Математический институт им. В. А. Стеклова
Российской академии наук

Лекционные курсы НОЦ

Выпуск 30

Издание выходит с 2006 года

Математические основы квантовой информатики

А. С. Холево

Математический институт им. В.А. Стеклова
Российской академии наук

МИАН

Москва
2018

ББК 22.311
Л43

Холево А. С.

Л43 Математические основы квантовой информатики – М.: МИАН, 2018. – 118 с. – (Лекционные курсы НОЦ, ISSN 2226-8782; Вып. 30).

ISBN 978-5-98419-080-7

Серия "Лекционные курсы НОЦ" – рецензируемое продолжающееся издание Математического института им. В.А. Стеклова Российской академии наук (МИАН). В серии "Лекционные курсы НОЦ" публикуются материалы специальных курсов, прочитанных в МИАН в рамках программы "Научно-образовательный центр МИАН".

DOI: https://doi.org/10.4213/lkn30

ISBN 978-5-98419-080-7

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 конечномерных 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" [9]), 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], [7], [9], рассматриваемую, однако, под углом ее сопоставления с классическим статистическим описанием. Последовательное развитие такой точки зрения привело во второй

[1] На самом деле, именно Фейнман сделал очень много для понимания квантовой теории: см., в частности, его обсуждение спиновых систем в книге [10].

половине XX века к обобщенной статистической модели квантовой теории [12], [13], 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 [14], [8]).

Автор благодарит Е. Р. Лубенец, замечания которой способствовали улучшению формулировок задач, и А. С. Кардашина за помощь в подготовке рисунков.

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 [7], [12]).

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:

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

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

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

Отметим еще, что индикатор E любого подмножества $\mathcal{E} \subseteq \Omega$ (т. е. "события" или "свойства") является случайной величиной, принимающей значения 0 или 1, что характеризуется алгебраическим свойством идемпотентности $E^2 = E$, а вероятность соответствующего события равна $\mathsf{M}_P E$.

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[1]

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

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 [7]:

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

Наблюдаемые, удовлетворяющие соотношению $E^2 = E$, являются матрицами проецирования на подпространства $\mathcal{E} \subseteq \mathbb{C}^d$, которые, таким образом, соответствуют квантовым "событиям" или "свойствам". При этом вероятность соответствующего события равна $\mathsf{M}_S E = \operatorname{Tr} SE$.

Если матрицы S, X, E 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 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: *эрмитовы матрицы X, Y коммутируют тогда и только тогда, когда они совместно диагонализуемы, т.е. существует ортонормированный базис из общих для них собственных векторов* (см. пп. (iii)–(iv) теоремы 4).

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.

[1] Неравенство $S \geqslant 0$ означает неотрицательность квадратичной формы матрицы S, см. раздел 1.3.

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 $\langle\phi|\psi\rangle$, $\phi, \psi \in \mathcal{H}$ [6]; 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)[2]. 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 $\{e_j\}_{j=1,\ldots,d}$ – ортонормированный базис (о.н.б.) в $\mathcal{H}$. Произвольный вектор $\psi \in \mathcal{H}$ может быть представлен в виде

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

что эквивалентно *соотношению полноты*

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

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

$$\langle\phi|\psi\rangle = \sum_{j=1}^{d} \langle\phi|e_j\rangle\langle e_j|\psi\rangle = \sum_{j=1}^{d} \overline{\langle e_j|\phi\rangle}\langle e_j|\psi\rangle$$

показывает, что отображение

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

является изометрическим изоморфизмом (взаимно-однозначным линейным отображением, сохраняющим скалярные произведения) гильбертовых пространств $\mathcal{H}$ и $\mathbb{C}^d$. В пространстве $\mathbb{C}^d$ выделяется *стандартный* о.н.б., для которого компоненты вектора $|e_j\rangle$ равны δ_{jk}, $k = 1, \ldots, d$.

Задача 1. Воспользовавшись соотношением полноты для данного базиса, запишите матричное представление для операторов в $\mathcal{H}$, аналогичное представлению векторов (1.2). Покажите, что матрица оператора A в о.н.б. $\{e_j\}_{j=1,\ldots,d}$ есть $[\langle e_j|A|e_k\rangle]_{j,k=1,\ldots,d}$.

[2]Нулевой вектор пространства $\mathcal{H}$ всегда обозначается 0, т.к. обозначения $|0\rangle$, $\langle 0|$ используются в другом смысле (см. раздел 1.7).

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 $\beta(\psi,\psi)$, $\psi\in\mathcal{H}$ (в вещественном случае подобное восстановление возможно лишь для симметричных форм). Благодаря этому, например, для доказательства операторного равенства $A=B$ достаточно установить равенство 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 $A=A^*$.

Задача 2. Докажите, что отображение $A\to A^*$ антилинейно и обладает свойствами $(A^*)^*=A$, $(AB)^*=B^*A^*$.

Задача 3. Докажите, что совокупность всех эрмитовых операторов в гильбертовом пространстве размерности d образует вещественное линейное пространство размерности d^2.

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

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

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 проектора P на область значений проектора 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 1. *Для любого эрмитова оператора A существует ортонормированный базис из собственных векторов $\{e_j\}$, которым отвечают вещественные собственные значения a_j.*

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

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

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

$$E_a = \sum_{j:a_j=a} |e_j\rangle\langle e_j|.$$

Набор различных собственных значений $\mathrm{spec}(A) = \{a\}$ называется *спектром* оператора A, а семейство проекторов $\{E_a\}$ – *спектральной мерой* 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*:

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

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

Для любой числовой функции f, определенной на множестве, содержащем все a_j, функция от оператора A определяется соотношением

$$f(A) = \sum_j f(a_j)|e_j\rangle\langle e_j| = \sum_{a\in\mathrm{spec}(A)} f(a)E_a. \tag{1.10}$$

Задача 4. Используя (1.9), покажите, что

$$\underbrace{A\times\cdots\times A}_{k} = \sum_{a\in\mathrm{spec}(A)} a^kE_a, \quad k = 2,\dots. \tag{1.11}$$

Поэтому для любого многочлена f оператор $f(A)$ может быть вычислен алгебраически, путем подстановки A в f вместо независимой переменной; с другой стороны, любую функцию на конечном множестве точек $\mathrm{spec}(A)$ можно заменить многочленом степени не выше d.

Как унитарные, так и эрмитовы операторы являются частными случаями *нормальных* операторов, то есть операторов, удовлетворяющих условию $A^*A = AA^*$. Для нормальных операторов спектральное разложение (1.7) имеет место с вообще говоря комплексными собственными значениями a. Соответственно обобщается и функциональное исчисление нормальных операторов.

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

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

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

A положителен тогда и только тогда, когда все главные (т.е. симметричные относительно главной диагонали) миноры неотрицательны.

Задача 5. 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

$$A = A_+ - A_-, \tag{1.12}$$

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

С другой стороны, всякий оператор A представляется комплексной линейной комбинацией двух эрмитовых операторов:

$$A = \frac{A + A^*}{2} + i\frac{A - A^*}{2i}.$$

В итоге получаем

Лемма 1. *Положительные операторы линейно порождают пространство всех операторов в гильбертовом пространстве* $\mathcal{H}$.

След оператора T определяется соотношением

$$\operatorname{Tr} T = \sum_{j=1}^{d} \langle e_j|T|e_j\rangle, \tag{1.13}$$

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

Задача 6. Покажите, что это определение не зависит от выбора базиса. Покажите, что функция $T \to \operatorname{Tr} T$ линейна и обладает свойствами:

$$\operatorname{Tr} A^* = \overline{\operatorname{Tr} A}, \tag{1.14}$$

$$\operatorname{Tr} AB = \operatorname{Tr} BA. \tag{1.15}$$

Покажите, что

$$\operatorname{Tr} A|\psi\rangle\langle\varphi| = \langle\varphi|A|\psi\rangle, \tag{1.16}$$

в частности

$$\operatorname{Tr} |\psi\rangle\langle\varphi| = \langle\varphi|\psi\rangle. \tag{1.17}$$

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

$$\operatorname{Tr} AB \geqslant 0 \tag{1.18}$$

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

1.4. Статистический постулат

Теперь мы обладаем необходимыми математическими средствами, чтобы сформулировать основной статистический постулат квантовой теории.

Всякая квантовая система характеризуется совокупностью статистических экспериментов, каждый из которых состоит из двух независимых стадий: приготовления (статистического)

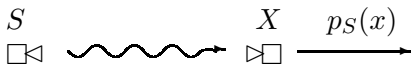

Рис. 1.1. Статистический эксперимент

состояния S и измерения наблюдаемой X. Квантовой системе сопоставляется гильбертово пространство $\mathcal{H}$ определенной размерности d. Состояния описываются *операторами плотности* S в $\mathcal{H}$, которые характеризуются условиями

$$S^* = S \geqslant 0, \quad \operatorname{Tr} S = 1. \tag{1.19}$$

(Вещественные) наблюдаемые описываются эрмитовыми операторами $X = X^*$, для которых имеет место спектральное разложение вида (1.8), т.е.

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

где x – различные собственные значения оператора X, а $\{E_x\}$ – его спектральная мера.

Результат статистического эксперимента, состоящего в приготовлении состояния S и измерении наблюдаемой X является случайной величиной, которая принимает значения $x \in \mathcal{X}$ с вероятностями

$$\mathsf{P}_S(X \models x) = \operatorname{Tr} S E_x. \tag{1.20}$$

Неотрицательность этих чисел следует из (1.18), а тот факт, что их сумма по всем x равна 1, из вторых равенств в (1.9) и (1.19). Таким образом, формула корректна, т.е. действительно задает распределение вероятностей на спектре наблюдаемой X. Согласно определению из теории вероятностей, математическое ожидание наблюдаемой X равно

$$\mathsf{M}_S X = \sum_{x \in \mathcal{X}} x \operatorname{Tr} S E_x = \operatorname{Tr} S X.$$

Формула (1.20) является наиболее полным выражением *статистического постулата* Борна–фон Неймана. Практически она означает, что эксперимент допускает неограниченное число независимых повторений, причем исходом каждой индивидуальной реализации эксперимента является то или иное собственное значение x оператора X, а частоты этих значений неограниченно приближаются к теоретическим величинам (1.20).

Заметим, что если в описанный выше картине мы ограничимся операторами, диагональными в фиксированном ортонормированном базисе, то мы фактически получим описание классической системы, которое имеет место, например в классической статистической механике.

В случае простого спектра наблюдаемой $X = \sum_j x_j |e_j\rangle\langle e_j|$ (все собственные x_j значения различны) из (1.20), (1.16) получаем

$$\mathsf{P}_S(X \models x_j) = \langle e_j | S | e_j \rangle, \tag{1.21}$$

а если состояние чистое (см. раздел 1.5), т. е. $S = |\psi\rangle\langle\psi|$, то

$$\mathsf{P}_S(X \models x_j) = |\langle e_j | \psi \rangle|^2.$$

Задача 7. Докажите формулу для дисперсии наблюдаемой X

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

Задача 8. Найдите математическое ожидание, дисперсию и распределение вероятностей наблюдаемой X в состоянии S, где

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

Множество всех квантовых состояний данной системы обладает статистически мотивированной структурой выпуклости, тогда как множество всех наблюдаемых имеет столь же мотивированную функциональную структуру. К изучению этих структур мы и переходим в следующих разделах.

1.5. 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 2 *(Минковский–Каратеодори). Пусть $\mathfrak{S}$ – компактное (ограниченное и замкнутое) 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 $S_j \in extr(\mathfrak{S})$:*

$$S = \sum_{j=1}^{n+1} p_j S_j, \quad S_j \in \mathrm{extr}(\mathfrak{S}). \tag{1.22}$$

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. Множество $\mathfrak{P}_{d-1}$ является множеством классических статистических состояний на конечном фазовом пространстве размера d. Его крайние точки называются *чистыми состояниями*.

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

$$\mathcal{F}\Big(\sum_j p_j S_j\Big) \leqslant (\geqslant) \sum_j p_j \mathcal{F}(S_j), \tag{1.23}$$

для любой выпуклой комбинации точек $S_j \in \mathfrak{S}$. В определении выпуклой функции обычно требуется выполнение этого условия для любых двух точек, а затем неравенство (1.23), называемое *неравенством Йенсена*, доказывается для любого конечного числа точек по индукции.

Функция является *аффинной*, если она как выпукла, так и вогнута, т.е.

$$\mathcal{F}\Big(\sum_j p_j S_j\Big) = \sum_j p_j \mathcal{F}(S_j). \tag{1.24}$$

Задача 9. Непрерывная выпуклая (в частности, аффинная) функция на компактном выпуклом множестве $\mathfrak{S}$ достигает своего максимума в крайней точке этого множества.

1.6. 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 $S \geqslant 0$, $\operatorname{Tr} S = 1$. Пусть $\mathcal{S}(\mathcal{H})$ – выпуклое множество всех операторов плотности. Выпуклая комбинация операторов плотности описывает смешивание соответствующих статистических ансамблей. Смесь $S = \sum_j p_j S_j$ получается, если взять ансамбли систем, приготовленных в состояниях S_j и смешать их с весами p_j.

В выпуклых множествах особо важны крайние точки, не представимые в виде нетривиальной смеси других точек. 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 3. *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

$$S = \sum_{j=1}^{d} s_j |e_j\rangle\langle e_j|, \quad s_j \geqslant 0, \qquad \sum s_j = 1, \tag{1.25}$$

где s_j – собственные значения, $|e_j\rangle$ – 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:

$$pS_1(I - S_1) + (1-p)S_2(I - S_2) + p(1-p)(S_1 - S_2)^2 = S - S^2 = 0. \tag{1.26}$$

Операторы $S_1(I-S_1), S_2(I-S_2)$ имеют неотрицательные собственные значения вида $s(1-s)$, $0 \leqslant s \leqslant 1$, и поэтому положительны. Кроме того, $(S_1 - S_2)^2 \geqslant 0$. Таким образом, 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. □

Для любого оператора плотности S *мера чистоты* определяется соотношением

$$P(S) = \operatorname{Tr} S^2 = \sum_j s_j^2, \tag{1.27}$$

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

$$H(S) = -\operatorname{Tr} S \log S = -\sum_j s_j \log s_j, \tag{1.28}$$

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 10. Покажите, что $1/d \leqslant P(S) \leqslant 1$, 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 $0 \leqslant H(S) \leqslant \log d$, причем $H(S) = 0$ тогда и только тогда, когда состояние S – чистое. $H(S) = \log d$ тогда и только тогда, когда состояние S – хаотическое.

Задача 11. Доказать, что если $\dim \mathcal{H} = d$, то $\mathcal{S}(\mathcal{H})$ погружается в вещественное пространство размерности $n = d^2 - 1$.

Спектральное разложение (1.25) показывает, что в случае множества квантовых состояний (как и для других выпуклых множеств с искривленной 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).

1.7. Двухуровневые системы. Квантовый бит

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

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

и множество всех классических состояний представляет собой единичный отрезок.

Наиболее простым, но важным примером квантовой системы является *q-бит* – двухуровневая квантовая система, $\dim \mathcal{H} = 2$. Будем использовать стандартный базис в $\mathbb{C}^2$:

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

Удобно ввести базис Паули в вещественном пространстве эрмитовых 2×2-матриц:

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

Задача 12. Матрицы Паули подчиняются следующей таблице умножения

$$\begin{gathered} \sigma_x^2 = I, \quad \sigma_y^2 = I, \quad \sigma_z^2 = I, \\ \sigma_x\sigma_y = i\sigma_z, \quad \sigma_y\sigma_z = i\sigma_x, \quad \sigma_z\sigma_x = i\sigma_y. \end{gathered} \tag{1.29}$$

Кроме того, $\operatorname{Tr}\sigma_x = \operatorname{Tr}\sigma_y = \operatorname{Tr}\sigma_z = 0$.

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

$$S(\vec{a}) = \frac{1}{2}\begin{bmatrix} 1+a_z & a_x - ia_y \\ a_x + ia_y & 1-a_z \end{bmatrix} = \frac{1}{2}(I + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z), \tag{1.30}$$

где $\vec{a} = (a_x, a_y, a_z)$. Условие $\det S \geqslant 0$ накладывает следующее ограничение на *параметры Стокса* (a_x, a_y, a_z):

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

Критерий Сильвестра показывает, что это условие необходимо и достаточно для того, чтобы $S \geqslant 0$. Таким образом, $\mathcal{S}(\mathcal{H})$ как выпуклое множество изоморфно единичному шару в $\mathbb{R}^3$.

Для меры чистоты (1.27) получаем

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

откуда следует, что чистые состояния соответствуют поверхности шара $|\vec{a}| = 1$ и составляют *сферу Блоха*.

Пусть $S(\vec{a}) = |\psi\rangle\langle\psi|$, где $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ – единичный вектор, так что $|c_0|^2 + |c_1|^2 = 1$. Тогда соответствующий вектор $\vec{a} = (a_x, a_y, a_z)$ на сфере Блоха однозначно определяется соотношениями

$$a_x + ia_y = c_0\bar{c}_1, \quad a_z = 2|c_0|^2 - 1. \tag{1.31}$$

Это соответствие не является взаимно однозначным: умножение вектора $|\psi\rangle$ на произвольный комплексный множитель α, по модулю равный единице, приводит к тому же состоянию $S(\vec{a})$, и к тому же вектору $\vec{a}$.

Задача 13. Для данного вектора чистого состояния q-бита найдите соответствующий вектор $\vec{a}$ на сфере Блоха:

$$|0\rangle, \quad |1\rangle, \quad \frac{1}{\sqrt{2}}\left(|0\rangle \pm |1\rangle\right), \quad \frac{1}{\sqrt{2}}\left(|0\rangle \pm i|1\rangle\right).$$

Чтобы описать обратное соответствие, введем углы Эйлера θ и ϕ так, что $a_z = \cos\theta$ и $a_x + ia_y = \sin\theta e^{i\phi}$, тогда

$$S(\vec{a}) = |\vec{a}\rangle\langle\vec{a}|, \tag{1.32}$$

где

$$|\vec{a}\rangle = \begin{bmatrix} \cos(\theta/2) \\ \sin(\theta/2)\, e^{i\phi} \end{bmatrix}. \tag{1.33}$$

Задача 14. Найдите углы Эйлера вектора $-\vec{a}$ и покажите, что $\langle\vec{a}|-\vec{a}\rangle = 0$. Таким образом, для каждого направления $\vec{a}$, векторы $|\pm\vec{a}\rangle$ образуют ортонормированный базис в $\mathbb{C}^2$.

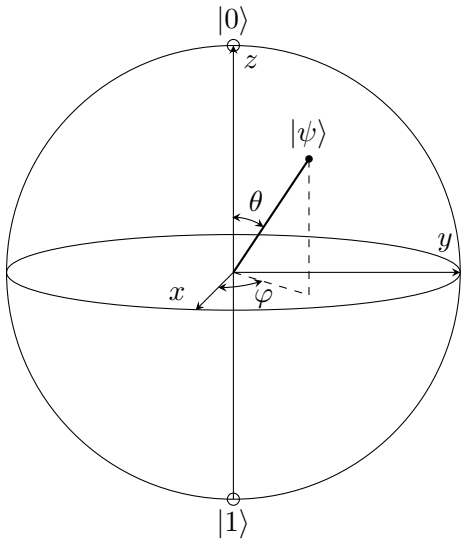

Рис. 1.2. Представление q-бита при помощи сферы Блоха

Задача 15. Для данного вектора $\vec{a}$ на сфере Блоха найдите соответствующий вектор чистого состояния $|\vec{a}\rangle$ и матрицу плотности в стандартном базисе:

$$\vec{a} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), \quad (0,1,0)\,.$$

Таким образом, мы дали описание состояний q-бита. Теперь перейдем к наблюдаемым. Используя тот факт, что матрицы Паули образуют базис в пространстве эрмитовых матриц, получаем, что всякая эрмитова 2×2- матрица однозначно представляется в виде

$$X = k_0 I + k_1 X(\vec{a}), \tag{1.34}$$

где $k_0 = \operatorname{Tr} X/2$,

$$X(\vec{a}) = \begin{bmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{bmatrix} = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z, \tag{1.35}$$

причем $|\vec{a}| = 1$. Таким образом, достаточно рассматривать наблюдаемые вида (1.35), т.к. остальные получаются из них путем изменения начала отсчета и масштаба. Для $|\vec{a}| = 1$ получаем из (1.30), (1.32)

$$|\vec{a}\rangle\langle\vec{a}| = \frac{1}{2}(I + X(\vec{a})). \tag{1.36}$$

Вычитая из (1.36) аналогичное равенство для $-\vec{a}$, получаем спектральное разложение эрмитова оператора (1.35)

$$X(\vec{a}) = |\vec{a}\rangle\langle\vec{a}| - |-\vec{a}\rangle\langle-\vec{a}|, \quad |\vec{a}| = 1. \tag{1.37}$$

Таким образом, наблюдаемая $X(\vec{a})$ имеет два собственных значения ± 1, которым отвечают собственные векторы $|\pm\vec{a}\rangle$ в $\mathbb{C}^2$.

Задача 16. Докажите формулу

$$X(\vec{a}_1)\,X(\vec{a}_2) = (\vec{a}_1 \cdot \vec{a}_2)I + i\,X(\vec{a}_1 \times \vec{a}_2). \tag{1.38}$$

Задача 17. Пользуясь формулой (1.38) и свойствами матриц Паули, покажите, что математическое ожидание наблюдаемой $X(\vec{b})$, $|\vec{b}| = 1$, в состоянии $S(\vec{a})$, $|\vec{a}| \leqslant 1$, равно

$$\operatorname{Tr} S(\vec{a})X(\vec{b}) = \vec{a} \cdot \vec{b}. \tag{1.39}$$

Задача 18. Пользуясь спектральным разложением (1.37), покажите, что измерение наблюдаемой $X(\vec{b})$, $|\vec{b}| = 1$ в состоянии $S(\vec{a})$, $|\vec{a}| \leqslant 1$, дает значения ± 1 с вероятностями

$$\mathsf{P}_{S(\vec{a})}(X(\vec{b}) \models \pm 1) = \frac{1}{2}\left(1 \pm \vec{a} \cdot \vec{b}\right). \tag{1.40}$$

Чему равна "вероятность перехода" $|\langle\vec{a}|\vec{b}\rangle|^2$, где $\|\vec{a}\| = \|\vec{b}\| = 1$?

Задача 19. Пользуясь соотношением (1.37), получите спектральное разложение оператора плотности (1.30):

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

Таким образом, собственные значения оператора плотности $S(\vec{a})$ суть $\frac{1\pm|\vec{a}|}{2}$. Отсюда, пользуясь определением (1.28), получаем энтропию состояния $S(\vec{a})$:

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

где $h(p) = -p\log p - (1-p)\log(1-p)$ – энтропия случайного бита. Эта функция постоянна на концентрических сферах $|\vec{a}| = r$, причем энтропия является вогнутой функцией, принимающей минимальное значение 0 на границе шара и максимальное 1 в его центре $\vec{a} = 0$. Этому соответствует *хаотическое* смешанное состояние $S(0) = I/2$.

Задача 20. Пусть $\vec{a}$ – случайный вектор, имеющий равномерное распределение на сфере Блоха. Найдите $\mathsf{M}S(\vec{a})$ и дайте статистическую интерпретацию ответа.

