обратной связью, может быть использован для передачи квантовой информации, см. [7].

Три пропускные способности (6.14), (6.21), (6.24) связаны соотношением $Q(\Phi) \le C(\Phi) \le C_{ea}(\Phi)$ и образуют основу для определения и изучения всего многообразия пропускных способностей квантового канала связи, которое возникает при применении дополнительных ресурсов, таких как обратная или прямая связь, коррелированность или сцепленность. Так, в квантовой теории информации изучаются классическая и квантовая пропускные способности с обратной связью (обозначаемые, соответственно, $C_{\leftarrow}, Q_{\leftarrow}$), а также классическая и квантовая пропускные способности с дополнительным независимым классическим двусторонним каналом (соответственно, $C_{\leftrightarrow}, Q_{\leftrightarrow}$). Для квантового канала имеет место следующая иерархия

$$\begin{array}{ccccccc} C_\chi & \le & C_{\leftarrow} & \le & C_{\leftrightarrow} & \le & C_{ea} \\ \text{VI} & & \text{VI} & & \text{VI} & & \text{VI} \\ Q & \le & Q_{\leftarrow} & \le & Q_{\leftrightarrow} & \le & Q_{ea} \end{array},$$

[6] I. Devetak, *The private classical information capacity and quantum information capacity of a quantum channel*, e-print no. quant-ph/0304127, 2003.

где $\leq$ следует понимать как “меньше или равно для всех каналов и строго меньше для некоторых”[7]. Известно также, что $C_{ea} = 2Q_{ea}$ и для ряда остальных пар возможны неравенства как в ту, так и в другую стороны.

[7] C. H. Bennett, I. Devetak, P. W. Shor and J. A. Smolin, *Inequalities and separations among assisted capacities of quantum channels*, arXiv:quant-ph/0406086 (2004).

Глава 7

Заключение. Другие направления

Дальнейшее развитие теории приводит к изучению квантовых *каналов с памятью* и *систем с многими пользователями*. Большой раздел квантовой теории информации посвящен исследованию систем с "непрерывными переменными", основанных на принципах квантовой оптики. Многие эксперименты по квантовой обработке информации реализованы именно в таких системах. Особенно важными здесь являются *гауссовские состояния*, включающие когерентные и сжатые состояния, реализуемые в лазерах и нелинейных квантовых оптических устройствах, и соответствующий класс преобразователей квантовой информации – *гауссовские каналы*. Для них получен ряд результатов, касающихся сцепленности состояний, пропускных способностей и других информационных характеристик[1].

В заключение перечислим ряд других направлений, некоторые из них были вкратце рассмотрены в настоящем курсе:

- Квантовая теория оценивания состояний. Томография квантовых состояний;
- Количественные характеристики сцепленности. Многочастичная сцепленность;
- Алгоритмы сжатия квантовой информации;
- Квантовые коды, исправляющие ошибки. Вычисления, устойчивые к ошибкам;
- Быстрые квантовые алгоритмы и квантовое моделирование. Сложность квантовых вычислений;
- Квантовая криптография.

[1] C. Weedbroock, S. Pirandola, R. García-Patrón, N. J. Cerf, T. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian Quantum Information, Reviews of Modern Physics **84**, 621 (2012).

Последние 20 лет характеризуются нарастающим потоком публикаций в этой области[2]. Основным и наиболее оперативным источником научной информации является электронный архив Корнеллского университета (прежде – исследовательского центра в Лос-Аламосе): http://xxx.arxiv.org/quant-ph/. Появились специализированные журналы: Quantum Information Processi International Journal of Quantum Information; Quantum Information and Compu Квантовый компьютер и квантовые вычисления. Эта тематика представлена в таких известных журналах как Physical Reviews; Physical Reviews Letters; IEEE Transactions on Information Theory; Communications on Mathema Physics; Journal of Mathematical Physics; Проблемы передачи информации, появилась и монографическая литература, см. библиографию в [13].

[2]C. M. Caves, Quantum Information Science: Emerging No More, arXiv:1302.1864.

Литература

[1] К. А. Валиев, А. А. Кокин, Квантовые компьютеры: надежды и реальность, 2-е изд., М.: ИКИ, 2004.

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

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

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

[5] А. И. Кострикин, Ю. И. Манин, Линейная алгебра и геометрия, М.: Наука, 1986.

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

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

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

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

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

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

[12] А. С. Холево, Статистическая структура квантовой теории, М.-Ижевск: 2003.

[13] А. С. Холево, Квантовые системы, каналы, информация. М.: МЦНМО 2010.

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