Физическими реализациями q-бита являются спин (внутренний угловой момент) электрона или атома со спином $1/2$, поляризация монохроматического фотона[3], любая квантовая система с двумя активными энергетическими уровнями, см. [1]. В таких системах наблюдаемая $X(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z$ описывает "проекцию спина" на направление $\vec{a} = (a_x\, a_y, a_z)$. Таким образом, *спин* – это вектор

$$\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$$

с (некоммутирующими) операторными компонентами $\sigma_x, \sigma_y, \sigma_z$. Состояние (1.36) описывает ансамбль (пучок частиц) со спином вдоль направления $\vec{a}$. *Эксперимент Штерна–Герлаха*, описывающий приготовление состояний $S(\vec{a})$ и измерение наблюдаемых $X(\vec{b})$ с помощью внешнего неоднородного магнитного поля, детально описан в лекциях Фейнмана [10].

1.8. Функции от наблюдаемой. Совместимые наблюдаемые

Рассмотрим вещественную наблюдаемую X со спектральным разложением

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

где x – различные собственные значения оператора X, а $\{E_x\}$ – его спектральная мера.

[3] В этом случае $\theta/2$ является углом линейной поляризации в плоскости, перпендикулярной направлению импульса, а ϕ – разностью фаз между горизонтальной и вертикальной компонентами электрического поля.

Для любой числовой функции $f(x)$, определенной на спектре $\mathcal{X}$, имеем

$$f(X) = \sum_{x\in\mathcal{X}} f(x)E_x.$$

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.

Вещественные наблюдаемые X, Y называются *совместимыми*, если найдется вещественная наблюдаемая Z, такая что $X = f(Z), Y = g(Z)$, где f, g – некоторые вещественные функции. Другими словами, наблюдаемые X, Y могут быть измерены в одном эксперименте (измерения Z), с последующим пересчетом результатов измерения.

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

Теорема 4. *Пусть X, Y –вещественные наблюдаемые со спектральными разложениями $X = \sum_x xE_x, Y = \sum_y yF_y$. Следующие утверждения эквивалентны*:

(*i*) *X, Y совместимы*;

(*ii*) *$[E_x, F_y] = 0$ для всех x, y*;

(*iii*) *$[X, Y] = 0$*;

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

Доказательство.

(*i*) $\Rightarrow$ (*ii*). Если X, Y совместимы, то

$$E_x = \sum_{z:f(z)=x} G_z, \quad F_y = \sum_{z:g(z)=y} G_z,$$

где $\{G_z\}$– спектральная мера наблюдаемой Z, откуда следует (*ii*), поскольку $[G_z, G_{z'}] \equiv 0$.

(*ii*) $\Rightarrow$ (*iii*). Очевидно, поскольку X является линейной комбинацией операторов E_x, а Y – линейной комбинацией операторов F_y.

(*iii*) $\Rightarrow$ (*iv*). Переходя к базису, в котором матрица X диагональна, мы можем считать, что $X = \operatorname{diag}[x_j]$, $Y = [y_{jk}]$. Из условия $XY - YX = 0$ получаем $(x_j - x_k)y_{jk} = 0$. Таким образом, $x_j \neq x_k$ влечет $y_{jk} = 0$. Группируя вместе одинаковые x_j, получаем, что матрицы X, Y можно представить в блочно-диагональном виде $X = \operatorname{diag}[x'_j I_j]$, $Y = \operatorname{diag}[Y_j]$, где все x'_j различны, I_j – единичные матрицы, размерности которых d_j равны кратности x'_j, а Y_j – эрмитовы $(d_j \times d_j)$-матрицы. Теперь в каждом блоке Y_j можно перейти к базису, в котором Y_j диагональна, при этом вид матрицы X не изменится.

(*iv*) $\Rightarrow$ (*i*). 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 $\{j\}$, причем $f(j) = x_j$, $g(j) = y_j$. $\square$

Таким образом, если наблюдаемые X, Y совместимы, то однозначно определены наблюдаемые вида

$$h(X, Y) = \sum_{x,y} h(x, y)E_x F_y,$$

где h – произвольная функция двух переменных, а

$$E_x F_y = F_y E_x = \sum_{z:f(z)=x, g(z)=y} G_z.$$

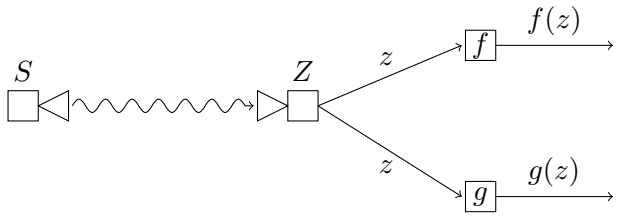

Рис. 1.3. Совместимые наблюдаемые

Отсюда видно, что *совместное распределение вероятностей* наблюдаемых X, Y в любом состоянии S может быть задано формулой

$$\mathsf{P}_S(X \models x, Y \models y) = \operatorname{Tr} SE_xF_y = \operatorname{Tr} SF_yE_x. \tag{1.41}$$

При этом

$$\mathsf{M}_S h(X,Y) = \operatorname{Tr} Sh(X,Y) = \sum_{x,y} h(x,y)\mathsf{P}_S(X = x, Y = y). \tag{1.42}$$

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

Задача 21. Обобщите соотношения (1.42), (1.41) на случай n попарно совместимых наблюдаемых.

Задача 22.

$$[\,X(\overrightarrow{n}), X(\overrightarrow{m})] = 2iX(\overrightarrow{n} \times \overrightarrow{m}).$$

Таким образом, наблюдаемые проекций спина на направления $\overrightarrow{n}, \overrightarrow{m}$ совместимы тогда и только тогда, когда эти направления коллинеарны. В частности, компоненты спина $\sigma_x, \sigma_y, \sigma_z$ не допускают совместного измерения.

Задача 23. Покажите, что вещественная наблюдаемая, совместимая со всеми квантовыми наблюдаемыми, является постоянной величиной, то есть задается оператором, кратным единичному оператору.

Задача 24. Докажите, что уравнение $XY - YX = cI$, где $c \neq 0$, не имеет решений в конечномерных матрицах X, Y.

Таким образом, наблюдаемые координаты и импульса Q, P, которые удовлетворяют каноническому коммутационному соотношению Гейзенберга

$$QP - PQ = i\hbar I, \tag{1.43}$$

(где $\hbar \approx 10^{-27}$ г см^{2}с$^{-1}$ – постоянная Планка), должны представляться "бесконечномерными матрицами", а точнее, операторами в бесконечномерном гильбертовом пространстве. Решение этого уравнения дают операторы $Q = x$, $P = \frac{\hbar}{i}\frac{d}{dx}$ в пространстве $\mathcal{H} = L^2(\mathbb{R})$ комплекснозначных функций на вещественной прямой $\mathbb{R}$ с интегрируемым квадратом модуля.

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 $X = [X_1, \dots, X_n]$ вещественных наблюдаемых положим $X_j^0 = X_j - I\mathsf{M}_S X_j$ и введем две вещественные матрицы: симметричную *матрицу ковариаций*

$$\mathsf{D}_S(X) = \left[\operatorname{Tr} SX_j^0 \circ X_k^0\right]_{j,k=1,\dots,n}, \tag{1.44}$$

и кососимметричную *коммутационную матрицу*

$$\mathsf{C}_S(X) = \left[i \operatorname{Tr} S[X_j, X_k]\right]_{j,k=1,\dots,n} = \left[i \operatorname{Tr} S[X_j^0, X_k^0]\right]_{j,k=1,\dots,n}. \tag{1.45}$$

Имеем

$$\mathsf{D}_S(X) \geqslant \frac{i}{2}\mathsf{C}_S(X) \tag{1.46}$$

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 $c_j \in \mathbb{C}$

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

где $Z = \sum_{j=1}^n c_j X_j$.

Для двух наблюдаемых $X_1 = X$ и $X_2 = Y$ неравенство (1.46) эквивалентно соотношению неопределенностей Шредингера–Робертсона

$$\mathsf{D}_S(X)\mathsf{D}_S(Y) \geqslant \operatorname{cov}_S(X,Y)^2 + \frac{1}{4}|\mathsf{M}_S[X,Y]|^2, \tag{1.47}$$

где

$$\mathsf{D}_S(X) = \operatorname{Tr} S(X - I\mathsf{M}_S(X))^2 \tag{1.48}$$

– дисперсия вещественной наблюдаемой X в состоянии S,

$$\operatorname{cov}_S(X,Y) = \mathsf{M}_S(X - I\mathsf{M}_S(X)) \circ (Y - I\mathsf{M}_S(Y)). \tag{1.49}$$

В самом деле, неравенство (1.46) принимает вид

$$\begin{bmatrix} \mathsf{D}_S(X) & \operatorname{cov}_S(X,Y) \\ \operatorname{cov}_S(X,Y) & \mathsf{D}_S(Y) \end{bmatrix} \geqslant \frac{i}{2}\begin{bmatrix} 0 & i\mathsf{M}_S[X,Y] \\ -i\mathsf{M}_S[X,Y] & 0 \end{bmatrix}.$$

Перенося правую часть налево, видим, что (1.47) означает неотрицательность определителя получившейся 2×2-матрицы. Остается использовать критерий Сильвестра неотрицательной определенности матрицы.

Если X, Y совместимые наблюдаемые, то величина $\operatorname{cov}_S(X,Y)$ представляют собой *ковариацию* величин X, Y в состоянии S; в этом случае $[X,Y] = 0$ и (1.47) 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. Отбрасывая неотрицательный первый член в правой части (1.47), получаем более известное неравенство

$$\mathsf{D}_S(X)\mathsf{D}_S(Y) \geqslant \frac{1}{4}|\mathsf{M}_S[X,Y]|^2. \tag{1.50}$$

Если формально подставить сюда наблюдаемые координаты и импульса $X = Q, Y = P$ и учесть (1.43), то получается соотношение неопределенностей Гейзенберга

$$\mathsf{D}_S(Q)\mathsf{D}_S(P) \geqslant \frac{\hbar^2}{4}, \tag{1.51}$$

из которого следует, что в любом квантовом состоянии неопределенности координаты и импульса связаны обратно пропорциональной зависимостью и не могуть быть сделаны одновременно сколь угодно малыми (строгое доказательство см., например, в [7]).

Задача 25. Докажите некоммутативное неравенство Коши–Буняковского

$$|\operatorname{Tr} SX^*Y|^2 \leqslant \operatorname{Tr} SX^*X \cdot \operatorname{Tr} SY^*Y, \tag{1.52}$$

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

Задача 26. Докажите неравенство

$$\mathsf{D}_{S(\overrightarrow{a})}X(\overrightarrow{n}) \cdot \mathsf{D}_{S(\overrightarrow{a})}X(\overrightarrow{m}) \geqslant (\overrightarrow{a}, \overrightarrow{n}, \overrightarrow{m})^2,$$

где в правой части – смешанное произведение трех векторов.

1.10. Апостериорное состояние. Последовательные измерения

Рассмотрим сначала классическую систему с фазовым пространством Ω, которая находится в (статистическом) состоянии $P = \{P(\omega)\}$. 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 $X(\omega) = x$:

$$P_x(\omega) = \frac{P(\omega)E_x(\omega)}{P\{X(\omega) = x\}}, \tag{1.53}$$

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

$$P(\omega) = \sum_x P_x(\omega)P\{X(\omega) = x\} \tag{1.54}$$

означает, что исходный "статистический ансамбль" разбивается на подансамбли, соответствующие различным значениям x случайной величины X, которые имеют "веса" $p_x = P\{X(\omega) = x\}$; 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, вероятности которых, согласно (1.21), равны $p_x = \langle e_x|S|e_x\rangle$. Постулируется, что в результате измерения с полученным исходом x система переходит в состояние

$$S_x = |e_x\rangle\langle e_x|. \tag{1.55}$$

Таким образом, статистический ансамбль после измерения разбивается на подансамбли, соответствующие различным исходам x с вероятностями p_x, и в целом описывается состоянием

$$S' = \sum_x p_x S_x = \sum_x |e_x\rangle\langle e_x|S|e_x\rangle\langle e_x|. \tag{1.56}$$

Постулат (1.55) мотивируется требованием *воспроизводимости*: при немедленном повторном измерении наблюдаемой X исход с вероятностью 1 должен совпасть с исходом x первого измерения.

В общем случае некоторые собственные числа X могут совпадать. Согласно *проекционному постулату фон Неймана - Людерса* [7], идеальное измерение наблюдаемой $X = \sum_x xE_x$ дает значение x с вероятностью

$$p_x = \operatorname{Tr} SE_x = \operatorname{Tr} E_x S E_x, \tag{1.57}$$

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

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

Поясним, что второе равенство в формуле (1.57) следует из того, что $E_x^2 = E_x$ и свойства следа (1.15). Для состояния всего ансамбля после измерения имеет место квантовый аналог формулы Байеса

$$S' = \sum_x p_x S_x = \sum_x E_x S E_x. \tag{1.58}$$

Отметим также, что если исходное состояние чистое, $S = |\psi\rangle\langle\psi|$, то апостериорное состояние также чистое, с вектором, который получается проецированием исходного вектора $|\psi\rangle$:

$$|\psi\rangle \longrightarrow |\psi_x\rangle = p_x^{-1/2} E_x |\psi\rangle, \quad p_x = \langle\psi|E_x|\psi\rangle. \tag{1.59}$$

Задача 27. Покажите, что если операторы S и X коммутируют, то проекционный постулат сводится к классической формуле (1.53), а формула (1.58) переходит в формулу Байеса (1.54).

Заметим, что в отличие от классической формулы Байеса (1.54), квантовое состояние (1.58) 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" (такое измерение называется неселективным). 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.

Феномен изменения квантового состояния при неселективном измерении лежит в основе протоколов квантовой криптографии, которые будут рассмотрены в разделе 3.6: если информация передается с помощью квантового состояния, то всякая попытка подслушивания требует измерения, которое с точки зрения легитимных участников является неселективным, поскольку они не могут знать его результатов. Однако сам факт подслушивания может быть установлен благодаря изменению состояния (1.58).

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

$$\begin{aligned} \mathsf{P}_S(X \models x, Y \models y) &= \mathsf{P}_S(Y \models y \,|X \models x)\, p_x = \operatorname{Tr} S_x F_y \cdot \operatorname{Tr} SE_x \\ &= \operatorname{Tr} E_x S E_x F_y = \operatorname{Tr} F_y E_x S E_x F_y. \end{aligned} \tag{1.60}$$

Заметим, что вообще говоря, $\mathsf{P}_S(Y \models y, X \models x) \neq \mathsf{P}_S(X \models x, Y \models y)$, 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 (1.60) совпадает с (1.41).

Задача 28. Идеальное квантовое измерение в общем случае удовлетворяет условию воспроизводимости: при повторном измерении наблюдаемой X исход с вероятностью 1 равен исходу первого измерения.

Обобщение соотношения (1.60) на случай последовательного измерения n наблюдаемых $X^1, \dots, X^n$ имеет вид

$$\mathsf{P}_S(X^1 \models x^1, \dots, X^n \models x^n) = \operatorname{Tr} E^n_{x^n} \dots E^1_{x^1} S E^1_{x^1} \dots E^n_{x^n}. \tag{1.61}$$

Задача 29. Покажите, что результаты последовательного измерения наблюдаемых с простым спектром образуют цепь Маркова.

Предположим, что над q-битом, приготовленном в хаотическом состоянии $S(0)$, производится измерение наблюдаемой $X(\vec{a})$. Согласно (1.40), в результате такого измерения получаются значения ± 1 с вероятностями $1/2$. При этом, согласно (1.55), получаются апостериорные состояния $S_{\pm 1} = |\pm\vec{a}\rangle\langle\pm\vec{a}|$. Предположим, что мы отбираем только те представители апостериорного ансамбля, для которых результат измерения равен $+1$. Такое измерение с последующим отбором в соответствии с полученным результатом (селективное измерение) является способом приготовления чистого состояния $|\vec{a}\rangle\langle\vec{a}|$, если первоначально имеется хаотический источник ("печка" в терминологии [10]).

Задача 30. Рассмотрите последовательное измерение наблюдаемых σ_x, σ_z в состоянии q-бита (1.30) и покажите, что

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

Задача 31. При каких значениях вещественных параметров a, b матрица

$$S = \frac{1}{3}\begin{bmatrix} 1 & a & 0 \\ a & 1 & ib \\ 0 & -ib & 1 \end{bmatrix}$$

является матрицей плотности? При каких a, b она является матрицей плотности чистого состояния?

Для состояния S найдите распределение вероятностей и апостериорные состояния при измерении наблюдаемой

$$X = \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{bmatrix},$$

где x_1, x_2, x_3 – произвольные вещественные числа.

1.11. Обратимые эволюции

Обратимые эволюции классической системы описываются взаимно-однозначными 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 $\Phi\colon \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$,

$$\Phi\Big[\sum_j p_j S_j\Big] = \sum_j p_j \Phi[S_j], \quad p_j \geqslant 0, \quad \sum_j p_j = 1, \quad S_j \in \mathcal{S}(\mathcal{H}).$$

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

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

Теорема 5. *Пусть* Φ – *аффинное взаимно-однозначное отображение выпуклого множества квантовых состояний* $\mathcal{S}(\mathcal{H})$ *на себя, тогда*

$$\Phi[S] = USU^*, \quad S \in \mathcal{S}(\mathcal{H}), \tag{1.62}$$

где U – *унитарный или антиунитарный оператор, определяемый этим соотношением с точностью до произвольного комплексного множителя, по модулю равного* 1.

Антиунитарный оператор U характеризуется свойствами:

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

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

$$S^\top = \bar{S} = \Lambda S \Lambda^*.$$

Доказательство теоремы 5 см., например, в [14].

Рассмотрим непрерывную однопараметрическую группу унитарных операторов $U_t; t \in \mathbb{R}$, т.е. семейство, удовлетворяющее условиям

1. $U_0 = I$;
2. $U_t U_s = U_{t+s}$;
3. функция $t \to U_t$ непрерывна.

Теорема Стоуна о непрерывных группах унитарных операторов утверждает, что

$$U_t = \mathrm{e}^{-itH}, \tag{1.63}$$

где H – эрмитов оператор. Параметр t обычно играет роль времени, а условие 2 означает однородность эволюции во времени, при этом H называется *гамильтонианом* или оператором энергии.

Теорема Стоуна является операторным аналогом известного в анализе факта, что непрерывные решения функционального уравнения $f(t+s) = f(t)f(s)$ суть экспоненциальные функции.

Задача 32. Покажите, что формула (1.63) задает непрерывную однопараметрическую группу унитарных операторов. Докажите (1.63) в предположении, что функция $t \to U_t$ непрерывно дифференцируема.

Дифференцируя соотношение (1.63) по t, получаем *уравнение Шредингера* для векторов чистых состояний $|\psi_t\rangle = U_t|\psi\rangle$

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

Из соотношений (1.62), (1.63) следует также, что всякая непрерывная однопараметрическая группа обратимых квантовых эволюций Φ_t, $t \in \mathbb{R}$, такая что $\Phi_0 = \mathrm{Id}$ (тождественное отображение), имеет вид

$$\Phi_t[S] = \mathrm{e}^{-itH} S \mathrm{e}^{itH}, \quad S \in \mathcal{S}(\mathcal{H}). \tag{1.64}$$

Дифференцируя соотношение (1.64) по t, получаем уравнение Лиувилля для оператора плотности $S_t = \Phi_t[S]$:

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

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.

Пример 2. *Обратимые эволюции q-бита.* Всякое аффинное взаимно-однозначное 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 R, причем в первом случае $\det R = 1$, а во втором $\det R = -1$. Пусть $S(\vec{a})$ – произвольное состояние q-бита (1.30). Тогда

$$S(R\vec{a}) = U(R) S(\vec{a}) U(R)^*, \tag{1.65}$$

где $U(R)$ – в первом случае унитарный, а во втором – антиунитарный оператор. Напомним, что $U(R)$ определяется этим соотношением с точностью до фазового множителя, по модулю равного 1. В случае унитарного оператора условимся устранить эту неоднозначность, потребовав $\det U(R) = 1$.

Если вектор $\vec{a}$ единичный, то $S(\vec{a})$ – чистое состояние (1.36), соответствующее направлению спина $\vec{a}$. При этом векторы чистых состояний в $\mathbb{C}^2$ преобразуются по формуле

$$|\vec{a}\rangle \to |R\vec{a}\rangle = \alpha U(R) |\vec{a}\rangle,$$

где α – некоторый множитель, по модулю равный 1.

Известная *теорема вращения* Эйлера утверждает, что всякое вращение в $\mathbb{R}^3$ можно реализовать как поворот вокруг некоторой оси.

Рассмотрим, например, поворот $R_{z,\varphi}$ вокруг оси z против часовой стрелки на угол φ. При этом единичный вектор $\vec{a}$ с углами Эйлера θ, ϕ преобразуется в вектор $R_{z,\varphi}\vec{a}$ с углами $\theta, \phi + \varphi$, поэтому вектор состояния (1.33) в $\mathcal{H}$ переходит в вектор

$$\begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} e^{i(\phi+\varphi)} \end{bmatrix} = e^{i\varphi/2} U(R_{z,\varphi}) |\vec{a}\rangle, \tag{1.66}$$

где

$$U(R_{z,\varphi}) = \begin{bmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{bmatrix} = \exp\left[-\frac{i\varphi}{2}\sigma_z\right]$$

– унитарная матрица с определителем, равным 1.

Задача 33. Покажите, что повороту $R_{\vec{a},\varphi}$ шара Блоха против часовой стрелки на угол φ вокруг оси $\vec{a}$ отвечает унитарный оператор

$$U(R_{\vec{a},\varphi}) = \exp\left[-\frac{i\varphi}{2} X(\vec{a})\right] = \cos\frac{\varphi}{2} I - i \sin\frac{\varphi}{2} X(\vec{a}), \tag{1.67}$$

где $X(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z$. Для доказательства второго соотношения используйте тождество $X(\vec{a})^2 = I$.

Задача 34. Унитарный оператор $U = \exp\left[-\frac{i\pi}{2}\sigma_\gamma\right] = -i\sigma_\gamma$ соответствует повороту R_γ в $\mathbb{R}^3$ на угол π относительно относительно координатной оси $\gamma = x, y, z$.

Отметим для будущего, что в квантовых вычислениях операция σ_x:

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

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.

Операция σ_x является квантовым аналогом классического отрицания NOT. Любопытно, что в квантовых вычислениях имеет смысл говорить об операции $\sqrt{\text{NOT}}$:

Задача 35. Покажите, что $\sigma_x = W^2$, где W – унитарный оператор

$$W = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}.$$

Задача 36. Покажите, что всякий унитарный оператор в пространстве векторов состояний q-бита имеет вид $U = \alpha U(R_{\vec{a},\varphi})$, где $|\alpha| = 1$. Найдите $\alpha, \vec{a}, \varphi$ для операции Адамара

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

Задача 37. Антиунитарный оператор $U = \Lambda$ комплексного сопряжения в базисе $|0\rangle, |1\rangle$ соответствует отражению R_{xz} относительно плоскости xz с $\det R_{xz} = -1$. При этом отображение $S(\vec{a}) \to S(R_{xz}\vec{a})$ сводится к транспонированию матрицы плотности $S(\vec{a})$.

Пример 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 (1.67), 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, и решение уравнения Шредингера имеет вид $|\psi_t\rangle = U(R_{z,\omega t})|\vec{a}\rangle$.

Изучением и решением уравнения Шредингера, а также связанными с этим вопросами занимается квантовая механика, см., например, [2], [10], [9]. 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.

1.12. Квантовый парадокс Зенона

Попытки рассмотрения непрерывных во времени измерений несовместимых наблюдаемых, опирающиеся на проекционный постулат, приводят к парадоксальным выводам, в основе которых лежит следующий математический факт. Пусть H – эрмитов оператор, а E – проектор в $\mathcal{H}$, так что $E^2 = E$. Тогда

$$\lim_{n\to\infty}\big(E\exp(itH/n)E\big)^n = \lim_{n\to\infty}\big[E + itEHE/n + o(1/n)\big]^n = E\exp(itEHE). \tag{1.68}$$

Рассмотрим квантовую частицу, эволюционирующую с гамильтонианом H на временном интервале $[0,t]$, и предположим, что в каждый момент времени tk/n, $k = 0,1,\dots,n$, производится измерение наблюдаемой E, т.е. проверка того факта, что частица находится в подпространстве $\mathcal{E}$, на которое проецирует E. Вероятность того, что во всех $n+1$ измерениях получен исход 1, в силу соотношения (1.61) равна

$$\mathsf{P}_{S_0}(1,\dots,1) = \mathrm{Tr}\big((E\exp(itH/n)E\big)^n S_0\big((E\exp(-itH/n)E\big)^n \tag{1.69}$$

и при $n\to\infty$ в силу (1.68) стремится к

$$\mathrm{Tr}\,E\exp(itEHE)S_0\exp(-itEHE)E = \mathrm{Tr}\,ES_0E = \mathrm{Tr}\,S_0E,$$

где S_0 – начальное состояние. Равенство получается, если воспользоваться свойством следа (1.15) и тем фактом, что проектор E коммутирует с любой функцией от EHE.

Если $ES_0E = S_0$, т.е. в начальный момент частица находится в подпространстве $\mathcal{E}$, то вероятность (1.69) равна 1 независимо от эволюции, т.е. при непрерывном измерении свойства E частица никогда не покидает соответствующее подпространство $\mathcal{E}$. Такое необычное поведение получило название “квантовый парадокс Зенона”.

Причина парадокса состоит в том, что измерение, описываемое проекционным постулатом, переводя состояние системы в состояние, отвечающее точно определенному значению наблюдаемой, производит конечное изменение, на фоне которого эффект эволюции за время t/n является пренебрежимо малым при $n\to\infty$. Чтобы избежать этого, необходимо ввести обобщение проекционного постулата, позволяющее рассматривать неточные измерения. При этом нетривиальный предельный процесс непрерывного измерения, включающий эволюцию, получается для последовательности неточных измерений, точность которых убывает пропорционально корню квадратному из числа измерений n (см. [13]).

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* (*кронекеровским*) *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 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"! Своеобразие и необычные возможности квантовой теории информации в значительной мере обусловлены свойствами составных квантовых систем.

Для более детального рассмотрения нам понадобится соответствующий математический аппарат.

2.2. Тензорное произведение гильбертовых пространств

Пусть $\mathcal{H}_i$ $(i = 1, 2)$ – гильбертовы пространства двух квантовых систем со скалярными произведениями $\langle\cdot|\cdot\rangle_i$. Совокупность двух квантовых систем 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] Англ. "entanglement", в российской физической литературе переводится как "запутанность".

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, оно и называется *тензорным произведением* $\mathcal{H}_1 \otimes \mathcal{H}_2$ гильбертовых пространств.

Задача 38. 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 $\dim \mathcal{H} = \dim \mathcal{H}_1 \cdot \dim \mathcal{H}_2$.

В частности, $\mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} = \mathbb{C}^{d_1 d_2}$ и стандартный базис в $\mathbb{C}^{d_1 d_2}$ состоит из всевозможных попарных произведений векторов стандартных базисов в $\mathbb{C}^{d_1}$ и $\mathbb{C}^{d_2}$.

Всякий вектор $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ записывается в виде

$$|\psi\rangle = \sum_{j=1}^{d_1} \sum_{k=1}^{d_2} c_{jk} |e_1^j\rangle \otimes |e_k^2\rangle = \sum_{j=1}^{d_1} |e_1^j\rangle \otimes |\psi_j\rangle, \tag{2.3}$$

или

$$|\psi\rangle = \begin{bmatrix} |\psi_1\rangle \\ \dots \\ |\psi_{d_1}\rangle \end{bmatrix}, \tag{2.4}$$

где компоненты $|\psi_j\rangle = \sum_{k=1}^{d_2} c_{jk} |e_k^2\rangle \in \mathcal{H}_2$, так что в общем случае $\mathcal{H}_1 \otimes \mathcal{H}_2$ изоморфно прямой ортогональной сумме d_1 слагаемых $\mathcal{H}_2 \oplus \dots \oplus \mathcal{H}_2$. Аналогичное, изоморфное представление получается, если поменять ролями $\mathcal{H}_1$ и $\mathcal{H}_2$.

Задача 39. Предполагая, что в $\mathcal{H}_1, \mathcal{H}_2$ используются стандартные базисы, запишите представление (2.4) для $|\psi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle$, где

$$|\phi_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad |\phi_2\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ \sqrt{2} \\ 0 \end{bmatrix}.$$

Для операторов X_1, X_2, действующих в пространствах $\mathcal{H}_1, \mathcal{H}_2$ соответственно, тензорное произведение задается соотношением

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

и продолжается по линейности на все пространство $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. В представлении (2.4) тензорного произведения $\mathcal{H}_1 \otimes \mathcal{H}_2$ этот оператор задается блочной $d_1 \times d_1$-матрицей $\left[\langle e_1^j | X_1 | e_1^l \rangle X_2 \right]$. Если при этом X_2 задается своей матрицей в базисе $\{e_2^k\}$, то $X_1 \otimes X_2$ сводится к кронекеровскому произведению матриц X_1, X_2.

В представлении (2.4) произвольный оператор X_{12} в $\mathcal{H}_1 \otimes \mathcal{H}_2$ действует как блочная матрица $X_{12} = [X_{jl}]_{j,l=1,\dots,d_1}$, элементы которой X_{jl} являются операторами в $\mathcal{H}_2$.

Задача 40. Найдите $(\sigma_x \otimes \sigma_y) |\psi\rangle$, где $|\psi\rangle = \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle)$.

Задача 41. Запишите матричный вид операторов $\sigma_x \otimes \sigma_z, \sigma_z \otimes \sigma_x, \sigma_y \otimes \sigma_x, \sigma_y \otimes \sigma_y$, действующих в представлении (2.4).

Если X_1, X_2 эрмитовы операторы со спектральным разложением

$$X_i = \sum_j x_i^j |e_i^j\rangle \langle e_i^j|, \quad i = 1, 2,$$

то $X_1 \otimes X_2$ является эрмитовым оператором со спектральным разложением

$$X_1 \otimes X_2 = \sum_j \sum_k x_1^j x_2^k |e_1^j \otimes e_2^k\rangle\langle e_1^j \otimes e_2^k|.$$

Отсюда следует, что если $X_1 \geqslant 0, X_2 \geqslant 0$, то $X_1 \otimes X_2 \geqslant 0$. В частности, если S_j, $j = 1, 2$, – 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_{12} действует в $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. *Частичный след* оператора T (по $\mathcal{H}_2$) обозначим $T_1 = \mathrm{Tr}_{\mathcal{H}_2} T_{12}$; это оператор в $\mathcal{H}_1$, ассоциированный с формой

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

Аналогично определяется частичный след $T_2 = \mathrm{Tr}_{\mathcal{H}_1} T_{12}$ по $\mathcal{H}_1$. Если $T = T_1 \otimes T_2$, то $\mathrm{Tr}_{\mathcal{H}_2}(T_1 \otimes T_2) = (\mathrm{Tr}\, T_2)T_1$. В представлении (2.4) частичные следы оператора $T = [T_{jl}]_{j,l=1,\ldots,d_1}$ находятся по формулам

$$\mathrm{Tr}_{\mathcal{H}_1} T = \sum_{j=1}^{d_1} T_{jj}, \quad \mathrm{Tr}_{\mathcal{H}_2} T = [\mathrm{Tr}\ T_{jl}]_{j,l=1,\ldots,d_1}. \tag{2.5}$$

Если S_{12} – оператор плотности квантового состояния, то его частичные следы S_1, S_2 являются операторами плотности. Они определяют *частичные состояния* подсистем $1, 2$.

Задача 42. Определение частичного следа не зависит от выбора ортонормированного базиса $\{e_2^k\}$.

Задача 43. Рассмотрим вектор чистого состояния $|\psi_{12}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Найдите матрицы плотности S_{12}, S_1, S_2.

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 $|\psi_{12}\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$, который является нетривиальной суперпозицией векторов-произведений, порождает чистое сцепленное состояние. Примером является *максимально сцепленное состояние*, которое порождается вектором

$$|\psi_{12}\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |e_j^1\rangle \otimes |e_j^2\rangle \tag{2.6}$$

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 $\mathcal{H}_{1,2}$.

Задача 44. Для максимально сцепленного состояния 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 6 *[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.*

[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.

Тогда S_1 и S_2 имеют одни и те же ненулевые собственные значения λ_j. Более того,

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

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

Доказательство теоремы. Пусть $\{e_j^1\}$ – ортонормированный базис в $\mathcal{H}_1$ из собственных векторов оператора S_1, тогда согласно (2.3) имеет место разложение

$$|\psi\rangle = \sum_j |e_j^1\rangle \otimes |h_j^2\rangle, \tag{2.8}$$

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

$$\sum_{j,k} \langle h_j^2|h_k^2\rangle |e_k^1\rangle\langle e_j^1| = S_1 = \sum_j \lambda_j |e_j^1\rangle\langle e_j^1|, \tag{2.9}$$

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 7 *[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 (2.7) с произвольным базисом $\{e_j^2\}$ в гильбертовом пространстве $\mathcal{H}_2$, изоморфном $\mathcal{H}_1$. Любой другой вектор $|\psi'\rangle$ имеет разложение (2.7) с другим базисом в $\mathcal{H}_2$. Остается заметить, что любые два базиса в гильбертовом пространстве связаны унитарным преобразованием. □

Из теоремы 6 вытекает равенство ненулевых собственных значений. Используя выражение для энтропии (1.28), получаем

Следствие. Энтропии частичных состояний S_1 и S_2 равны между собой:

$$H(S_1) = H(S_2) = -\sum_j \lambda_j \log \lambda_j. \tag{2.10}$$

Энтропия частичных состояний является мерой сцепленности чистого состояния $S_{12} = |\psi\rangle\langle\psi|$ составной системы: она принимает минимальное значение 0 тогда и только тогда, когда S_{12} несцеплено, а максимальное значение $\log d$ тогда и только тогда, когда S_{12} максимально сцеплено.

Смешанное состояние S_{AB} называется *разделимым* или несцепленным, если оно является смесью состояний-произведений, а все состояния, не сводящиеся к таковым – *сцепленными*. Сцепленность представляет собой чисто квантовое свойство, лишь отчасти родственное классической коррелированности.

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 состояния $\mathrm{Tr}_B\, S_{AB}$, тогда как для смешанных состояний имеется целый ряд существенно различных характеристик, важнейшей из которых является *сцепленность формирования*

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

где минимум берется по всевозможным ансамблям чистых состояний S_{ψ_i}, представляющим состояние S_{AB}:

$$S_{AB} = \sum_i p_i S_{\psi_i}.$$

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 из сцепленных измерений над 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 5.3).

2.4. Два q-бита

Ортонормированный базис в пространстве двух q-битов образован четырьмя векторами-произведениями

$$|00\rangle = |0\rangle_1 \otimes |0\rangle_2, \quad |01\rangle = |0\rangle_1 \otimes |1\rangle_2, \quad |10\rangle = |1\rangle_1 \otimes |0\rangle_2, \quad |11\rangle = |1\rangle_1 \otimes |1\rangle_2. \tag{2.11}$$

Следуя принятой в квантовой информатике терминологии мы будем называть этот базис, как и его очевидное обобщение на случай n q-битов, *вычислительным*. 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

$$\begin{aligned} S_1 &= \mathrm{Tr}_2\, |\psi\rangle\langle\psi| = \\ &= \left(|a|^2 + |b|^2\right)|0\rangle\langle 0| + (a\bar{c} + b\bar{d})|0\rangle\langle 1| + (\bar{a}c + \bar{b}d)|1\rangle\langle 0| + \left(|c|^2 + |d|^2\right)|1\rangle\langle 1|. \end{aligned}$$

Собственные числа матрицы плотности находятся из характеристического уравнения

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

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

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

Энтропия частичного состояния (2.10) является строго монотонно возрастающей функцией C.

Величина C называется "согласованностью" (concurrence) и также может служить 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 $\lambda_\pm = \frac{1}{2}$.

Задача 45. Является ли вектор $|\psi\rangle = \frac{1}{2}\left(|00\rangle - |01\rangle - |10\rangle + |11\rangle\right)$ сцепленным?

Задача 46. Проверьте, что четыре максимально сцепленных вектора

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \tag{2.12}$$

$$|\beta_{01}\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right), \tag{2.13}$$

$$|\beta_{10}\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right), \tag{2.14}$$

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right) \tag{2.15}$$

образуют ортонормированный базис в системе из двух q-битов. Этот базис называется *базисом Белла*.

Задача 47. Покажите, что три наблюдаемые $\sigma_\gamma \otimes \sigma_\gamma$, $\gamma = x, y, z$, являются совместимыми. Все векторы базиса Белла являются собственными векторами для каждой из этих трех наблюдаемых.

2.5. Парадокс ЭПР. Неравенство Белла

Квантовая сцепленность отражает необычные корреляции составных квантовых систем, которые описываются тензорным (а не декартовым, как в классической механике) произведением подсистем. Обычно сцепленность возникает в результате квантового взаимодействия подсистем.

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 результирующее состояние их спинов, называемое *синглетным*, описывается вектором

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}\big[\,|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle\big],$$

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. В краткой записи

$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}\big[\,|01\rangle - |10\rangle\big].$$

Каждая из компонент описывает состояние с разнонаправленными спинами, а $|\beta_{11}\rangle$ – их суперпозиция, которую невозможно представить в виде произведения векторов состояний, относящихся к разным частицам. Синглетное состояние – канонический пример чистого 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 (1.35) проекции спина $X(\vec{a})$ для одной частицы и $X(\vec{b})$ для другой. Операторы $X = X(\vec{a}) \otimes I$, $Y = I \otimes X(\vec{b})$ коммутируют, следовательно соответствующие наблюдаемые совместимы и их ковариация дается выражением (1.49).

Задача 48. Используя выражения для матричных элементов,

$$\langle 0|X(\vec{a})|0\rangle = a_z, \quad \langle 1|X(\vec{a})|0\rangle = a_x + ia_y, \quad \langle 1|X(\vec{a})|1\rangle = -a_z, \tag{2.16}$$

вытекающие из (1.35), покажите, что в синглетном состоянии среднее значение и дисперсия каждой наблюдаемой равны

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

так что $X(\vec{a}), X(\vec{b})$ принимают значения ± 1 с равными вероятностями $1/2$, а ковариация между спинами задается формулой

$$\langle \beta_{11}|X(\vec{a}) \otimes X(\vec{b})|\beta_{11}\rangle = -\vec{a} \cdot \vec{b}. \tag{2.17}$$

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 (2.16) и соотношения $X(\vec{a})^2 = I$ следует

$$\langle \beta_{11}|[X(\vec{a}) \otimes I + I \otimes X(\vec{a})]^2|\beta_{11}\rangle = 0,$$

откуда

$$\big[X(\vec{a}) \otimes I + I \otimes X(\vec{a})\big]|\beta_{11}\rangle = 0,$$

или

$$\big[X(\vec{a}) \otimes X(\vec{a})\big]|\beta_{11}\rangle = -|\beta_{11}\rangle.$$

Это означает, что если A и B измеряют проекцию спина в одном и том же направлении $\vec{a}$, то результаты их измерений будут случайны, но с вероятностью 1 противоположны, каково бы ни было направление $\vec{a}$.

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

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" системы.

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 (2.17) 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* $|X_j| \leqslant 1$, $|Y_k| \leqslant 1$. *Тогда для любого распределения вероятностей на* Ω *корреляции этих величин удовлетворяют неравенству*

$$|\mathsf{M}X_1Y_1 + \mathsf{M}X_1Y_2 + \mathsf{M}X_2Y_1 - \mathsf{M}X_2Y_2| \leqslant 2, \tag{2.18}$$

где M *– соответствующее математическое ожидание.*

Доказательство получается усреднением элементарного неравенства

$$-2 \leqslant X_1Y_1 + X_1Y_2 + X_2Y_1 - X_2Y_2 \leqslant 2,$$

которое в свою очередь вытекает из цепочки соотношений

$$\begin{aligned} |X_1Y_1 + X_1Y_2 + X_2Y_1 - X_2Y_2| &\leqslant |Y_1 + Y_2| + |Y_1 - Y_2| \\ &= \sqrt{2(Y_1^2 + Y_2^2 + |Y_1^2 - Y_2^2|)} \\ &= \sqrt{4\max(Y_1^2, Y_2^2)} \leqslant 2. \end{aligned}$$

Принцип локальности, или, лучше сказать, *разделимости* в данной модели заключается в том, что физическая наблюдаемая для второй системы описывается одной и той же случайной величиной (Y_1 в случае первой и третьей корреляций, Y_2 в других двух случаях) независимо от того, какая величина – X_1 или X_2 измеряется во первой системе. Это условие кажется настолько естественным, что оно даже трудно уловимо. Однако именно оно запрещает мгновенное влияние выбора измерения, проводящегося в одной системе, на представление измерений в другой, пространственно удаленной, системе. Если от него отказаться, то интересующие нас четыре физические корреляции могут быть любыми величинами из отрезка $[-1,1]$ и граница в неравенстве (2.18) увеличивается до 4.

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

При этом система приготавливается в одном и том же синглетном состоянии. В каждом из 4-х экспериментов подсчитываются корреляции между результатами измерений в первом и втором q-битах.

Предположим, что для этих экспериментов существует классическое вероятностное описание, удовлетворяющее сформулированному выше принципу локальности (разделимости). Математически это означает, что найдется вероятностное пространство (Ω, P) и случайные величины X_1, X_2, Y_1, Y_2 на этом пространстве, принимающие значения ± 1 (как и спиновые переменные $\widehat{X}_1, \widehat{X}_2, \widehat{Y}_1, \widehat{Y}_2$), и такие что

$$\mathsf{M}_P X_j Y_k = \mathsf{M}_S \widehat{X}_j \widehat{Y}_k = -\vec{a}_j \cdot \vec{b}_k \quad j,k = 1,2.$$

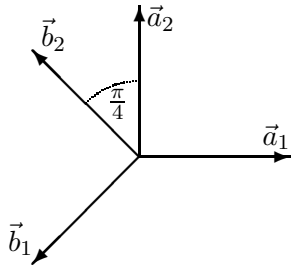

Рис. 2.1. Выбор векторов $\vec{a}_j$ и $\vec{b}_k$

Подстановка таких значений корреляций (из формулы (2.17)) в левую часть формулы (2.18) дает значение $2\sqrt{2}$, нарушающее неравенство. Отсюда следует, что либо квантовая теория дает неправильные выражения для реальных 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 теории[3].

Еще более необычными свойствами в плане сцепленности обладает состояние Гринбергера–Хорна–Цайлингера (GHZ), которое задается следующим вектором в пространстве трех q-битов

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right). \tag{2.19}$$

Рассмотрим четыре квантовые наблюдаемые

$$\begin{aligned}
\widehat{A}_1 &= \sigma_y \otimes \sigma_y \otimes \sigma_x,\\
\widehat{A}_2 &= \sigma_y \otimes \sigma_x \otimes \sigma_y,\\
\widehat{A}_3 &= \sigma_x \otimes \sigma_y \otimes \sigma_y,\\
\widehat{B} &= \sigma_x \otimes \sigma_x \otimes \sigma_x
\end{aligned}$$

Задача 49. Покажите, что

$$\widehat{A}_j|\psi\rangle = -|\psi\rangle, \quad j = 1,2,3, \qquad \widehat{B}|\psi\rangle = |\psi\rangle. \tag{2.20}$$

Предположим, что существует классическое описание спиновых переменных в виде функций $X_j(\omega), Y_j(\omega)$; $j = 1,2,3$, на гипотетическом фазовом пространстве Ω, принимающих значения ± 1, так что $X_j^2 = Y_j^2 = 1$. Рассмотрим классический аналог наблюдаемых $\widehat{A}_j$:

$$\begin{aligned}
A_1 &= Y_1 Y_2 X_3,\\
A_2 &= Y_1 X_2 Y_3,\\
A_3 &= X_1 Y_2 Y_3,\\
B &= X_1 X_2 X_3
\end{aligned}$$

Пусть $\omega \in \Omega$ – любая фазовая точка (классическое чистое состояние), для которой

$$A_j(\omega) = -1, \quad j = 1,2,3.$$

[3]Впечатляющий эксперимент описан в недавней работе R. Hanson et al., Experimental loophole-free violation of a Bell inequality using entangled electron spins separated by 1.3 km, arXiv:1508.05949.

Поскольку тождественно $B = A_1A_2A_3$, имеем $B(\omega) = -1$, в противоположность квантовым соотношениям (2.20).

Задача 50. Докажите формулу для частичного следа GHZ-состояния

$$\mathrm{Tr}_3\ |\psi\rangle\langle\psi| = \frac{1}{2}\big(|00\rangle\langle 00| + |11\rangle\langle 11|\big).$$

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. Крупье наугад 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, так что произведение равно $= 1$, B – ± 1 в своем столбце, так что произведение равно $= -1$. 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

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

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 (соответственно, B) произведение всех матричных элементов должно равняться 1 (соответственно, -1). Можно доказать, что наилучшая стратегия состоит в выборе матрицы, в которой лишь один элемент является разным для A и B. В этом случае они выигрывают в 8 из 9 случаев. Таким образом, вероятность успеха всегда будет < 1.

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

Рассмотрим *состояние Белла* $S = |\beta_{00}\rangle_{AB}\langle\beta_{00}|$,

$$|\beta_{00}\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B\right) = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right). \tag{2.21}$$

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

$$|\beta_{0000}\rangle = |\beta_{00}\rangle_{A_1B_1} \otimes |\beta_{00}\rangle_{A_2B_2}.$$

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 $X_{ij}^2 = \sigma_0 \otimes \sigma_0 \equiv I$, так что 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

$$\left(X_{ij}^A \otimes X_{ij}^B\right) |\beta_{0000}\rangle = |\beta_{0000}\rangle, \quad i, j = 1, 2, 3,$$

где X_{ij}^A (соответственно, X_{ij}^B) – оператор X_{ij} в системе $A = A_1A_2$ (соответственно, $B = B_1B_2$). Это равенство говорит, что если вся система $A_1A_2B_1B_2$ приготовлена в состоянии $|\beta_{0000}\rangle$, то произведение результатов измерений A и B в любой клетке ij будет равно 1, т.е. результаты совпадут.

Задача 51. Докажите утверждения 1-3.

Квантовая стратегия удовлетворяет всем правилам игры. Именно использование квантовых информационных протоколов 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 (2.18) 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

$$|\mathsf{M}_S X_1Y_1 + \mathsf{M}_S X_1Y_2 + \mathsf{M}_S X_2Y_1 - \mathsf{M}_S X_2Y_2| \leqslant 2\sqrt{2}. \tag{2.22}$$

Доказательство вытекает из тождества

$$(X_1Y_1 + X_1Y_2 + X_2Y_1 - X_2Y_2)^2 = 4I - [X_1, X_2][Y_1, Y_2], \tag{2.23}$$

с учетом того, что $\|[X_1, X_2]\| \leqslant 2$, так что норма выражения (2.23) не превосходит 8.

Задача 52. Докажите тождество (2.23).

Для системы из двух q-битов AB равенство в (2.22) 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 (2.21).

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 (2.18))[4].

[4]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. (5.30)):

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 (т.е. квантовой сцепленности) 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,

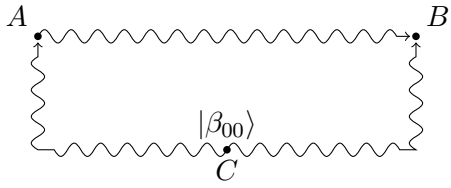

Рис. 3.1. Сверхплотное кодирование

описываемое о.н.б. $|e_y\rangle$ в $\mathcal{H}_B$. Легко видеть, что результирующая переходная вероятность

$$P(y|x) = \mathrm{Tr}(U_x \otimes I_B) S_{AB} (U_x \otimes I_B)^* (I_A \otimes |e_y\rangle\langle e_y|) = \langle e_y| \, S_B \, |e_y\rangle \tag{3.2}$$

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 этой точки зрения, также представляет 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 (2.12)–(2.15) в системе из двух q-битов AB может быть получен из одного вектора $|\beta_{00}\rangle$ действием "локальных" унитарных операторов, т.е. операторов, действующих нетривиально только в пространстве q-бита A, а именно

$$|\beta_{10}\rangle = (\sigma_z \otimes I)|\beta_{00}\rangle, \quad |\beta_{01}\rangle = (\sigma_x \otimes I)|\beta_{00}\rangle, \quad |\beta_{11}\rangle = i(\sigma_y \otimes I)|\beta_{00}\rangle.$$

Таким образом, если AB изначально находятся в *сцепленном* состоянии $|\beta_{00}\rangle$, участник A может закодировать 2 бита классической информации в 4 состояния базиса Белла, производя только локальные операции, а затем (физически) послать свой q-бит участнику B по идеальному квантовому каналу. Тогда, производя измерение в базисе Белла, участник B получает 2 бита классической информации. Дополнительным преимуществом такого протокола является его защищенность по отношению к возможному перехвату третьей стороной q-бита, посланного от A к B: полученное частичное состояние является хаотическим (задача 44), т.е. перехватчик не получает никакой информации.

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

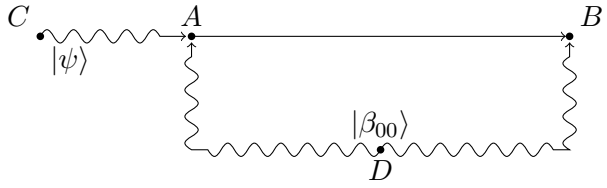

Рис. 3.2. Телепортация квантовых состояний

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 сцепленность между входом и выходом канала связи. Заметим, что свести передачу произвольного квантового состояния к только передаче классической информации, не используя дополнительного квантового ресурса, невозможно: поскольку классическая информация копируема, это означало бы возможность копирования и квантовой информации.

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

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

1. Перед началом передачи система AB приготовляется в максимально сцепленном состоянии с вектором
$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right).$$
2. C передает A произвольное чистое состояние
$$|\psi\rangle = a|0\rangle + b|1\rangle.$$

[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.

Совокупность трех систем CAB описывается вектором состояния

$$(a|0\rangle + b|1\rangle) \otimes \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}\big[a|000\rangle + b|100\rangle + a|011\rangle + b|111\rangle\big].$$

3. Затем
   (a) A производит некоторое обратимое преобразование состояния системы CA;
   (b) A производит измерение (с 4 исходами, что составляет 2 бита классической информации). Преобразование и измерение будут описаны ниже.
4. A посылает результат измерения B по классическому каналу связи.
5. В зависимости от полученного результата измерения B производит некоторое преобразование и получает это произвольное $|\psi\rangle$.

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

$$|00\rangle \to |00\rangle, \quad |01\rangle \to |01\rangle, \quad |10\rangle \to |11\rangle, \quad |11\rangle \to |10\rangle, \tag{3.3}$$

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 вычислительный базис переходит в себя, следовательно, в 4-х мерном пространстве CA этому преобразованию соответствует унитарный оператор (перестановка векторов базиса). 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

$$|0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle, \qquad |1\rangle \to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle, \tag{3.4}$$

т.е. базис поворачивается (с изменением ориентации) на угол $\pi/4$.

Вектор начального состояния всей системы CAB есть

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

После действия CNOT на CA получаем

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

Потом H действует на C:

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

Выделяя состояние системы CA, получаем вектор состояния системы CAB:

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

Теперь производится измерение в системе CA, проецирующее этот вектор на один из 4-х базисных векторов $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ согласно формуле (1.59), в которой $E_x = |x\rangle\langle x| \otimes I_B$, $p_x =$

1/4; $x = 00, 01, 10, 11$. Исход измерения x посылается от A к B по классическому (идеальному) каналу связи.

Результатом является одно из четырех равновероятных апостериорных состояний системы CAB с вектором

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

В зависимости от полученного значения $x = 00, 01, 10, 11$ участник B применяет к своему состоянию соответствующий унитарный оператор Паули

$$\sigma_{00} = I, \quad \sigma_{01} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_{10} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad -i\sigma_{11} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

во всех четырех случаях преобразующий вектор состояния системы B

$$a|0\rangle + b|1\rangle, \quad a|1\rangle + b|0\rangle, \quad a|0\rangle - b|1\rangle, \quad a|1\rangle - b|0\rangle$$

в вектор $|\psi\rangle = a|0\rangle + b|1\rangle$, который описывал состояние системы C до начала протокола. При этом противоречия с невозможностью копирования квантового состояния не возникает, так как частичное состояние C в конце протокола есть $|0\rangle\langle 0|$ или $|1\rangle\langle 1|$, в зависимости от исхода измерения. Если же исход не учитывается, то получается хаотическое состояние. Более подробно:

Задача 53. Матрица плотности системы CAB после выполнения протокола телепортации имеет вид

$$S_{CAB} = \frac{1}{2} I_C \otimes \frac{1}{2} I_A \otimes |\psi\rangle\langle\psi|.$$

Таким образом, разрушается как исходное состояние C, так и сцепленность между A и B.

Конструкции протоколов сверхплотного кодирования и телепортации допускают обобщение на случай пространства произвольной конечной размерности.

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].

Задача 54. Рассмотрим систему двух q-битов. Покажите, что применение операции Адамара к первому q-биту, за которым следует операция CNOT, переводит вычислительный базис в базис Белла. Последовательное применение этих операций в обратном порядке переводит базис Белла в вычислительный базис. Таким образом, измерение в базисе Белла (например, в протоколе сверхплотного кодирования) может быть сведено к измерению в вычислительном базисе.

При всей фундаментальной значимости протоколов сверхплотного кодирования и телепортации, демонстрирующих удивительные свойства квантовой сцепленности, с чисто утилитарной точки зрения могут быть более простые решения для рассматриваемых задач. В любом случае создание сцепленности предполагает возможность установления квантовых каналов от источника сцепленности к участникам A и B. Но тогда возможно более экономным решением было бы установление дополнительного прямого канала от A к B.

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

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

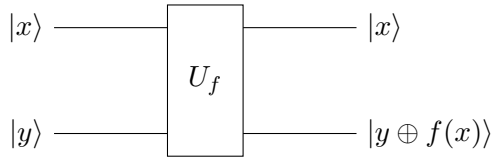

Рис. 3.3. Квантовый "оракул"

Ю. И. Манина "Вычислимое и невычислимое", 1980 г. Простейшие, но довольно искусственные примеры таких задач рассмотрели Дойч 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 Дойча. Обозначим $B = \{0, 1\}$. Пусть $f\colon\ B \to B$ – некоторая булева (двоичная) функция. Она либо постоянна: $f = \text{const}$, если $f(0) = f(1)$, либо непостоянна: $f \neq \text{const}$, если $f(0) \neq f(1)$. Очевидный классический алгоритм, позволяющий определить, какая из этих двух возможностей имеет место, предполагает вычисление обоих ее значений: $f(0)$ и $f(1)$.

Дойч предложил алгоритм, который позволяет решить *квантовый аналог* этой задачи, используя лишь один акт *квантового вычисления* функции f. Рассмотрим пространство двух q-битов, основного и вспомогательного, в котором задан унитарный оператор U_f, действующий по формуле

$$|x\rangle \otimes |y\rangle \xrightarrow{U_f} |x\rangle \otimes |y \oplus f(x)\rangle.$$

Если вспомогательный q-бит приготовлен в состоянии с вектором $|0\rangle$, то

$$|x\rangle \otimes |0\rangle \xrightarrow{U_f} |x\rangle \otimes |f(x)\rangle.$$

В этом смысле оператор U_f задает квантовое вычисление функции f. Отметим свойство оператора U_f, которое будет использовано в дальнейшем:

$$U_f\left(|x\rangle \otimes |-\rangle\right) = (-1)^{f(x)}\left(|x\rangle \otimes |-\rangle\right), \tag{3.5}$$

т.е. векторы $|x\rangle\otimes|-\rangle$, $x = 0, 1$, являются собственными векторами оператора U_f с собственными значениями $(-1)^{f(x)}$. Здесь $|-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)$.

Доказательство.

$$\begin{aligned} U_f\left(|x\rangle \otimes \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)\right) &= |x\rangle \otimes \frac{1}{\sqrt{2}}\big(|f(x)\rangle - |1 \oplus f(x)\rangle\big) \\ &= |x\rangle \otimes \begin{cases} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), & f(x) = 0 \\ \frac{1}{\sqrt{2}}(|1\rangle - |0\rangle), & f(x) = 1 \end{cases} \end{aligned}$$

Дадим пошаговое описание квантового алгоритма.

1. Основной q-бит приготавливается в состоянии с вектором $|0\rangle$, вспомогательный – в состоянии с вектором $|1\rangle$.
2. К обоим q-битам применяется операция Адамара

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

в результате чего состояния q-битов преобразуются по формулам

$$H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) = |+\rangle, \quad H|1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) = |-\rangle.$$

3. К полученному состоянию применяется оператор U_f. Используя свойство (3.5), получаем

$$\begin{aligned} U_f\left(H\otimes H\right)\left(|0\rangle\otimes|1\rangle\right) &= \frac{1}{2}U_f\left(|0\rangle + |1\rangle\right)\otimes\left(|0\rangle - |1\rangle\right) \\ &= \frac{1}{2}\left((-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\right)\otimes\left(|0\rangle - |1\rangle\right). \end{aligned}$$

4. К полученному состоянию применяется оператор $H\otimes I$:

$$\begin{aligned} &\left(H\otimes I\right)U_f\left(H\otimes H\right)\left(|0\rangle\otimes|1\rangle\right) \\ &\qquad = \frac{1}{2}\left(\left[(-1)^{f(0)} + (-1)^{f(1)}\right]|0\rangle + \left[(-1)^{f(0)} - (-1)^{f(1)}\right]|1\rangle\right)\otimes|-\rangle. \end{aligned}$$

5. Производя измерение над основным q-битом в вычислительном базисе $\{|0\rangle, |1\rangle\}$, с вероятностью 1 получаем исход 0, если $f(0) = f(1)$, или исход 1, если $f(0) \neq f(1)$.

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. Алгоритм Саймона. Обозначим $B^n = B^{\times n}$ множество всех двоичных последовательностей длины n. Множество B^n с операциями покомпонентного двоичного сложения и умножения на коэффициенты из B образует линейное пространство над полем B [6]. Пусть задано отображение $f\colon B^n \to B^n$. Считается известным, что оно является "периодическим", то есть $f(x) = f(y) \Leftrightarrow y = x\oplus\xi$, где $\xi \in B^n$ – двоичный (булев) вектор, $\xi \neq 0$. Здесь $\oplus$ обозначает покомпонентное сложение двоичных векторов по модулю 2 (логическая операция "XOR").

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. Для доказательства предположим, что все $2^n - 1$ возможных значений ξ равновероятны. После вычисления s значений функции $f(x_j)$, $j = 1,\ldots,s$, сравнивая эти значения во всевозможных парах точек, мы можем обнаружить, что $f(x_j) = f(x_k)$ и тогда найти $\xi = x_j \oplus x_k$. В худшем случае все значения $f(x_j)$ различны и тогда мы можем исключить $s(s-1)/2$ значений ξ. Вероятность того, что $(s+1)$-е вычисление даст нужный результат, не превосходит

$$\frac{s}{2^n - 1 - s(s-1)/2} \leqslant \frac{s}{2^n - s^2},$$

поскольку $f(x_{(s+1)})$ должно совпасть с одним из s значений $f(x_j)$, $j = 1,\ldots,s$. Вероятность того, что потребуется $m+1$ актов вычисления f для нахождения ξ, оценивается как

$$p \leqslant \sum_{s=1}^{m} \frac{s}{2^n - s^2} \leqslant \frac{m^2}{2^n - m^2},$$

откуда

$$m \geqslant \sqrt{p/(1-p)}\,2^{n/2}. \tag{3.6}$$

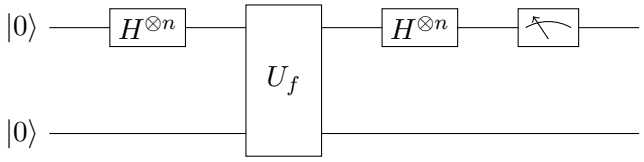

Рис. 3.4. Алгоритм Саймона

В теоретической информатике задачи, требующие для своего решения экспоненциальное число неких элементарных операций, принято считать "трудными", см. подробнее [5]. 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

$$H_n = \underbrace{H \otimes \cdots \otimes H}_{n} = H^{\otimes n}.$$

Рассмотрим *квантовый регистр* – физическую систему из n q-битов; информация будет задаваться состоянием этой системы. Если x пробегает B^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 $|0\rangle = |00\dots\rangle$, затем к каждому q-биту применяется операция Адамара:

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

   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 и некоторых других алгоритмах соответствующее вычисление описывается эффективно). В принципе, он может быть составлен из некоторых элементарных одно- и двух-кубитных операций, если известно, как

само отображение f составлено из элементарных логических операций. Здесь $|z\rangle$ – состояние *вспомогательного регистра*, который введен, чтобы сделать операцию вычисления функции обратимой. Если исходно этот регистр находится в основном состоянии $|00\dots\rangle$, то

$$\left(\sum_x |x\rangle\right) \otimes |00\dots\rangle \xrightarrow{U_f} \sum_x |x\rangle \otimes |f(x)\rangle.$$

3. Вновь применяя операцию Адамара, получаем вектор состояния

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

4. Поскольку $\xi \neq 0$, отображение f принимает 2^{n-1} разных значений (Каждое значение встречается два раза, как $f(x)$ и как $f(x\oplus\xi)$. Поэтому вектор состояния можно переписать в виде

$$\frac{1}{2^n} \sum_{y\in B^n} \sum_{f(x)\,\text{разные}} [(-1)^{x\cdot y} + (-1)^{(x+\xi)\cdot y}] |y\rangle \otimes |f(x)\rangle.$$

Коэффициент в суперпозиции равен $2(-1)^{x\cdot y}$, если $y\cdot\xi = 0 \,(\mathrm{mod}\, 2)$, и 0 в противном случае. Поэтому вектор состояния принимает вид

$$\frac{1}{2^{n-1}} \sum_{y\cdot\xi=0} \sum_{f(x)\,\text{разные}} (-1)^{x\cdot y} |y\rangle \otimes |f(x)\rangle.$$

Измеряя основной регистр в вычислительном базисе, получаем одно из 2^{n-1} различных значений y, удовлетворяющих уравнению $y\cdot\xi = 0$, каждое с вероятностью $2^{-(n-1)}$.
Таким образом, получается случайный равномерно распределенный булев 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 $(\mathrm{mod}\, 2)$, что позволяет найти вектор ξ.

Лемма 2. *Пусть $y_1(\omega), \dots, y_{n-1}(\omega)$ вероятностно независимые, равномерно распределенные случайные векторы из гиперплоскости $y\cdot\xi = 0$. Тогда*

$$\mathsf{P}\{y_1(\omega), \dots, y_{n-1}(\omega) \text{ линейно независимы}\} \geqslant \frac{1}{4}.$$

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

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

Тогда

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

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

Следовательно

$$\begin{aligned}
&\mathsf{P}\{y_1(\omega), \ldots, y_{n-1}(\omega) \text{ линейно независимы}\} \\
&\qquad = \left(1 - \frac{1}{2}\right) \ldots \left(1 - \frac{1}{2^{n-1}}\right) \\
&\qquad \geqslant \frac{1}{2} \cdot \left(1 - \frac{1}{4}\right) \left(1 - \frac{1}{8}\right) \ldots \\
&\qquad \geqslant \frac{1}{2} \cdot \left(1 - \frac{1}{4} - \frac{1}{8} - \ldots\right) = \frac{1}{4},
\end{aligned}$$

где многократно использовано неравенство $(1-a)(1-b) \geqslant 1-a-b$ при $a, b \geqslant 0$.

5) Повторяем всю процедуру m раз, где $(1-1/4)^m \leqslant \varepsilon$, так что $m \geqslant c \log \frac{1}{\varepsilon}$. Тогда с вероятностью $p = 1-\varepsilon$ получим по крайней мере $n-1$ линейно независимых булевых векторов, ортогональных ξ, а значит, и сам вектор ξ.

Квантовый алгоритм дает вероятностный ответ и требует лишь $O(n \log \frac{1}{\varepsilon})$ применений[3] оператора U_f вместо $O\left(2^{n/2}\sqrt{\frac{1}{\varepsilon}}\right)$ вычислений значения f для классического вероятностного алгоритма, как следует из (3.6). Вновь, такое радикальное ускорение достигается благодаря "квантовому параллелизму" при действии оператора U_f.

3.4.3. Замечания об алгоритме Шора. Алгоритм, предложенный Шором в 1994 г., эффективно решает задачу нахождения множителя большого составного 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, широко используемой, в частности, для шифрования данных в интернете. 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 не выше $O(n^3)$. Алгоритм использует сведение задачи факторизации к нахождению периода функции $f(x) = a^x \pmod N$, 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). Реализация алгоритма Шора поставила бы под угрозу методы шифрования данных с открытым ключом, что явилось причиной повышенного интереса к идее квантового компьютера. С другой стороны, были предложены и реализованы квантовые методы распределения секретного ключа (см. п. 3.6), использование которых позволяет принципиально достичь защищенности каналов связи от несанкционированного прослушивания даже с использованием квантового компьютера.

Подробнее об алгоритме Шора и квантовых вычислениях см. в [8, 4].

[3] Отметим, что алгоритм решения системы линейных уравнений n-го порядка имеет полиномиальную сложность $O(n^{2,376})$ (Coppersmith, Winograd).

3.4.4. Алгоритм Гровера. Этот алгоритм решает квантовый аналог задачи поиска. Более точно, предполагается, что задана булева функция $F\colon B^n \to B$, 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 $\sim \sqrt{N} = 2^{n/2}$ шагов, при этом решение носит вероятностный характер.

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

$$U_F|x\rangle = |x\rangle, \quad x \neq x_0, \qquad U_F|x_0\rangle = -|x_0\rangle. \tag{3.7}$$

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

$$H_n J H_n = 2H_n|0\rangle\langle 0|H_n - I = 2|\psi(\theta_0)\rangle\langle\psi(\theta_0)| - I. \tag{3.8}$$

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

Лемма 3. *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 $\varphi = 2\theta_0$ в плоскости, натянутой на базис $|\bar{x}_0\rangle$, $|x_0\rangle$ в направлении от $|\bar{x}_0\rangle$ к $|x_0\rangle$.

Доказательство. Используя (3.7), (3.8), получаем

$$U|x_0\rangle = -H_n J H_n|x_0\rangle = -2|\psi(\theta_0)\rangle\langle\psi(\theta_0)|x_0\rangle + |x_0\rangle$$

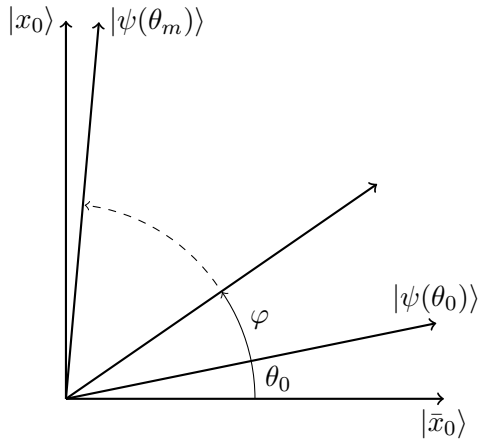

Рис. 3.5. Алгоритм Гровера

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

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

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

где

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

При этом конечное состояние $|\psi(\theta_m)\rangle$ становится очень близким к искомому:

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

причем тем ближе, чем больше N.

Если в завершение производится измерение в вычислительном базисе, то вероятность получить результат x_0 очень близка к единице, а именно

$$|\langle x_0|\psi(\theta_m)\rangle|^2 = \sin^2\theta_m = 1 - O(1/N).$$

Доказано, что этот алгоритм оптимален, в том смысле, что порядок $\sim \sqrt{N} = 2^{n/2}$ нельзя улучшить. Это, в частности, означает, что невозможно достичь экспоненциального ускорения в решении других трудных (NP-полных) задач, используя "квантовый перебор" возможных решений. К сожалению, данный алгоритм не позволяет достичь ускорения в классической задаче поиска в неструктурированной базе данных [8].

3.4.5. 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 унитарной динамики квантовых систем. Его подробное рассмотрение дано в [8]. 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") (3.3).
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 (3.4)), однокубитной операции

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

и двух-q-битной операции CNOT. Известны и другие конечные универсальные наборы мало-q-битных операций.
3. С другой стороны, всякое квантовое вычисление, использующее операцию Адамара (3.4), фазовую операцию

$$S = T^2\colon |0\rangle \to |0\rangle, \qquad |1\rangle \to i|1\rangle,$$

и операцию CNOT, а также приготовление состояний и измерение в вычислительном базисе, эффективно моделируется классическим вычислением (теорема Готтесмана–Нилла, п. 10.5.4 [8]).
4. Для любого фиксированного конечного универсального набора операций 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 этого 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 [8].

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 использовать 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, составляющая теорию кодирования.

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 мала. Рассмотрим код $0 \to 00$, $1 \to 11$. Хотя этот код и позволяет обнаружить и исправить некоторые ошибки, он имеет существенный недостаток: например, в ситуации одной ошибки во втором или первом бите $00 \to 01$, $11 \to 01$ мы не можем сказать, какое сообщение было закодировано. Но от этого недостатка можно избавиться, если добавить еще один разряд: $0 \to 000$, $1 \to 111$. Такой код будет уже помехоустойчивым по отношению к ошибке в любом одном бите, в том смысле, что испорченные слова, получающиеся из 000 и 111, никогда не совпадают и поэтому безошибочно различимы.

В квантовой статистике безошибочная различимость чистых состояний (а мы здесь будем иметь дело только с ними) равносильна ортогональности векторов состояний. Прямолинейное обобщение классического рецепта повторения сообщений на квантовый случай наталкивается на трудность – квантовую информацию невозможно размножить. По самому существу квантовой информации, при передаче через канал с ошибками безошибочно должны приниматься не только базисные состояния, но и всевозможные их суперпозиции $|\psi\rangle$. Таким образом, прямолинейное обобщение кода повторения

$$|\psi\rangle \to |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle$$

является неосуществимым. На первый взгляд, задача кажется неразрешимой, тем не менее в работах Шора и Стина независимо были построены первые примеры квантовых кодов, исправляющих ошибки. Многие авторы сделали вклад в последующее развитие теории, фрагменты которой представлены в этом разделе, см. обзор в [8].

Следуя классической аналогии, рассмотрим сначала код

$$|\,0\rangle \to |\,000\rangle, \quad |\,1\rangle \to |\,111\rangle. \tag{3.9}$$

Реализующий его унитарный оператор может быть построен с помощью двух вспомогательных q-битов, находящихся в начальном состоянии $|0\rangle$:

$$|\widetilde{\psi}\rangle = (CNOT_{13} \cdot CNOT_{12})(|\psi\rangle \otimes |0\rangle \otimes |0\rangle),$$

где $CNOT_{12}$ обозначает операцию $CNOT$ над первым и вторым q-битами. Произвольное чистое состояние $|\psi\rangle = a|\,0\rangle + b|\,1\rangle$ сигнального q-бита при таком способе кодирования переходит в состояние трех кодовых q-битов $|\widetilde{\psi}\rangle = a|\,000\rangle + b|\,111\rangle$. Как показано ниже, такой код позволяет исправить "переворот бита", т. е. переход $|\,0\rangle \leftrightarrow |\,1\rangle$ в любом одном q-бите. Это означает следующее: любые два ортогональных вектора кода остаются ортогональными, если не более чем в одной из кодовых позиций 1,2,3 происходит переворот бита.

В общем случае обнаружение и исправление ошибки можно осуществить путем измерения специального набора *проверочных операторов*, исходы которых образуют *синдром ошибки*, указывающий на кодовую позицию, в которой произошла ошибка, и последующего применения соответствующего унитарного *исправляющего оператора*. Для кода (3.9) возможным набором проверочных операторов является коммутирующая пара

$$\sigma_z \otimes \sigma_z \otimes I, \quad I \otimes \sigma_z \otimes \sigma_z. \tag{3.10}$$

Собственными значениями каждого из этих операторов являются ± 1, причем собственное значение $1\,(-1)$ для первого оператора отвечает совпадению (несовпадению) 1-го и 2-го битов, а для второго – 2-го и 3-го битов. При этом соответствие между парой собственных значений, парой общих собственных векторов синдромом ошибки и исправляющим оператором имеет вид

$(1,1)$	$\|000\rangle$, $\|111\rangle$	нет ошибки	$I \otimes I \otimes I$
$(-1,1)$	$\|100\rangle$, $\|011\rangle$	ошибка в 1-й позиции	$\sigma_x \otimes I \otimes I$
$(-1,-1)$	$\|010\rangle$, $\|101\rangle$	ошибка во 2-й позиции	$I \otimes \sigma_x \otimes I$
$(1,-1)$	$\|001\rangle$, $\|110\rangle$	ошибка в 3-й позиции	$I \otimes I \otimes \sigma_x$

Однако код (3.9) не исправляет "переворот фазы" типа $|\,0\rangle \leftrightarrow |\,0\rangle, |\,1\rangle \leftrightarrow -|\,1\rangle$. В самом деле, в результате такой фазовой ошибки в любом одном бите получим $a|\,000\rangle - b|\,111\rangle$ вместо $a|\,000\rangle + b|\,111\rangle$, так что нельзя даже определить, в каком бите произошла ошибка.

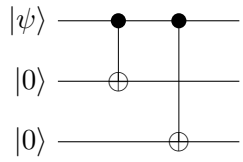

Рис. 3.6. Схема кодирования для трех-q-битового кода, исправляющего классические ошибки

Теперь заметим, что преобразование Адамара

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

отображает переворот фазы в переворот бита и наоборот, а именно

$$H\,\sigma_x\,H = \sigma_z, \quad H\,\sigma_z\,H = \sigma_.$$

Преобразуя соответствующим образом код (3.9), получим новый код

$$|\,0\rangle \to \frac{1}{2^{3/2}}(|\,0\rangle + |\,1\rangle)(|\,0\rangle + |\,1\rangle)(|\,0\rangle + |\,1\rangle) = |\,{+}{+}{+}\rangle, \tag{3.11}$$

$$|\,1\rangle \to \frac{1}{2^{3/2}}(|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle)(|\,0\rangle - |\,1\rangle) = |\,{-}{-}{-}\rangle, \tag{3.12}$$

который исправляет переворот фазы в любом одном q-бите (но не исправляет переворот бита). При этом проверочные операторы

$$\sigma_x \otimes \sigma_x \otimes I, \quad I \otimes \sigma_x \otimes \sigma_x, \tag{3.13}$$

а таблица обнаружения и исправления ошибок принимает вид

$(1,1)$	$\|\,{+}{+}{+}\rangle, \|\,{-}{-}{-}\rangle$	нет ошибки	$I \otimes I \otimes I$
$(-1,1)$	$\|\,{-}{+}{+}\rangle, \|\,{+}{-}{-}\rangle$	ошибка в 1-й позиции	$\sigma_z \otimes I \otimes I$
$(-1,-1)$	$\|\,{+}{-}{+}\rangle, \|\,{-}{+}{-}\rangle$	ошибка во 2-й позиции	$I \otimes \sigma_z \otimes I$
$(1,-1)$	$\|\,{+}{+}{-}\rangle, \|\,{-}{-}{+}\rangle$	ошибка в 3-й позиции	$I \otimes I \otimes \sigma_z$

Задача 55. Постройте схему кодирования для трех-q-битового кода, исправляющего фазовые ошибки.

Код Шора, который исправляет как переворот бита, так и переворот фазы в одном q-бите, получается комбинированием кодов (3.9), (3.12) и требует девять q-битов

$$\begin{aligned}|\,0\rangle &\to \frac{1}{2^{3/2}}(|\,000\rangle + |\,111\rangle)(|\,000\rangle + |\,111\rangle)(|\,000\rangle + |\,111\rangle)\\ |\,1\rangle &\to \frac{1}{2^{3/2}}(|\,000\rangle - |\,111\rangle)(|\,000\rangle - |\,111\rangle)(|\,000\rangle - |\,111\rangle)\end{aligned} \tag{3.14}$$

Проверочные операторы также являются комбинацией проверочных операторов для кодов (3.9), (3.12), см. (3.18). Используя соотношение $-i\sigma_y = \sigma_x\sigma_z$, получаем, что эти операторы позволяют обнаружить и комбинацию битовой и фазовой ошибки, а используя тот факт, что матрицы Паули (включая единичную) образуют базис в пространстве всех 2×2-матриц, можно показать, что этот код позволяет обнаружить и исправить любую ошибку, возникающую в результате применения *произвольного* оператора в одном (любом) из q-битов. Проверка этого факта может быть осуществлена с помощью критерия, который будет получен далее.

Задача 56. Напишите уравнения (постройте схемы) для унитарных операторов, реализующих код (3.11) и код Шора, используя вспомогательные q-биты, операции Адамара и CNOT.

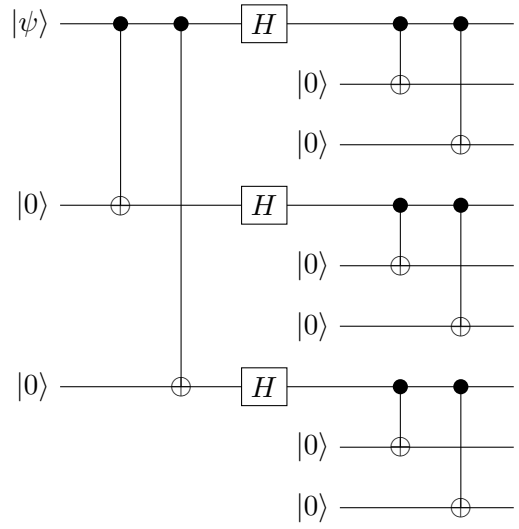

Рис. 3.7. Схема кодирования для девяти-q-битового кода Шора

3.5.2. Общая формулировка. Пусть $\mathcal{M}$, $\mathcal{N}$ гильбертовы пространства, первое из которых является пространством векторов кодируемых состояний, а второе – кодирующим пространством. *Кодом* называется изометрическое отображение $V\colon \mathcal{M} \to \mathcal{N}$, переводящее состояния S в кодированные состояния VSV^* в 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 – фиксированные операторы *элементарных ошибок. Код $\mathcal{L}$ исправляет ошибки класса $\mathcal{E}$*, если для любых $\phi, \psi \in \mathcal{L}$, таких, что $\langle\phi|\,\psi\rangle = 0$, имеет место $\langle\phi|\,B_i^* B_j\,|\psi\rangle = 0$, для всех $i, j = 1, \ldots p$. Заметим, что в силу линейности этого условия по B_i^*, B_j, аналогичное условие выполняется для произвольных операторов ошибок из класса $\mathcal{E} = \mathrm{Lin}(B_1, \ldots, B_p)$.

Смысл этого условия в том, что ошибки не нарушают ортогональности векторов состояний кода.

Равносильное условие[4] состоит в том, что для какого-либо ортонормированного базиса $\{|\,k\rangle\}$ в $\mathcal{L}$ выполняется

$$\langle k|\,B_i^* B_j\,|\,k\rangle = \langle l|\,B_i^* B_j\,|\,l\rangle \quad \text{для всех } k, l, \tag{3.15}$$

$$\langle k|\,B_i^* B_j\,|\,l\rangle = 0, \quad \text{для } k \neq l. \tag{3.16}$$

В самом деле, разлагая векторы ϕ, ψ по базису $\{|\,k\rangle\}$, получаем

$$\langle\phi|\,B_i^* B_j\,|\,\psi\rangle = \langle k|\,B_i^* B_j\,|\,k\rangle\langle\phi|\,\psi\rangle,$$

откуда следует условие исправления ошибок. Обратное утверждение получаем, рассматривая пары ортогональных векторов $|\,k\rangle, |\,l\rangle$ и $|\,k\rangle + |\,l\rangle, |\,k\rangle - |\,l\rangle$ при $k \neq l$.

Задача 57. Проверьте выполнение равенств (3.15), (3.16) для кода (3.9) и операторов элементарных ошибок $I, \sigma_x^{(1)}, \sigma_x^{(2)}, \sigma_x^{(3)}$, представляющих переворот одного из трех битов.

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

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

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 задаваться произвольным оператором. Операторами элементарных ошибок в каждом q-бите служат матрицы Паули

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

причем σ_x описывает переворот бита, σ_z переворот фазы, а $\sigma_y = i\sigma_x\sigma_z$ – их комбинацию. Вместе с единичным оператором $\sigma_0 = I$, который соответствует отсутствию ошибки, они образуют базис в пространстве наблюдаемых q-бита.

Пример Шора демонстрирует возможность исправления ошибок из $\mathcal{E}(n,1)$, если $n \geqslant 9$ (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 в каждом элементе схемы не превосходит некоторого порогового значения (fault-tolerant quantum computing) [4], [8], раздел 10.6. Различные оценки дают пороговое значение от 10^{-6} до 10^{-4}, что в принципе представляется достижимой величиной.

Задача 58. Проверьте выполнение равенств (3.15), (3.16) для кода Шора и операторов элементарных ошибок $\sigma_\gamma^{(1)}, \sigma_\gamma^{(2)}, \sigma_\gamma^{(3)}$, $\gamma = 0, x, y, z$.

3.5.3. Симплектические коды. Рассмотрим поле $B = \{0,1\}$ с обычными бинарными операциями сложения и умножения. Заметим, что правила умножения (1.29) для матриц Паули 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 форму 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 коммутационные соотношения [5]

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

Для системы из n q-битов рассмотрим $2n-$мерное векторное пространство B^{2n} над полем B, состоящее из векторов $f = (\gamma_1, \ldots, \gamma_n)$. Введем эрмитовы операторы $\sigma(f) = \sigma_{\gamma_1}^1 \otimes \cdots \otimes \sigma_{\gamma_n}^n \in$

B^{2n}, удовлетворяющие каноническим коммутационным соотношениям

$$\sigma(f)\sigma(g) = i^{\Delta(f,g)}\sigma_{f+g} = (-1)^{\Delta(f,g)}\sigma(g)\sigma(f), \quad f,g \in B^{2n}, \tag{3.17}$$

где $\Delta(f,g) = \Delta(\gamma_1,\gamma_1') + \cdots + \Delta(\gamma_n,\gamma_n')$, если $g = (\gamma_1',\ldots,\gamma_n')$. Форма $\Delta(f,g)$ определяет в B^{2n} структуру симплектического пространства, откуда и происходит название "симплектический код".

Заметим, что операторы $\sigma(g), \sigma(f)$ коммутируют (антикоммутируют) тогда и только тогда, когда $\Delta(f,g) = 0 \pmod 2$, (соответственно $\Delta(f,g) \neq 0 \pmod 2$). Пусть $g_1,\ldots,g_{n-k}$ линейно независимые векторы из B^{2n} такие, что $\Delta(g_i,g_j) = 0 \pmod 2$ для всех i,j. Тогда эрмитовы операторы $\sigma(g_1),\ldots,\sigma(g_{n-k})$ коммутируют между собой, следовательно, представляют совместно измеримые наблюдаемые, принимающие значения ± 1. *Симплектическим кодом с проверочным операторами* $\sigma(g_1),\ldots,\sigma(g_{n-k})$ называется линейное подпространство

$$\mathcal{L} = \left\{\psi \in (\mathbb{C}^2)^{\otimes n}\colon \quad \sigma(g_j)\psi = \psi, \quad j = 1,\ldots,n-k\right\}.$$

Легко видеть, что $\dim \mathcal{L} = 2^k$.

Пусть $\mathcal{E}$ класс ошибок, порождаемый элементарными ошибками вида $\sigma(f)$, $f \in E$, где E – некоторое подмножество B^{2n}. В случае $\mathcal{E} = \mathcal{E}(n,m)$ имеем $E = E(n,m) = \left\{g \in B^{2n}\colon \operatorname{wt}(g) \leqslant m\right\}$, где вес $\operatorname{wt}(g)$ равен числу ненулевых компонент вектора g. Из канонических коммутационных соотношений (3.17) следует, что для любого кодового вектора $|\psi\rangle \in \mathcal{L}$ и элементарной ошибки $\sigma(f)$, вектор $\sigma(f)|\psi\rangle$ является собственным вектором проверочных операторов $\sigma(g_j)$, $j = 1,\ldots,n-k$, с собственными значениями $(-1)^{\Delta(g_j,f)}$:

$$\sigma(g_j)[\sigma(f)|\psi\rangle] = (-1)^{\Delta(g_j,f)}\sigma(f)\sigma(g_j)|\psi\rangle = (-1)^{\Delta(g_j,f)}[\sigma(f)|\psi\rangle].$$

Совокупность этих значений образует *синдром ошибки*. Ошибки $\sigma(f_1),\sigma(f_2)$ *неразличимы*, если их синдромы совпадают, т.е. $\Delta(g_j,f_1) = \Delta(g_j,f_2) \pmod 2$. (Отметим, что в силу двоичной природы операций в B^{2n} вектор $f_1 - f_2$ совпадает с $f_1 + f_2$.)

Обозначим через G $(n-k)$-мерное подпространство пространства B^{2n}, порожденное векторами $g_1,\ldots,g_{n-k}$. Отметим, что $\Delta(f,g) = 0 \pmod 2$, $f,g \in G$, поэтому $G \subset G^\perp$, где

$$G^\perp = \left\{f \in B^{2n}\colon \Delta(f,g) = 0 \pmod 2, \quad g \in G\right\}.$$

Таким образом, ошибки $\sigma(f_1),\sigma(f_2)$ неразличимы, если $f_1 - f_2 \in G^\perp$. Ошибки *эквивалентны*, если $f_1 - f_2 \in G$. Поскольку $G \subset G^\perp$, эквивалентные ошибки неразличимы, но обратное, вообще говоря, неверно.

Теорема 8 [5]. *Симплектический код $\mathcal{L}$ исправляет ошибки класса $\mathcal{E}$, если любые две неразличимые ошибки $\sigma(f_1),\sigma(f_2)$ эквивалентны.*

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 $d = \min\left\{\operatorname{wt}(g)\colon g \in G^\perp \setminus G\right\} \geqslant 3$, то данный код исправляет любые ошибки в $m = \left[\frac{d-1}{2}\right]$ из n q-битов.

Доказательство. Проверим выполнение условия исправления ошибок. Пусть ψ,φ – ортогональные векторы из $\mathcal{L}$, и $f_1,f_2 \in E$. Если ошибки $\sigma(f_1),\sigma(f_2)$ различимы, то векторы $\sigma(f_1)\psi,\sigma(f_2)\varphi$ являются собственными векторами коммутирующих проверочных операторов

[5] 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.

с различными наборами собственных значений и, следовательно, они ортогональны. Если же 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. Последнее вытекает из канонических коммутационных соотношений и того факта, что $f_1 - f_2$ разлагается по базису $g_1, \ldots, g_{n-k}$.

Процедура исправления элементарной ошибки состоит из двух этапов: сначала производится измерение проверочных операторов, в результате чего находится синдром ошибки; после этого ошибка определяется с точностью до эквивалентности; применяя оператор ошибки, получаем исходное состояние. По поводу исправления произвольной ошибки из $\mathcal{E}$ см. раздел 10.5.5 в [8].

Задача 59. Убедитесь, что код Шора является симплектическим кодом с проверочными операторами $\sigma(g_j); j = 1, \ldots, 8$, где

$$\begin{array}{lcccccccccc}
g_1 & = (z & z & 0 & 0 & 0 & 0 & 0 & 0 & 0), \\
g_2 & = (0 & z & z & 0 & 0 & 0 & 0 & 0 & 0), \\
g_3 & = (0 & 0 & 0 & z & z & 0 & 0 & 0 & 0), \\
g_4 & = (0 & 0 & 0 & 0 & z & z & 0 & 0 & 0), \\
g_5 & = (0 & 0 & 0 & 0 & 0 & 0 & z & z & 0), \\
g_6 & = (0 & 0 & 0 & 0 & 0 & 0 & 0 & z & z), \\
g_7 & = (x & x & x & x & x & x & 0 & 0 & 0), \\
g_8 & = (0 & 0 & 0 & x & x & x & x & x & x).
\end{array} \tag{3.18}$$

Первые шесть операторов обнаруживают позицию, в которой произошел переворот бита, а последние два – блок, в котором произошел переворот фазы (т.е. изменение знака).

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: $y_k = x_k \oplus \kappa_k$, $x_k = y_k \oplus \kappa_k$. Здесь $x = (x_1, \ldots, x_m)$ – кодируемое сообщение, а $y = (y_1, \ldots, y_m)$ – закодированное сообщение, которое A пересылает B для последующего декодирования. Такой способ кодирования/декодирования носит название *шифр Вернама*.

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 $\{|+\rangle, |-\rangle\}$, точнее, с вероятностью $1/2$ переходят в состояния $|+\rangle$ или $|-\rangle$, и наоборот, состояния $|+\rangle, |-\rangle$ при измерении в базисе $\{|0\rangle, |1\rangle\}$ переходят в состояния этого базиса.

3.6.1. Протокол $BB84$. Дадим пошаговое описание протокола, предложенного Беннетом и Брассаром в 1984 г. Всюду далее δ – малое положительное число, которое подбирается позднее для получения нужного уровня вероятности.

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 $\left\{\frac{1}{2}, \frac{1}{2}\right\}$).
2. Участник A создает чистое состояние с вектором

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

где

$$|\psi_{00}\rangle = |0\rangle, \qquad |\psi_{10}\rangle = |1\rangle,$$
$$|\psi_{01}\rangle = |+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \qquad |\psi_{11}\rangle = |-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right). \tag{3.19}$$

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 $\approx N/2 = (2+\delta/2)n \gtrsim 2n$ (см. ниже соотношение (3.20)). 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 в передачу по квантовому каналу, поэтому A и B прекращают этот раунд протокола и пытаются устранить возможность вмешательства.
5. A и B выполняют тесты, чтобы определить величину возмущения из-за возможного шума или подслушивания. Для этого A наугад выбирает n битов просеянного ключа и сообщает их значения (вместе с их номерами) участнику B по открытому каналу. Вероятность получить $\leqslant \nu n$ ошибок в этих n контрольных битах и при этом $\geqslant (\nu+\varepsilon)n$ ошибок в остальных битах имеет порядок $\exp[-O(\varepsilon^2 n)]$ (см. Упражнение 12.27 в [8]). Поэтому при больших n можно считать, что доля несовпадений в этих оставшихся секретными битах практически равна ν. Если A и B обнаруживают, что ν превосходит некоторое пороговое значение ν_t, то они решают, что было вмешательство и также прекращают этот раунд протокола. Величина ν_t зависит от того, что известно о возможностях перехватчика E. В литературе приводятся значения $\nu_t = 0.11 \div 0.25$, в зависимости от вида атак перехватчика и используемых средств противодействия. Подробнее об этом см. в [3].
6. Если же установлено, что $\nu < \nu_t$, то A и B выполняют "согласование информации", чтобы устранить несовпадающие биты и "усиление конфиденциальности" по оставшимся $\approx n$

a	0	1	0	0	0	1	0	1
b	0	0	1	0	0	1	1	1
b'	0	0	1	1	1	0	1	1
$b = b'$	y	y	y	n	n	n	y	y
key	0	1	0				0	1

Рис. 3.8. Пример реализации протокола BB84 (N=8)

битам, чтобы свести на нет информацию, которую могла перехватить E в ходе предыдущих открытых операций. При этом весьма полезной оказывается верхняя квантовая оценка количества информации перехватчика E, вытекающая из неравенства (5.31), см. главу 5. Последние операции относятся к области классической информатики; их описание можно найти, например, в [8]. В итоге A и B получают $m \lesssim n$ битов совместного секретного ключа.

Покажем, что при малых δ

$$\mathsf{P}\left\{2n \leqslant \left(\#k : b_k = b'_k\right) \leqslant (2+\delta)n\right\} \geqslant 1 - 2\exp\left[-\frac{1}{8}n\left(\delta^2 + o(\delta^2)\right)\right]. \tag{3.20}$$

Вводя бит несовпадения $\nu_k = b_k \oplus b'_k$, имеем $\mathsf{P}\{\nu_k = 0\} = \mathsf{P}\{\nu_k = 1\} = \frac{1}{2}$, $\mathsf{M}\nu_k = \frac{1}{2}$. Интересующая нас вероятность равна ($N = (4+\delta)n$)

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

в силу симметрии распределения суммы относительно ее математического ожидания $N/2 = (2+\delta/2)n$, при этом

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

где $\varepsilon = \frac{1}{2} - \frac{2n}{(4+\delta)n} = \frac{1}{8}\left(\delta + o(\delta)\right)$. Заметим, что из неравенства Чебышева, используемого при доказательстве закона больших чисел, вытекает

$$\mathsf{P}\left\{\left|\frac{1}{N}\sum_{k=1}^{N}\left(\nu_k - \frac{1}{2}\right)\right| > \varepsilon\right\} \leqslant \frac{\mathsf{D}\nu_k}{N\varepsilon^2} = \frac{1}{4N\varepsilon^2} \to 0$$

при $N \to \infty$. Интересующая нас гораздо более точная экспоненциальная оценка (3.20) вытекает из *неравенства Чернова*:

$$\mathsf{P}\left\{\frac{1}{N}\sum_{k=1}^{N}\left(\nu_k - \frac{1}{2}\right) \leqslant -\varepsilon\right\} = \mathsf{P}\left\{\frac{1}{N}\sum_{k=1}^{N}\left(\nu_k - \frac{1}{2}\right) \geqslant \varepsilon\right\} \leqslant \exp\left[-2N(\varepsilon^2 + o(\varepsilon^2))\right]. \tag{3.21}$$

Доказательство. Пусть $s > 0$, тогда

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

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

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 (3.21), а учитывая, что $N = (4+\delta)n$, имеем (3.20).

3.6.2. Протокол $B92$. Этот протокол, предложенный Беннетом в 1992 г., 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 $a'_k = 1$. Исход "плюс" кодируется как 0, а "минус" как 1. Таким образом, результаты измерений B 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 $b_k = 1$. Полученные таким образом секретные биты $a_k = 1 - a'_k$ составляют просеянный ключ.
4. Следующие шаги аналогичны соответствующим шагам протокола $BB84$.

Задача 60. Покажите, что $\mathsf{P}\{b_k = 1\} = 1/4$. Поэтому для создания просеянного ключа той же длины потребуется примерно вдвое больше битов, чем в протоколе $BB84$.

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

$$|\psi_k\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|++\rangle + |--\rangle), \tag{3.22}$$

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

Задача 61. Проверьте второе равенство, используя определения (3.19).

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

Из равенства (3.22) следует, что на если на k-м шаге A и B проводят измерения в одинаковых базисах: $\{|0\rangle, |1\rangle\}$ либо $\{|+\rangle, |-\rangle\}$, то они получают совпадающие результаты, $\mathsf{P}\{a'_k = b'_k\} = 1$.

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

Описанные выше протоколы реализованы в эксперименте; более того, для первых двух созданы коммерческие образцы оборудования. Доказательство их стойкости является непростой задачей, которая требует привлечения всего инструментария квантовой теории информации [8].

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[6].

[6] 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 (1.20).

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 9. *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*

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

для которого

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

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 (1.18), вероятностная нормировка – из второго условия в (4.2), а аффинность – из линейности следа в (4.3). Доказательство прямого утверждения дано в [12]. Оно основывается на двух математических фактах: операторы плотности линейно порождают пространство $\mathcal{L}$ всех линейных операторов в $\mathcal{H}$ (ср. Лемму 1 в разделе 1.3 ч. I); и всякая аффинная функция на выпуклом множестве квантовых состояний $\mathcal{S}$ однозначно продолжается до комплексно-линейной функции на $\mathcal{L}$.

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

Разложение единицы называется *ортогональным*, если

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

Задача 62. Второе условие является следствием первого. Таким образом, ортогональное разложение единицы характеризуется свойством: все операторы M_x – проекторы.

Как было установлено при рассмотрении стандартной статистической модели квантовой механики, (см. раздел 1.3), спектральное разложение

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

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

Основываясь на этих понятиях, можно было бы назвать *обобщенной квантовой наблюдаемой* со значениями в произвольном конечном множестве $\mathcal{X}$ 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.

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

$$S = \sum_{\omega} s_\omega |\omega\rangle\langle\omega|, \quad s_\omega \geqslant 0, \quad \sum s_\omega = 1,$$

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

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

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 $M_x = \sum\limits_{\omega} M(x|\omega)|\omega\rangle\langle\omega|$. Тогда собственные числа удовлетворяют условиям): $0 \leqslant M(x|\omega) \leqslant 1$ 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 10. *Всякая четкая наблюдаемая $M \in \mathfrak{M}_\mathcal{X}$ экстремальна. 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 (1.26)

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

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 4 п. 1.8 операторы 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 $M^\pm(x_0|\omega_0) = M(x_0|\omega_0) \pm \epsilon$, $M^\pm(x_1|\omega_0) = M(x_1|\omega_0) \mp \epsilon$, где ϵ достаточно малое положительное число, 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 измерений с двумя исходами (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 $\{|\psi_j\rangle,\ j=1,\dots,n\}\subset\mathcal{H}$ называется *переполненной*, если

$$\sum_{j=1}^{n}|\psi_j\rangle\langle\psi_j|=I. \tag{4.6}$$

Другими словами

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

С необходимостью $n\geqslant d$, 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 63. Система $\{|\psi_j\rangle\}$ является переполненной тогда и только тогда, когда
1) система полна, т.е. $\{|\psi_j\rangle,\ j=1,\dots,n\}^{\perp}=\{0\}$;
2) матрица $P=[\langle\psi_j|\psi_k\rangle]_{j,k=1,\dots,n}$ 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

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

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 $P=[\langle\psi_j|\psi_k\rangle]_{j,k=1,\dots,n}$

задает проекцию пространства $\widetilde{\mathcal{H}} = \mathbb{C}^n$ на $V\mathcal{H}$. Рассмотрим стандартный 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 $\{|\psi_x\rangle\}$ связано разложение единицы, т.е. наблюдаемая

$$M_x = |\psi_x\rangle\langle\psi_x|. \tag{4.7}$$

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

$$M_x = G^{-1/2}|\phi_x\rangle\langle\phi_x|G^{-1/2} \tag{4.8}$$

задает наблюдаемую[2], в определенном смысле "измеряющую" состояния $|\phi_x\rangle\langle\phi_x|$.

Теорема 11. *Наблюдаемая* (4.7) *является экстремальной тогда и только тогда, когда операторы* M_x *линейно независимы.*

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

$$\sum_x c_x|\psi_x\rangle\langle\psi_x| = 0. \tag{4.9}$$

Взяв достаточно малое $\epsilon > 0$, определим

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

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 (4.9) 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

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

Умножая справа и слева на эрмитов оператор $I - \frac{|\psi_x\rangle\langle\psi_x|}{\langle\psi_x|\psi_x\rangle}$, который является проектором на подпространство, ортогональное вектору $|\psi_x\rangle$, 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, № 2, 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 12. *Пусть $\vec{a}_j$, $j = 1, \dots, m$, – система единичных векторов в $\mathbb{R}^3$, такая что $\sum_{j=1}^m \vec{a}_j = 0$. Тогда векторы $\sqrt{2/m}|\vec{a}_j\rangle$, $j = 1, \dots, m$, образуют переполненную систему в пространстве q-бита $\mathcal{H}$, так что*

$$\frac{2}{m}\sum_{j=1}^m |\vec{a}_j\rangle\langle\vec{a}_j| = I. \tag{4.10}$$

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

Доказательство. Первое утверждение, т.е. соотношение (4.10), непосредственно следует из (1.36). Также используя это соотношение получаем, что линейная зависимость операторов $|\vec{a}_j\rangle\langle\vec{a}_j|$, равносильная, в силу теоремы 11, неэкстремальности наблюдаемой, означает, что

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

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 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

$$M_1 = \frac{2}{3}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_{2,3} = \frac{2}{3}\begin{bmatrix} 1/4 & \pm\sqrt{3}/4 \\ \pm\sqrt{3}/4 & 3/4 \end{bmatrix}. \tag{4.11}$$

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

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

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

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

причем константа оказывается равной $[d^2(d+1)]^{-1}$. В самом деле, учитывая, что $\operatorname{Tr} M_j^2 = 1/d^2$, имеем

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

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

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

Задача 64. 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 $\operatorname{Tr} SM_k = \operatorname{Tr} S'M_k$, где S' – оператор в правой части равенства.

Восстановление состояния по статистике измерений (одного или целого ряда) называют *томографией* квантового состояния. Например, формула (1.30) 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:

$$|\langle e_j|h_k\rangle|^2 = \frac{1}{d}, \tag{4.12}$$

(где $d = 2$) для всех j, k. В общем случае, базисы в $\mathcal{H}$, $\dim \mathcal{H} = d$, обладающие свойством (4.12), называются *равнонаклоненными* (mutually unbiased). Это понятие было введено знаменитым физиком Дж. Швингером в 1960 г. Доказано, что количество равнонаклоненных базисов в $\mathbb{C}^d$ не превосходит $d+1$. Существование $d+1$ равнонаклоненных базисов доказано для размерностей вида p^m, где p – простое число; для других размерностей (уже для $d = 6$) максимальное количество $\mathfrak{M}(d)$ равнонаклоненных базисов неизвестно, хотя имеются различные оценки [4]. Существует гипотеза, что $\mathfrak{M}(6) = 3$.

Измерения в равнонаклоненных базисах удобны для томографии квантовых состояний; они также находят применения в квантовой криптографии и при построении квантовых кодов, исправляющих ошибки.

4.6. Теорема Наймарка

Геометрический смысл неортогональных разложений единицы проясняет следующая теорема.

Теорема 13. *Пусть $\{M_x\}_{x\in\mathcal{X}}$ – разложение единицы в гильбертовом пространстве $\mathcal{H}$, $\dim \mathcal{H} = d$, $|\mathcal{X}| = n$. Существует гильбертово пространство $\widetilde{\mathcal{H}}$, $\dim \widetilde{\mathcal{H}} \leqslant n \cdot d$, изометрический оператор $V\colon \mathcal{H} \to \widetilde{\mathcal{H}}$ и ортогональное разложение единицы $\{E_x\}$ в $\widetilde{\mathcal{H}}$, такие, что*

$$M_x = V^* E_x V. \tag{4.13}$$

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.

[4] https://en.wikipedia.org/wiki/Mutually_unbiased_bases

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

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

Оно отличается от скалярного произведения тем, что соответствующая квадратичная форма может быть вырождена. Обозначим $\mathcal{H}_0 = \{\Psi \in \mathcal{H}_n\colon \langle\Psi|\Psi\rangle = 0\}$ и рассмотрим факторпространство $\mathcal{H}_n/\mathcal{H}_0$. 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

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

Этот оператор сохраняет псевдоскалярное произведение:

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

поскольку $\sum M_x = I$. Отсюда вытекает, что после факторизации получается изометрический оператор V из $\mathcal{H}$ в $\widetilde{\mathcal{H}}$.

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 $\langle\psi|V^*E_yV|\psi\rangle = \langle\psi|M_y|\psi\rangle$, откуда следует (4.13). □

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*

$$M_j = \mathrm{Tr}_{\mathcal{H}_0}(I \otimes |\psi_0\rangle\langle\psi_0|)E_j. \tag{4.14}$$

Доказательство. Согласно теореме Наймарка $M_j = V^*\widetilde{E_j}V$, где $V\colon \mathcal{H} \to \widetilde{\mathcal{H}}$ – 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,

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

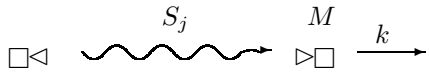

Рис. 4.1. Различение состояний

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

$$|\psi_0\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad |\phi \otimes \psi_0\rangle = \begin{bmatrix} |\phi\rangle \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Обозначая E_j подходящие расширения операторов $\widetilde{E_j}$, имеем для $\psi \in \mathcal{H}$:

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

т.е. соотношение (4.14). □

Итак, всякую наблюдаемую можно реализовать в виде четкой наблюдаемой в составной системе за счет добавления вспомогательной системы (квантового генератора случайности), 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 квантовая система a priori может находиться в одном из состояний S_j, $j = 1, \ldots, n$. Над системой можно производить произвольное измерение. Требуется найти *оптимальное* измерение, позволяющее наилучшим образом выяснить, в каком из этих состояний действительно находится данная система. Такая постановка задачи характерна для теории связи, где j – это сигнал, передаваемый квантовым носителем информации. Она основывается на аналогии с задачей различения гипотез в математической статистике [11].

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 $p_M(k|j) = \mathrm{Tr}\, S_j M_k$. 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.

4.7.2. Различение по максимуму правдоподобия. Будем максимизировать вероятность правильного решения

$$\mathcal{P}\{M\} = \sum_{j=1}^{n} \pi_j \operatorname{Tr} S_j M_j = \operatorname{Tr}\Bigl(\sum_{j=1}^{n} W_j M_j\Bigr),$$

где $W_j = \pi_j S_j \geqslant 0$.

Множество наблюдаемых, по которым ведется оптимизация

$$\mathfrak{M}_n = \Bigl\{ M = \{M_k\}_{k=1,\dots,n} \colon M_k \geqslant 0,\ \sum_{k=1}^{n} M_k = I \Bigr\}$$

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

$$\mathcal{P}\Bigl\{ \sum p_\lambda M^\lambda \Bigr\} = \sum p_\lambda \mathcal{P}\{M^\lambda\}.$$

Оптимизация аффинной функции, заданной на выпуклом множестве – типичная задача выпуклого программирования. Из общих фактов теории вытекает, что средняя вероятность правильного решения $\mathcal{P}\{M\}$ достигает максимума в крайней точке множества $\mathfrak{M}_n$. Однако оптимальная наблюдаемая, как правило, неединственна и в этом случае не все оптимальные наблюдаемые обязаны быть экстремальными.

Теорема 14. *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) $\Lambda^0 \geqslant W_k$.

При этом

$$\Lambda^0 = \sum_{k=1}^{n} W_k M_k^0 = \sum_{k=1}^{n} M_k^0 W_k \tag{4.15}$$

и

$$\max\{\mathcal{P}\{M\} \colon M \in \mathfrak{M}_n\} = \operatorname{Tr}\Lambda^0. \tag{4.16}$$

Доказательство. Докажем достаточность условий теоремы.

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

$$\begin{aligned}\mathcal{P}\{M\} = \operatorname{Tr}\sum_k W_k M_k &\overset{2)}{\leqslant} \operatorname{Tr}\sum_k \Lambda^0 M_k \\ &= \operatorname{Tr}\Lambda^0 \overset{1)}{=} \operatorname{Tr}\sum_k W_k M_k^0 = \mathcal{P}\{M^0\}.\end{aligned}$$

Здесь был использован простой факт, вытекающий из (1.18): Для $B \geqslant 0$ и A_1, A_2, таких что $A_1 \leqslant A_2$, имеет место $\operatorname{Tr} A_1 B \leqslant \operatorname{Tr} A_2 B$, 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

$$\operatorname{Tr}\sum_k W_k X_k^2 - \operatorname{Tr}\Lambda\biggl(\sum_k X_k^2 - I\biggr), \tag{4.17}$$

где Λ эрмитов оператор, играющий роль операторного множителя Лагранжа, по всевозможным наборам эрмитовых операторов X_k. Пусть X_k^0 оптимальный набор, положим $X_k = X_k^0 + \epsilon Y_k$, и рассмотрим (4.17) как функцию от ϵ. Рассматривая коэффициенты при ϵ и ϵ^2, получаем условия

$$\begin{aligned}&\operatorname{Tr}[(W_k - \Lambda)X_k^0 + X_k^0(W_k - \Lambda)]Y_k = 0,\\ &\operatorname{Tr}(W_k - \Lambda)Y_k^2 \leqslant 0\end{aligned}$$

для произвольных эрмитовых Y_k, т.е.

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

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

Соотношения (4.15), (4.16) следуют из условий 1), 2).

Задача 65. Имеет место соотношение двойственности

$$\max\{\mathcal{P}\{M\}\colon M \in \mathfrak{M}_n\} = \min\{\operatorname{Tr}\Lambda\colon \Lambda \geqslant W_k, k = 1, \ldots, n\}, \tag{4.18}$$

при этом операторный множитель Лагранжа Λ^0 является единственным решением двойственной задачи в правой части (4.18).

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

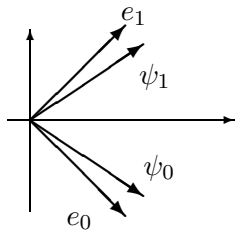

Рис. 4.2. Различение двух чистых состояний.

Пример 1. *Различение классических состояний.* Пусть операторы $W_k = \pi_k S_k$ коммутируют, тогда существует ортонормированный базис $\{|\omega\rangle\}$, в котором они все диагонализуются

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

Тогда решение дается формулами

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

где $\Lambda^0(\omega) = \max\limits_k W_k(\omega)$ – верхняя огибающая функций $W_k(\omega) = \pi_k S_k(\omega)$; $k = 1, \ldots, n$;

$$M_k^0 = \sum_\omega \mathbf{1}_{\Omega_k}(\omega)|\omega\rangle\langle\omega|,$$

где $\mathbf{1}_{\Omega_k}$ обозначает индикатор подмножества Ω_k, и $\Omega_k \subset \{\omega\colon \Lambda^0(\omega) = W_k(\omega)\}$ образуют разбиение множества $\Omega = \{\omega\}$. Легко проверить, что условия 1), 2) теоремы при этом выполняются.

Это сводится к принципу *максимального правдоподобия* в классической статистике: k-е решение необходимо принимать для тех ω, для которых значение $\pi_k S_k(\omega)$ 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 10).

Пример 2. *Различение двух квантовых состояний.* Произвольная наблюдаемая с двумя значениями имеет вид $M = \{M_0, M_1\}$, $M_{0,1} \geqslant 0$, $M_1 = I - M_0$, причем четкие наблюдаемые характеризуются условием $M_0^2 = M_0$, которое в точности соответствует крайним точкам "некоммутативного отрезка" $\{0 \leqslant M_0 \leqslant I\}$. Таким образом, для различения двух состояний достаточно четких наблюдаемых.

Задача 66. Докажите, что крайними точками выпуклого множества $\{0 \leqslant X \leqslant I\}$ являются проекторы и только они.

Приведем явное решение задачи различения двух состояний. Пусть S_0, S_1 произвольные операторы плотности, $M = \{M_0, M_1\}$ – оптимальная наблюдаемая. Согласно (4.15), оператор

$$\Lambda^0 = \pi_0 S_0 M_0 + \pi_1 S_1 M_1 = \pi_1 S_1 + (\pi_0 S_0 - \pi_1 S_1) M_0$$

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

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

Очевидным решением этих двух неравенств 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. По построению условия 1), 2) теоремы выполняются. При этом

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

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

Рассмотрим специально случай чистых состояний. 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. [11], гл. IV), для которого

$$\max \mathcal{P}\{M^0\} = \frac{1}{2}\big(1 + \sqrt{1 - 4\pi_0\pi_1\,|\,\langle\psi_1|\psi_0\rangle|^{\,2}}\big).$$

Пусть теперь $\pi_0 = \pi_1 = 1/2$, тогда оптимальный базис расположен симметрично по отношению к $|\psi_0\rangle, |\psi_1\rangle$ (рис. 4.2) и

$$\max \mathcal{P}\{M^0\} = \frac{1}{2}\big(1 + \sqrt{1 - |\,\langle\psi_1|\psi_0\rangle|^{\,2}}\big). \tag{4.19}$$

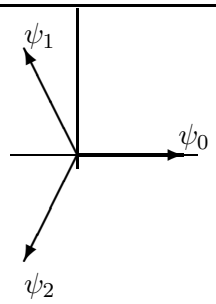

Рис. 4.3. Векторы трех состояний

Приведенное решение единственно в подпространстве, натянутом на векторы $|\psi_0\rangle, |\psi_1\rangle$. В ортогональном дополнении к этому подпространству его можно продолжить произвольным разложением единицы. Это дает представление о возможной неединственности решения задачи различения состояний.

Пример 3. Рассмотрим три чистых состояния q-бита $S_j = |\vec{a}_j\rangle\langle\vec{a}_j|$, $j = 1, 2, 3$, где

$$|\vec{a}_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\vec{a}_{2,3}\rangle = \begin{bmatrix} 1/2 \\ \pm\sqrt{3}/2 \end{bmatrix} \tag{4.20}$$

векторы состояний в $\mathcal{H} = \mathbb{C}^2$, соответствующие равноугольной конфигурации $\vec{a}_1 = (0, 0, 1)$, $\vec{a}_{2,3} = (\pm\sqrt{3}/2, 0, -1/2)$ в $\mathbb{R}^3$.

Рассмотрим соответствующее неортогональное разложение единицы $M_j^0 = \frac{2}{3}S_j$, $j = 1, 2, 3$ в $\mathcal{H}$, которое описывается соотношениями (4.11), и покажем, что в случае равновероятных состояний, $\pi_j = 1/3$, разложение единицы $\{M_j^0\}$ дает оптимальную наблюдаемую. Проверим условия теоремы 14. Поскольку $S_j^2 = S_j$, то

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

так что выполнено условие 2): $I/3 = \Lambda^0 \geqslant S_j/3$, 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 $M_1 = |e_1\rangle\langle e_1|$, $M_2 = |e_2\rangle\langle e_2|$, $M_3 = 0$, где $|e_1\rangle, |e_2\rangle$, – произвольный базис. Таким образом, задача сводится к оптимальному различению только двух состояний S_1, S_2, для которых $\langle\vec{a}_1|\vec{a}_2\rangle = 1/2$. 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.

4.7.3. "Безошибочное" различение состояний. Вернемся к задаче различения двух чистых состояний $S_0 = |\psi_0\rangle\langle\psi_0|$, $S_1 = |\psi_1\rangle\langle\psi_1|$. Для простоты будем считать, что они равновероятны. Рассматриваемое нами обобщение понятия квантовой наблюдаемой позволяет ввести статистическую процедуру M, обладающую интересным свойством: вероятности ошибочных решений равны нулю, $p_M(1|0) = p_M(0|1) = 0$. При этом, однако, приходится допустить третий исход измерения "?", означающий уклонение от определенного решения. Именно, рассмотрим разложение единицы с тремя исходами $0, 1, ?$, заданное соотношениями

$$M_0 = \frac{I - |\psi_1\rangle\langle\psi_1|}{1 + |\langle\psi_1|\psi_0\rangle|}, \quad M_1 = \frac{I - |\psi_0\rangle\langle\psi_0|}{1 + |\langle\psi_1|\psi_0\rangle|}, \quad M_? = I - M_0 - M_1.$$

Очевидно, что $M_0 \geqslant 0$, $M_1 \geqslant 0$.

Задача 67. Покажите, что $M_? \geqslant 0$. Используйте тот факт, что минимальное собственное значение оператора $|\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|$ равно $1 - |\langle\psi_1|\psi_0\rangle|$.

Имеем

$$p_M(0|1) = \langle\psi_1|M_0|\psi_1\rangle = 0, \quad p_M(1|0) = \langle\psi_0|M_1|\psi_0\rangle = 0. \tag{4.21}$$

Средняя вероятность правильного решения при этом равна

$$\frac{1}{2}\langle\psi_0|M_0|\psi_0\rangle + \frac{1}{2}\langle\psi_1|M_1|\psi_1\rangle = 1 - |\langle\psi_1|\psi_0\rangle|,$$

что строго меньше величины (4.19) для оптимальной наблюдаемой, если $0 < |\langle\psi_1|\psi_0\rangle| < 1$. На первый взгляд, свойство "безошибочности" (4.21) должно было бы привести к увеличению вероятности правильного решения, что, однако, противоречило бы оптимальности. Разрешение парадокса в том, что полные вероятности ошибок должны включать в себя и вероятности неопределенного исхода ?, так что в среднем "безошибочная" наблюдаемая $\{M_0, M_1, M_?\}$ проигрывает оптимальной. Тем не менее свойство (4.21) оказывается полезным и используется в некоторых приложениях.

4.7.4. "Степень совпадения" и другие меры близости двух состояний. В качестве меры близости двух квантовых состояний S_0, S_1 естественно использовать следовую норму

$$\|S_0 - S_1\|_1 = \sum_j |t_j|,$$

где t_j –собственные числа оператора $T = S_0 - S_1$. Однако нахождение собственных чисел, за исключением простых частных случаев (см. ниже), обычно представляет трудную задачу, и поэтому более удобными оказываются другие меры совпадения. Ниже мы дадим их описание и рассмотрим взаимосвязи между ними.

Задача 68. Для двух состояний q-бита (см. (1.30))

$$\|S(\vec{a}_1) - S(\vec{a}_0)\|_1 = |\vec{a}_1 - \vec{a}_0|.$$

В общем случае вопрос упрощается, если хотя бы одно из состояний – чистое.

Лемма 4. *Пусть ψ – единичный вектор и S – произвольное состояние. Имеют место неравенства*

$$2\left[1 - \langle\psi|S|\psi\rangle\right] \leqslant \| |\psi\rangle\langle\psi| - S \|_1 \leqslant 2\sqrt{1 - \langle\psi|S|\psi\rangle}. \tag{4.22}$$

Если $S = |\phi\rangle\langle\phi|$ – также чистое состояние, то второе неравенство превращается в равенство, т.е.

$$\| \,|\,\psi\rangle\langle\psi|\, - |\,\phi\rangle\langle\phi|\,\|_1 = 2\sqrt{1 - |\,\langle\psi|\,\phi\rangle|^2}. \tag{4.23}$$

Доказательство. Равенство (4.23) получается аналогично соотношению (4.19)). Для вычисления следовой нормы достаточно найти собственные значения оператора $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$, который имеет ранг 2.

Обозначая t_j собственные числа оператора $T = |\psi\rangle\langle\psi| - S$, имеем

$$\|| \,\psi\rangle\langle\psi| - S\|_1 = 2 \sum_{j:t_j>0} t_j \geqslant 2\langle\psi|T|\,\psi\rangle = 2\,[1 - \langle\psi|\,S|\,\psi\rangle]\,,$$

т.е. получаем первое неравенство.

Пусть S – произвольный оператор плотности; рассмотрим его спектральное разложение $S = \sum \lambda_j |e_j\rangle\langle e_j|$. Используя выпуклость нормы, вогнутость функции $\sqrt{\ }$ и соотношение (4.23), получаем

$$\begin{aligned}\|| \,\psi\rangle\langle\psi| - S\|_1 &\leqslant \sum_j \lambda_j \|| \,\psi\rangle\langle\psi| - |e_j\rangle\langle e_j|\|_1 \\ &= 2\sum_j \lambda_j \sqrt{1 - \langle\psi|e_j\rangle\langle e_j|\,\psi\rangle} \leqslant 2\sqrt{1 - \langle\psi|\,S|\,\psi\rangle}.\end{aligned}$$

□

Для чистого состояния $|\,\psi\rangle\langle\psi|$ и произвольного состояния S величина

$$F(|\psi\rangle\langle\psi|, S) = \sqrt{\langle\psi|\,S|\,\psi\rangle} \tag{4.24}$$

называется *степенью совпадения* (*fidelity*) (состояний $|\psi\rangle\langle\psi|$ и S). Очевидно, что $F \leqslant 1$, причем из доказанной леммы следует, что равенство имеет место тогда и только тогда, когда $S = |\psi\rangle\langle\psi|$.

Степень совпадения произвольных состояний S_0, S_1 определяется соотношением

$$F(S_0; S_1) = \mathrm{Tr}\,\sqrt{\sqrt{S_0}S_1\sqrt{S_0}}.$$

Если $S_0 = |\,\psi\rangle\langle\psi|$, то $\sqrt{S_0} = |\,\psi\rangle\langle\psi|$, и мы приходим к выражению (4.24). Неравенства (4.22) обобщаются следующим образом (см. [8], п. 9.2.3)

$$2[1 - F(S_0; S_1)] \leqslant \|\,S_0 - S_1\|_1 \leqslant 2\sqrt{1 - F(S_0; S_1)^2}.$$

По сравнению со следовой нормой более простой является норма Гильберта–Шмидта

$$\|\,S_0 - S_1\|_2 = \sqrt{\mathrm{Tr}(S_0 - S_1)^2},$$

вычисление которой не требует знания собственных чисел t_j. Соотношение между нею и следовой нормой имеет вид

$$\|\,S_0 - S_1\|_2 \leqslant \|\,S_0 - S_1\|_1 \leqslant \sqrt{d}\|\,S_0 - S_1\|_2,$$

что вытекает из выражения

$$\|\,S_0 - S_1\|_2 = \sqrt{\sum_j t_j^2}$$

и неравенства Коши–Буняковского.

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 $\mathcal{X} = \{1, \dots, |\mathcal{X}|\}$, 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 69. Пусть $\mathcal{X} = \{0, 1\}$, тогда всякое распределение на $\mathcal{X}$ имеет вид $\{p, 1-p\}$, $0 \leqslant p \leqslant 1$ (случайный бит). Докажите, что двоичная энтропия

$$h(p) = -p \log p - (1-p) \log(1-p), \tag{5.2}$$

является неотрицательной вогнутой функцией p, которая обращается в 0 при $p = 0, 1$ и принимает максимальное значение, равное 1, при $p = 1/2$. Постройте график этой функции.

Задача 70. Используя вогнутость функции $\log x$, покажите, что $0 \leqslant H(X) \leqslant \log |\mathcal{X}|$, 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 $w = (x_1, \dots, x_n)$ 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 15. *Если $X_1, \dots, X_n, \dots$ независимые и одинаково распределенные случайные величины с распределением $p = \{p_x\}$, то*

$$-\frac{1}{n}\sum_{i=1}^{n} \log p_{X_i} \longrightarrow H(X) \quad \text{по вероятности.} \tag{5.3}$$

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

$$\mathsf{P}\left\{\left| -\frac{1}{n}\sum_{i=1}^{n} \log p_{X_i} - H(X)\right| < \delta\right\} > 1 - \epsilon. \tag{5.4}$$

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

$$p_w = p_{x_1} \cdot \ldots \cdot p_{x_n} = 2^{-n\left(-\frac{1}{n}\sum_{i=1}^{n} \log p_{x_i}\right)} \tag{5.5}$$

мы теперь можем использовать соотношение (5.4), чтобы ввести понятие *типичного слова*: слово w длины n, имеющее вероятность p_w, называется δ-типичным, если

$$2^{-n(H(X)+\delta)} < p_w < 2^{-n(H(X)-\delta)}. \tag{5.6}$$

Поскольку эти неравенства равносильны неравенствам под знаком вероятности в соотношении (5.4), то левая часть в этом соотношении равна вероятности множества всех δ-типичных слов.

Отсюда вытекают следующие свойства типичных слов:

1. Существует не более $2^{n(H(X)+\delta)}$ типичных слов.
2. Для $\epsilon > 0$ и достаточно больших n существует, по крайней мере, $(1-\epsilon)2^{n(H(X)-\delta)}$ типичных слов.
3. При этом множество не-типичных слов имеет вероятность $\leqslant \epsilon$.

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.

Задача 71. Докажите обратное утверждение: любой код, использующий двоичные последовательности длины $n(H(X)-\delta)$, имеет асимптотически неисчезающую вероятность ошибки, стремящуюся к единице при $n \to \infty$.

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*:

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

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:

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

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

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y|X), \tag{5.8}$$

где в последней формуле уже $H(Y)$ может быть интерпретирована, как информационное содержание выхода, а $H(Y|X)$ как его бесполезная составляющая, обусловленная *шумом*.

Задача 72. Покажите, что взаимная информация всегда неотрицательна: тот факт, что $H(X) \geqslant H(X|Y)$ вытекает из вогнутости функции $-x\log x$.

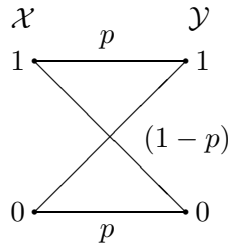

Рис. 5.1. Двоичный симметричный канал

Отсюда также следует свойство субаддитивности энтропии:

$$H(X,Y) \leqslant H(X) + H(Y).$$

Далее, $I(X;Y) = 0$ (т.е. $H(X,Y) = H(X) + H(Y)$) тогда и только тогда, когда X и Y – независимые случайные величины: $p_{x,y} = p_x \cdot p_y$.

Шенноновская пропускная способность определяется как

$$C = \max_{\{p_x\}} I(X;Y) = \max_{\{p_x\}} [H(Y) - H(Y|X)], \tag{5.9}$$

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. Используя двоичную энтропию (5.2), взаимную информацию можно записать как $I(X;Y) = H(Y) - H(Y|X) = H(Y) - h(p)$. 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 $x_1, x_2, \ldots,$ 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 канала без памяти со входом X^n. Очевидно, что последовательность

$$C_n = \max_{X^n} I(X^n;Y^n)$$

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

Лемма 5.

$C_n = nC$.

Доказательство. Покажем, что

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

Имеет место *цепное правило* для условной энтропии:

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

которое легко доказать по индукции, используя формулу:

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

Тогда взаимная информация

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

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

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

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 получено значение $y^n\in V^{(j)}$, $j=1,\dots,N$, то принимается решение, что было послано слово $w^{(j)}$; если же получено $y^n\in V^{(0)}$, то никакого определенного решения не принимается. Таким образом, *максимальная вероятность ошибки* такого кода есть

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

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

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

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

Лемма 6. *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 $p(V^{(j)}|w^{(j)})\geqslant 2\epsilon$, так что построить требуемый N-подкод невозможно. Тогда средняя ошибка $2N$-кода ограничена снизу величиной $\overline{P}_e(W,V)\geqslant\frac{1}{2N}2\epsilon(N+1)>\epsilon$, что противоречит предположению. □

Теорема 16 (Теорема кодирования для канала без памяти). *Пусть*

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

минимальная средняя ошибка для всевозможных кодов размера N 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 7 (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*

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

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

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

где был использован тот факт, что $H(X|E = 0, Y)$ также равна 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 свойству аддитивности и неравенству Фано

$$\begin{aligned} nC = C_n \geqslant I(Z; Y^n) = H(Z) - H(Z|Y^n) \geqslant \\ \geqslant \log N - 1 - \underbrace{P\{\widehat{Z}(Y^n) \neq Z\}}_{= \overline{P}_e(W, V)} \log N. \end{aligned} \tag{5.19}$$

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

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

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

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

□

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 друг от друга 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

$$N \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}. \tag{5.20}$$

Таким образом, $N \approx 2^{nC}$. Конечно, это рассуждение в высшей степени эвристично; строгое доказательство, реализующее эту идею, можно найти, например, в [14], [15].

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 усредненной степенью совпадения (ср. (4.24))

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

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:

$$H(S) = -\sum_j s_j \log s_j = -\operatorname{Tr} S \log S. \tag{5.21}$$

Далее нам понадобятся элементарные свойства квантовой энтропии:

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

1) $0 \leqslant H(S) \leqslant \log d$, 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).

Задача 73. Докажите свойства 1)-3), используя соответствующие свойства классической энтропии Шеннона.

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

Теорема 17 [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* $\mathcal{H}_d$ *размерности* $d \leqslant 2^{n(H(\overline{S}_p)-\delta)}$ *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 чистых состояний, поскольку $2^{nH(\overline{S}_p)}$ есть критическое значение размерности гильбертова пространства. (Напомним классический результат: пусть имеется источник, посылающий символы $1, \ldots, a$ с вероятностями $p_1, \ldots, p_a$, тогда количество слов, асимптотически безошибочно пересылаемых источником, есть $N \sim 2^{nH(p)}$, где $H(p) = -\sum_x p_x \log p_x$).

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

$$\overline{S}_p = \sum_j \lambda_j |e_j\rangle\langle e_j|. \tag{5.22}$$

Пусть $J = (j_1, \ldots, j_n)$, $\lambda_J = \lambda_{j_1} \ldots \lambda_{j_n}$, $|e_J\rangle = |e_{j_1}\rangle \otimes \cdots \otimes |e_{j_n}\rangle$, 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

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

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

$$\dim E\mathcal{H}^{\otimes n} = \operatorname{Tr} E \leqslant \operatorname{Tr} \frac{\overline{S}_p^{\otimes n}}{2^{-n(H(\overline{S}_p)+\delta)}} \leqslant 2^{n(H(\overline{S}_p)+\delta)}. \tag{5.23}$$

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

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

$$\begin{aligned} F_n &= \sum_w p_w \langle\psi_w|S_w|\psi_w\rangle = \sum_w p_w \langle\psi_w|E|\psi_w\rangle \\ &= \operatorname{Tr} E(\sum_w p_w|\psi_w\rangle\langle\psi_w|) = \operatorname{Tr} E S_p^{\otimes n} = \sum_{J\in J_{n,\delta}} \lambda_J. \end{aligned} \tag{5.24}$$

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

$$\begin{aligned} &\mathsf{P}\{2^{-n(H(\overline{S}_p)+\delta))} < \lambda_J < 2^{-n(H(\overline{S}_p)-\delta))}\} = \\ &\qquad = \mathsf{P}\left\{H(\overline{S}_p) - \delta < -\frac{1}{n}\sum_{k=1}^n \log\lambda_{j_k} < H(\overline{S}_p) + \delta\right\} \\ &= \mathsf{P}\left\{\left|-\frac{1}{n}\sum_{k=1}^n \log\lambda_{j_k} - H(\overline{S}_p)\right| < \delta\right\}, \end{aligned} \tag{5.25}$$

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 $\delta/2$. Фиксируем $\varepsilon > 0$. Тогда, согласно свойствам типичного подпространства, для достаточно больших n правая часть оценивается как

$$\begin{aligned} \operatorname{Tr}\overline{S}_p^{\otimes n} E P_d + \operatorname{Tr}\overline{S}_p^{\otimes n}(1-E)P_d &\leqslant \|\overline{S}_p^{\otimes n}E\|\ \operatorname{Tr} P_d + \operatorname{Tr}\overline{S}_p^{\otimes n}(1-E) \leqslant \\ &\leqslant d2^{-n(H(\overline{S}_p)-\delta/2)} + \frac{\varepsilon}{2} \leqslant 2^{-n\delta/2} + \frac{\varepsilon}{2} < \varepsilon. \end{aligned} \tag{5.26}$$

□

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 $w = \{x_1, \dots, x_n\}$, которому сопоставляется состояние S_w:

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

Предположение о том, что w кодируется в тензорное произведение состояний S_{x_j}, соответствует понятию 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

$$I_1(p, M) = \sum_x p_x \sum_y p_M(y|x) \left[\log p_M(y|x) - \log \sum_z p_M(y|z) p_z \right], \tag{5.27}$$

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

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

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

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 $W = \{w^{(1)}, \dots, w^{(N)}\}$ вместе с разложением единицы $M = \{M_j\}$ в $\mathcal{H}^{\otimes n}$ с исходами $j = 0, 1, \dots, N$; исход 0 означает уклонение от принятия решения.

Средняя ошибка кода равна

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

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

$$C_\chi = \max_p \left\{ H\left(\sum_x p_x S_x\right) - \sum_x p_x H(S_x)\right\}, \tag{5.28}$$

где $H(S)$ – энтропия фон Неймана (5.21).

Теорема 18 (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 16) вытекает, что утверждение теоремы 18 выполняется с заменой C_χ на C_∞. Таким образом, утверждение теоремы 18 состоит в том, что $C_\infty = C_\chi$.

Если состояния $S_x = |\psi_x\rangle\langle\psi_x|$ чистые, то из (5.28) вытекает

$$C_\chi = \max_p H\left(\sum_x p_x |\psi_x\rangle\langle\psi_x|\right). \tag{5.29}$$

Из свойства 1) энтропии следует, что всегда

$$C_\chi \leqslant \log d. \tag{5.30}$$

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

[3] A. S. Holevo, "The capacity of quantum channel with general signal states", IEEE Trans. Inform. Theory, 1998, v. 44, No. 1, 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.

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 (4.20) из трех равновероятных "равноугольных" векторов ψ_1, ψ_2, ψ_3. Тогда

$$\sum_{x=1}^{3} p_x |\psi_x\rangle\langle\psi_x| = \frac{1}{2} I$$

и, как следует из (5.29), 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 $p_1 = p_2 = 1/2$, $p_3 = 0$ и соответствующего оптимального измерения для двух равновероятных состояний S_1, S_2 (см. пример 1 в разделе 4.3).

5.4. Квантовая граница классической информации и доказательство обратной теоремы

Теорема 19 (Квантовая граница классической информации). *Для любого распределения p и любой наблюдаемой M*

$$I_1(p, M) \leqslant H\Big(\sum p_x S_x\Big) - \sum p_x H(S_x), \tag{5.31}$$

где $I_1(p, M)$ – количество информации в (5.27), *причем имеет место строгое неравенство, если среди операторов $p_x S_x$ есть некоммутирующие между собой.*

Задача 74. Если все операторы $p_x S_x$ коммутируют, то равенство достигается для наблюдаемой $M = \{|e_k\rangle\langle e_k|\}$, где $\{e_k\}$ – о.н.б. из общих собственных векторов операторов $p_x S_x$.

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 (5.31).

Отметим также, что очевидным следствием неравенства (5.31) является вогнутость квантовой энтропии как функции операторов плотности:

$$H\Big(\sum p_x S_x\Big) \geqslant \sum p_x H(S_x). \tag{5.32}$$

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

$$\chi(t) = H(S_t) - (1 - t)H(S_0) - tH(S_1), \quad t \in [0, 1]. \tag{5.33}$$

[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.

[5] А. С. Холево, "Некоторые оценки для количества информации, передаваемого квантовым каналом связи", Пробл. передачи информ., 1973, т. 9, № 3, 3–11.

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

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

где $p = \{1-t, t\}$. Заметим, что

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

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

$$\chi(t)'' \leqslant I_M(t)'', \quad t \in [0,1]. \tag{5.34}$$

Отсюда, очевидно, следует

$$\chi(t) \geqslant I_M(t), \quad t \in [0,1]. \tag{5.35}$$

Положим $D = S_1 - S_0$ и пусть

$$S_t = \sum_k s_k E_k$$

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

Лемма 8.

$$\chi''(t) = -\sum_{k,j} (\operatorname{Tr} E_k D E_j D) f(s_k, s_j), \quad t > 0, \tag{5.36}$$

где

$$f(a,b) = \frac{\log a - \log b}{a-b}, \quad a \neq b, \quad f(a,a) = a^{-1}. \tag{5.37}$$

Используя элементарное неравенство

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

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

$$\chi''(t) \leqslant -\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j}, \tag{5.38}$$

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 75. Покажите, что оператор

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

является решением уравнения

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

[6] ibid.

причем

$$\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j} = \operatorname{Tr} D L_t = \operatorname{Tr} S_t L_t^2. \tag{5.39}$$

Оператор L_t является некоммутативным аналогом логарифмической производной семейства S_t, а (5.39) – аналогом информационного количества Фишера в математической статистике.

Из (5.38), (5.39) вытекает, что

$$\chi''(t) \leqslant -\operatorname{Tr} S_t L_t^2, \tag{5.40}$$

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

Пусть теперь $M = \{M_y\}$ – произвольная наблюдаемая, $P_t(y) = \operatorname{Tr} S_t M_y = (1-t)P_0(y) + tP_1(y)$ – ее распределение в состоянии S_t. Положим также $D(y) = P_1(y) - P_0(y) = \operatorname{Tr} D M_y$. Применяя полученные результаты к диагональной матрице $\operatorname{diag}[P_t(y)]$ в роли состояния S_t и учитывая коммутативность диагональных матриц, получаем вместо (5.40)

$$I_M''(t) = -\sum_y \frac{D(y)^2}{P_t(y)}. \tag{5.41}$$

Доказательство неравенства (5.34). Имеем

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

где $A = \sqrt{S_t}\sqrt{M_y}$, $B = \sqrt{S_t} L_t \sqrt{M_y}$. В силу некоммутативного неравенства Коши–Буняковского (1.52) получаем $D(y)^2 \leqslant \operatorname{Tr} S_t M_y \cdot \operatorname{Tr} L_t S_t L_t M_y$. Подставляя в (5.41), имеем

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

причем при $[S_0, S_1] \neq 0$ 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 $W = \{w^{(1)}, \ldots, w^{(N)}\}$ 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: $H(X|Y) \leqslant 1 + \mathsf{P}(X \neq Y)\log N$. Таким образом, $\max I(X,Y) \geqslant nR(1 - p_e(n, 2^{nR})) - 1$ (это повторение доказательства слабого обращения классической теоремы Шеннона).

[7] ibid.

Из (5.31) вытекает неравенство

$$I(X;Y) \leqslant \max_p \biggl[H\Bigl(\sum_w p_w S_w - \sum_w p_w H(S_w) \Bigr) \biggr] = C_\chi^{(n)}.$$

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

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

$$\max_{p_{ij}} \biggl[H\Bigl(\sum_{ij} p_{ij} S_i^1 \otimes S_j^2 \Bigr) - \sum_{ij} p_{ij} H(S_i^1 \otimes S_j^2) \biggr] = \max_{p_i^1} [\dots] + \max_{p_j^2} [\dots].$$

Очевидно, что $\max\limits_{p_{ij}} \geqslant \max\limits_{p_{ij} = p_i^1 p_j^2}$, тогда в силу свойства аддитивности квантовой энтропии

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

откуда $C_\chi^{(n)} \geqslant nC_\chi$.

Обратное неравенство вытекает из свойства субаддитивности квантовой энтропии (см., например, [8], п. 11.3.4), из которого вытекает

$$H\Bigl(\sum p_{ij} S_i^1 \otimes S_j^2 \Bigr) \leqslant H\Bigl(\sum p_i^1 S_i^1 \Bigr) + H\Bigl(\sum p_i^2 S_i^2 \Bigr),$$

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

Окончательно, $nC_\chi \geqslant nR[1 - p_e(n, 2^{nR})] - 1$, т.е. $p_e(n, 2^{nR}) \geqslant 1 - C_\chi/R - 1/nR$, 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 случае чистых состояний[8] $S_x = |\psi_x\rangle\langle\psi_x|$, когда

$$C_\chi = \max_p H\Bigl(\sum_x p_x S_x \Bigr).$$

Доказательство непросто уже в этом случае, тогда как классический аналог этой проблемы тривиален, поскольку различные чистые состояния с необходимостью ортогональны. Доказательство в общем случае см. в [14].

Доказательство. Пусть $R < C_\chi$. Докажем, что $p_e(n, 2^{nR}) \to 0$. Рассмотрим среднюю вероятность ошибки кода

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

[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.

которая зависит от выбора слов $W = \{w^{(j)}\}$ 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

$$|\widetilde{\psi}_{w^{(j)}}\rangle = E|\psi_{w^{(j)}}\rangle \in \mathcal{H}^{n,\delta} \tag{5.42}$$

и пусть $G = \sum\limits_{j=1}^{N} |\widetilde{\psi}_{w^{(j)}}\rangle\langle\widetilde{\psi}_{w^{(j)}}|$ – оператор Грама системы (5.42). Оператор Грама всегда можно обратить на подпространстве $\mathcal{H}^{n,\delta}$. Обозначим $G^{-1/2}$ корень из обобщенного обратного к G, равного 0 на ортогональном дополнении к $\mathcal{H}^{n,\delta}$. Для данного набора кодовых слов W введем наблюдаемую M, обобщающую "square-root measurement" (4.8):

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

Тогда средняя ошибка кода (W, M) равна

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

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

$$\begin{aligned}\overline{P}_e(W; M) &\leqslant \frac{2}{N}\sum_{j=1}^{N}\bigl(1 - \langle\widetilde{\psi}_{w^{(j)}}|G^{-1/2}|\widetilde{\psi}_{w^{(j)}}\rangle\bigr)\\ &= \frac{1}{N}\sum_{j=1}^{N}\bigl(2\operatorname{Tr} S_{w^{(j)}} - 2\operatorname{Tr} S_{w^{(j)}}G^{-1/2}\bigr).\end{aligned} \tag{5.43}$$

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

$$-2x^{-1/2} \leqslant -3 + x, \quad x \geqslant 0,$$

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

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

Подставляя в (5.43), получаем

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

Учитывая, что

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

и

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

получаем окончательную оценку

$$\overline{P}_e(W;M) \leqslant \frac{1}{N}\sum_{j=1}^{N}\left[2\operatorname{Tr} S_{w^{(j)}}(I-E) + \sum_{k\neq j}\operatorname{Tr} ES_{w^{(j)}}ES_{w^{(k)}}E\right]. \tag{5.44}$$

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

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

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

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

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

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

$$\begin{aligned}\ll\overline{P}_e(W,M)\gg \;&\leqslant 2\operatorname{Tr}\overline{S}_p^{\otimes n}(I-E) + (N-1)\operatorname{Tr}(E\overline{S}_p^{\otimes n}E)^2\\ &\leqslant 2\operatorname{Tr}\overline{S}_p^{\otimes n}(I-E) + (N-1)\|\overline{S}_p^{\otimes n}E\|.\end{aligned}$$

Согласно свойствам типичного подпространства для достаточно больших n

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

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

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

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

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:

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

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

1) $\Phi\left[\sum c_j T_j\right] = \sum c_j \Phi[T_j]$ (*линейность*);

2) $T \geqslant 0 \Rightarrow \Phi[T] \geqslant 0$ (*положительность*);

3) $\operatorname{Tr} \Phi[T] = \operatorname{Tr} T$ (*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 [14].

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

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

имеет место

$$[\Phi[X_{ij}]]_{i,j=1,\ldots,m} \geqslant 0.$$

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

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

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

Тогда условие полной положительности можно переформулировать в виде:

Для любого $m = 1, 2, \ldots$ 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:

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

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

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

Задача 76. 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 $\Phi \otimes \mathrm{Id}$:

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

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, $\mathcal{H}_E$ – пространство окружения E. Эволюция составной системы является обратимой и описывается унитарным оператором U.

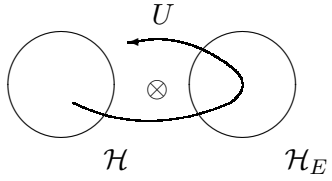

Рис. 6.1. Открытая квантовая система

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

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

Задача 77. Докажите, что отображение (6.3) является квантовым каналом. Указание:

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

Задача 78. 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 полной положительности 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 нами ранее.

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 *– фиксированные 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

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

где $\{e_j\}$ – о.н.б. в $\mathcal{H}'$, $\{M_j\}$ – 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 79. Докажите полную положительность отображений (6.4), (6.5), получив для них представление Крауса (см. следующий раздел).

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 приводит к следующему представлению:

Теорема 20 (Представление Крауса). *Линейное отображение* $\Phi\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ *является каналом, т.е. 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 $D \leqslant dd'$, *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. Пусть $\{e_j\}$ – о.н.б. в $\mathcal{H}$. Рассмотрим "максимально сцепленный" вектор

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

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

$$S_{12}=(\Phi\otimes\mathrm{Id})\,[|\psi_{12}\rangle\langle\psi_{12}|]=\sum_{j,k=1}^{d}\Phi[|e_j\rangle\langle e_k|]\otimes|e_j\rangle\langle e_k| \tag{6.8}$$

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

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

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

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

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

получаем

$$\begin{aligned}\langle\psi'|\Phi[X]|\psi'\rangle&=\sum_{j,k=1}^{d}\langle e_j|X|e_k\rangle\langle\psi'|\Phi[|e_j\rangle\langle e_k|]|\psi'\rangle=\sum_{j,k=1}^{d}\langle e_j|X|e_k\rangle\langle\psi'\otimes e_j|S_{12}|\psi'\otimes e_k\rangle\\&=\sum_{j,k=1}^{d}\langle e_j|X|e_k\rangle\sum_{l=1}^{D}\langle\psi'\otimes e_j|\Psi_l\rangle\langle\Psi_l|\psi'\otimes e_k\rangle=\sum_{j,k=1}^{d}\langle e_j|X|e_k\rangle\sum_{l=1}^{D}\langle\psi'|V_l|e_j\rangle\langle e_k|V_l^*|\psi'\rangle,\end{aligned}$$

то есть

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

Используя сохранение следа, получаем

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

для любого оператора X, откуда следует (6.7).

Доказательство того факта, что отображение (6.6) является каналом, несложно и предоставляется в качестве простого упражнения. □

Оператор (6.8), называемый *оператором Чоя–Ямилковского*, позволяет однозначно восстановить канал Φ:

Задача 80. Докажите формулу обращения

$$\Phi[S]=\operatorname{Tr}_2 S_{12}(I_1\otimes S^{\top}), \tag{6.9}$$

где $S^{\top}$ – оператор, имеющий в базисе $\{e_j\}$ матрицу, транспонированную к матрице оператора S.

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

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

$$\mathcal{K}=\mathcal{H}\otimes\mathcal{H}_E\approx\underbrace{\mathcal{H}\oplus\cdots\oplus\mathcal{H}}_{D},$$

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

$$V=\begin{bmatrix}V_1\\ \vdots\\ V_D\end{bmatrix}. \tag{6.10}$$

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:

$$U=\begin{bmatrix}V_1 & \dots & \dots\\ \vdots & \vdots & \vdots\\ V_D & \dots & \dots\end{bmatrix}. \tag{6.11}$$

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

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

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

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

так что начальное состояние системы+окружения есть

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

Оператор унитарной эволюции преобразует его в

$$U\big(S\otimes|\psi_E\rangle\langle\psi_E|\big)U^*=\begin{bmatrix}V_1SV_1^* & \dots & \dots\\ \vdots & \ddots & \vdots\\ \dots & \dots & V_DSV_D^*\end{bmatrix}.$$

Беря частичный след по окружению, получаем

$$\operatorname{Tr}_E U\big(S\otimes|\psi_E\rangle\langle\psi_E|\big)U^*=\sum_{l=1}^{D}V_lSV_l^*=\Phi[S].$$

□

Найдем еще, как изменится состояние окружения в результате взаимодействия с системой. Беря частичный след по системе, получаем

$$S'_E = \mathrm{Tr}_{\mathcal{H}}\, U\big(S \otimes |\psi_E\rangle\langle\psi_E|\big)U^* = \mathrm{Tr}_{\mathcal{H}}\, VSV^* = [\mathrm{Tr}\; V_k S V_l^*]_{k,l=1,\ldots,D}\,, \tag{6.12}$$

где V – оператор (6.10), а частичный след вычислен по правилу (2.5).

6.3. Q-битные каналы

Всякий канал Φ в пространстве $\mathbb{C}^2$ является аффинным преобразованием множества состояний q-бита и поэтому 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 себя, и которое удовлетворяют некоторым дополнительным ограничениям, вытекающим из условия полной положительности. При этом имеем

$$\Phi[S(\vec{a})] = S(T\vec{a} + \vec{b}). \tag{6.13}$$

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 (6.13), 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), при этом центр шара остается на месте. Это имеет место тогда и только тогда, когда канал Φ *бистохастический*, т.е. переводит хаотическое состояние в хаотическое.

Задача 81. Покажите, что в этом случае

$$\Phi[S] = \sum_{\gamma=0,x,y,z} \mu_\gamma \sigma_\gamma S \sigma_\gamma, \tag{6.14}$$

где

$$\begin{aligned} \mu_0 &= \frac{1}{4}\left(1 + \lambda_x + \lambda_y + \lambda_z\right), \qquad \mu_x = \frac{1}{4}\left(1 + \lambda_x - \lambda_y - \lambda_z\right), \\ \mu_y &= \frac{1}{4}\left(1 - \lambda_x + \lambda_y - \lambda_z\right), \qquad \mu_z = \frac{1}{4}\left(1 - \lambda_x - \lambda_y + \lambda_z\right), \end{aligned} \tag{6.15}$$

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

[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.*

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 (6.15) 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.* Докажите тождество

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

Из него следует, что *деполяризующий канал*

$$\Phi[S] = (1-p)S + \frac{p}{2} I \operatorname{Tr} S \tag{6.16}$$

допускает представление Крауса

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

Из уравнений (6.15) находим $\lambda_x = \lambda_y = \lambda_z = 1-p$, что соответствует равномерному сжатию шара Блоха в $|1-p|$ раз. При $p \leqslant 1$ формула (6.16) описывает смесь (с вероятностями $1-p$ и p) идеального канала и "полностью деполяризующего" канала, который отображает любое состояние в хаотическое. Однако отображение Φ остается вполне положительным и при $1 < p \leqslant 4/3$.

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) [8].

Преобразование матрицы плотности имеет вид (6.13), где

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

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. Процессы квантовых измерений

Важнейший пример необратимой эволюции – изменение состояния квантовой системы в результате производимого над ней измерения. Рассмотрим ортогональное разложение единицы $E = \{E_x\}$,

$$E_x E_{x'} = \delta_{xx'} E_x, \quad \sum_x E_x = I,$$

другими словами, четкую наблюдаемую в пространстве $\mathcal{H}$. Согласно проекционному постулату фон Неймана–Людерса изменение состояния при идеальном (неселективном) измерении этой наблюдаемой описывается соотношением (1.58):

$$S' = \sum_x E_x S E_x = \Phi[S].$$

Идеальное квантовое измерение удовлетворяет гипотезе воспроизводимости: при повторном измерении исход с вероятностью 1 равен исходу первого измерения (в предположении, что никаких изменений между измерениями не произошло). Большинство реальных процедур измерения не удовлетворяют этому ограничению, и мы переходим к описанию таких неидеальных измерений.

Рассмотрим следующий процесс *косвенного* измерения: система $\mathcal{H}$ в начальном состоянии S взаимодействует с некоторой пробной системой $\mathcal{H}_E$, находящейся в состоянии S_E, над которой затем производится идеальное измерение четкой наблюдаемой E^0. Тогда, в соответствии с постулатом фон Неймана-Людерса, вероятность исхода x равна

$$p_x = \operatorname{Tr} U(S \otimes S_E)U^*(I \otimes E^0_x),$$

а апостериорное состояние описывается оператором плотности S_x, удовлетворяющим соотношению

$$p_x S_x = \operatorname{Tr}_{\mathcal{H}_E} U(S \otimes S_E)U^*(I \otimes E^0_x) \equiv \Phi_x[S]. \tag{6.17}$$

Таким образом, состояние системы после измерения может быть записано как

$$\Phi[S] = \sum_x p_x S_x = \sum_x \Phi_x[S].$$

Отображения Φ_x являются вполне положительными, причем их сумма Φ сохраняет след. Любое такое семейство отображений $\{\Phi_x\}$ называется *инструментом*. Инструмент описывает статистику и апостериорные состояния неидеального измерения, которое в общем случае не обязано удовлетворять гипотезе воспроизводимости.

Заметим, что

$$p_x = \operatorname{Tr} S M_x, \quad \text{где} \quad M_x = \Phi^*_x[I]$$

– разложение единицы, описывающее наблюдаемую, связанную с данным косвенным измерением. Обратно, по данной наблюдаемой $M = \{M_x\}$ можно построить связанные с ней (неединственный) инструмент и косвенное измерение. Рассмотрим произвольное представление $M_x = V^*_x V_x$; построим по нему унитарный оператор (6.11) в пространстве $\mathcal{H} \otimes \mathcal{H}_E$, тогда семейство вполне положительных отображений

$$\Phi_x[S] = \operatorname{Tr}_{\mathcal{H}_E} U\big(S \otimes |\psi_E\rangle\langle\psi_E|\big)U^*\big(I \otimes |e_x\rangle\langle e_x|\big), \tag{6.18}$$

где $|e_x\rangle$ – базисный вектор-столбец с единицей на x-м месте, так что $|\psi_E\rangle = |e_1\rangle$, образует инструмент. При этом

$$M_x = \operatorname{Tr}_{\mathcal{H}_E}\big(I \otimes |\psi_E\rangle\langle\psi_E|\big)U^*\big(I \otimes |e_x\rangle\langle e_x|\big)U = \Phi^*_x[I]. \tag{6.19}$$

Таким образом, процесс косвенного измерения наблюдаемой M реализуется пробной системой $\mathcal{H}_E$ в начальном состоянии $S_E = |\psi_E\rangle\langle\psi_E|$, унитарным оператором U в пространстве $\mathcal{H} \otimes \mathcal{H}_E$, и четкой наблюдаемой $\{|e_x\rangle\langle e_x|\}$ в $\mathcal{H}_E$.

6.5. ПРОПУСКНЫЕ СПОСОБНОСТИ КВАНТОВОГО КАНАЛА

6.5.1. 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 монографиях [14], [16].

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.

6.5.2. 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

$$\Phi[S] = \mathrm{Tr}_{\mathcal{H}_E}\, U\,(S \otimes S_E)\, U^*. \tag{6.20}$$

С точки зрения 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 (6.16) описывает смесь идеального канала и полностью деполяризующего канала.

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

$$w \overset{\text{кодирование}}{\longrightarrow} S_w \overset{\text{канал}}{\longrightarrow} S'_w \overset{\text{декодирование}}{\longrightarrow} w' \tag{6.21}$$

Применение квантовой теоремы кодирования (теорема 18 в разделе 5.3) дает следующее выражение для классической пропускной способности канала Φ

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} C_\chi(\Phi^{\otimes n}), \tag{6.22}$$

где

$$C_\chi(\Phi) = \max_{p_i, S_i} \left\{ H\left(\sum_i p_i \Phi\left[S_i\right] \right) - \sum_i p_i H\left(\Phi\left[S_i\right]\right) \right\} \tag{6.23}$$

квантовый аналог шенноновской формулы (5.9). Если эта величина обладает свойством *аддитивности*, $C_\chi(\Phi^{\otimes n}) = nC_\chi(\Phi)$, то

$$C(\Phi) = C_\chi(\Phi).$$

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].

Вычислим величину $C_\chi(\Phi)$ для произвольного q-битного бистохастического канала.

Лемма 10. *Пусть Φ бистохастический канал в $\mathbb{C}^2$, тогда*

$$C_\chi(\Phi) = 1 - \min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S)). \tag{6.24}$$

Доказательство. Неравенство $\leqslant$ в (6.24) вытекает из того, что для любого канала

$$C_\chi(\Phi) \leqslant \log\dim\mathcal{H} - \min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S)), \tag{6.25}$$

так что остается доказать неравенство $\geqslant$. Поскольку энтропия – вогнутая функция состояния, минимум достигается на чистом состоянии S. Беря равновероятно чистые состояния $S_0 = S, S_1 = I - S$, получаем

$$C_\chi(\Phi) \geqslant H\left(\frac{1}{2}\Phi(I)\right) - \frac{1}{2}\big[H(\Phi(S)) + H(\Phi(I-S))\big].$$

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 (6.24). □

При вычислении $\min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S))$, в силу унитарной инвариантности энтропии достаточно рассмотреть случай Φ вида (6.14). Отметим также, что собственные значения оператора плотности (1.30) в $\mathcal{H}_2$ равны $(1 \pm |\vec{a}|)/2$, а значит, энтропия равна

$$H(S) = h\left(\frac{1-|\vec{a}|}{2}\right). \tag{6.26}$$

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

$$C_\chi(\Phi) = 1 - h\left(\frac{1-\max_\gamma|\lambda_\gamma|}{2}\right). \tag{6.27}$$

[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.

6.5.3. 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 сверхплотным кодированием, двукратен (см. п. 3.2). 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]

$$C_{ea}(\Phi) = \max_S I(S, \Phi), \tag{6.28}$$

где $I(S, \Phi)$ – *квантовая взаимная информация* между передатчиком и приемником, задаваемая соотношением

$$I(S, \Phi) = \{H(S) + H(\Phi[S]) - H(S; \Phi)\}. \tag{6.29}$$

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 (6.12). Таким образом,

$$H(S; \Phi) = H\left([\mathrm{Tr}\ V_k S V_l^*]_{k,l=1,\ldots,D}\right). \tag{6.30}$$

Для q-битного унитального канала $\Phi = \Lambda$ обменная энтропия $H(\bar{S}, \Phi)$, где $\bar{S}$ – хаотическое состояние, может быть вычислена с помощью представления Крауса (6.14). Используя свойства матриц Паули, получаем

$$H(\bar{S}, \Phi) = -\sum_{\gamma=0,x,y,z} \mu_\gamma \log \mu_\gamma.$$

Отсюда

$$C_{ea}(\Phi) = I(\bar{S}, \Phi) = 2\log 2 + \sum_{\gamma=0,x,y,z} \mu_\gamma \log \mu_\gamma,$$

где использован тот факт, что максимум взаимной информации в данном случае достигается на хаотическом состоянии.

Сравним это с пропускной способностью $C(\Phi) = C_\chi(\Phi)$, даваемой формулой (6.27), для деполяризующего канала (6.16). Для этого канала $\lambda_x = \lambda_y = \lambda_z = 1 - p$, $\mu_x = \mu_y = \mu_z = p/4$, $\mu_0 = (1 - 3p/4)$, откуда

$$C(\Phi) = 1 + (p/2)\log(p/2) + (1 - p/2)\log(1 - p/2),$$
$$C_{ea}(\Phi) = 2 + (1 - 3p/4)\log(1 - 3p/4) + (3p/4)\log(p/4).$$

При $p = 0$ получаем $C = 1$, $C_{ea} = 2$ (сверхплотное кодирование, п. 3.2).

Задача 82. Доказать, что в пределе сильного шума $p \to 1$, когда обе пропускные способности стремятся к нулю, $C_{ea}/C \to 3$ (использовать правило Лопиталя).

[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; Arxiv: quant-ph/0106052.

6.5.4. Квантовая пропускная способность. Преобразование квантового состояния $S \to \Phi[S]$ можно рассматривать как передачу квантовой информации. Открытие *квантовых кодов*, *исправляющих ошибки* (раздел 3.5) поставило вопрос об асимптотически (при $n \to \infty$) безошибочной передаче квантовых состояний каналом $\Phi^{\otimes n}$. При этом *квантовая пропускная способность* $Q(\Phi)$ определяется как максимальное количество передаваемой квантовой информации и связана с размерностью подпространства векторов 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,

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_{S^{(n)}} I_c\big(S^{(n)}, \Phi^{\otimes n}\big). \tag{6.31}$$

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.

6.5.5. 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. [8].

Три пропускные способности (6.22), (6.28), (6.31) связаны соотношением $Q(\Phi) \leqslant C(\Phi) \leqslant C_{ea}(\Phi)$ 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

$$\begin{array}{ccccccc} C_\chi & \leqslant & C_{\leftarrow} & \leqslant & C_{\leftrightarrow} & \leqslant & C_{ea} \\ \text{VI} & & \text{VI} & & \text{VI} & & \text{VI} \\ Q & \leqslant & Q_{\leftarrow} & \leqslant & Q_{\leftrightarrow} & \leqslant & Q_{ea} \end{array},$$

где $\leqslant$ 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.

[6] I. Devetak, "The private classical information capacity and quantum information capacity of a quantum channel", e-print no. quant-ph/0304127, 2003.

[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:

- Количественные характеристики сцепленности. Многочастичная сцепленность;
- Алгоритмы сжатия квантовой информации;
- Быстрые квантовые алгоритмы и квантовое моделирование. Адиабатические вычисления. Сложность квантовых вычислений;
- Квантовые коды, исправляющие ошибки. Вычисления, устойчивые к ошибкам;
- Квантовая криптография;
- Квантовая теория оценивания состояний. Квантовая метрология. Томография квантовых состояний;
- Квантовая термодинамика.

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 Processing; International Journal of Quantum Information; Quantum Information and Computation. Эта тематика представлена в таких известных журналах как Physical Reviews; Physical Reviews Letters; IEEE Transactions on Information Theory; Communications on Mathematical Physics; Journal of Mathematical Physics; Проблемы передачи информации, появилась и монографическая литература, см. библиографию в [14], [16].

[1] A. S. Holevo, V. Giovannetti, "Quantum channels and their entropic characteristics", Rep. Prog. Phys., 75 (2012), 046001.

[2] C. M. Caves, "Quantum Information Science: Emerging No More", arXiv:1302.1864.

Ответы и указания к решению задач

1. Подставляя в тождество $A = I\,A\,I$ разложение единичного оператора (1.3), получаем

$$A = \sum_{j,k=1}^{d} |e_k\rangle\langle e_k|A|e_j\rangle\langle e_j| = \sum_{j,k=1}^{d} a_{kj}|e_k\rangle\langle e_j|, \tag{7.1}$$

где $a_{kj} = \langle e_k|A|e_j\rangle$ – элементы матрицы оператора A в базисе $\{|e_j\rangle\}$, поскольку

$$A|\psi\rangle = \sum_{k=1}^{d} |e_k\rangle \sum_{j=1}^{d} a_{kj}\langle e_j|\psi\rangle.$$

3. Произвольная эрмитова $d \times d$-матрица имеет d свободных вещественных параметров на диагонали и $\frac{d(d-1)}{2}$ свободных комплексных параметров выше диагонали, причем каждый из них сводится к двум вещественным параметрам. Итого эрмитова $d \times d$-матрица имеет $d + 2\frac{d(d-1)}{2} = d^2$ свободных вещественных параметров.

5. Если $A = B^*B$, то $\langle\psi|A|\psi\rangle = \|B\psi\|^2 \geqslant 0$ для всех ψ, т.е. $A \geqslant 0$. Обратно, пусть $A \geqslant 0$, тогда имеем спектральное разложение (1.7), где $a_j \geqslant 0$. Полагая $B = \sum_{j=1}^{d} \sqrt{a_j}|e_j\rangle\langle e_j|$, имеем $B^* = B \geqslant 0$ и $B^2 = A$.

6. Инвариантность следа относительно выбора базиса. Пусть $\{h_l\}$ – другой о.н.б., тогда, подставляя первое равенство из разложения (7.1), и пользуясь соотношением полноты для нового базиса, получаем

$$\sum_{l=1}^{d}\langle h_l|A|h_l\rangle = \sum_{l=1}^{d}\sum_{j,k=1}^{d} \langle h_l|e_k\rangle\langle e_k|A|e_j\rangle\langle e_j|h_l\rangle$$
$$= \sum_{j,k=1}^{d} \langle e_k|A|e_j\rangle \sum_{l=1}^{d} \langle e_j|h_l\rangle\langle h_l|e_k\rangle = \sum_{j,k=1}^{d} \langle e_k|A|e_j\rangle\langle e_j|e_k\rangle = \sum_{j=1}^{d}\langle e_j|A|e_j\rangle.$$

Тождество (1.15) вытекает из симметрии выражения

$$\operatorname{Tr} AB = \sum_{j,k=1}^{d} \langle e_k|A|e_j\rangle\langle e_j|B|e_k\rangle$$

по отношению к индексам j, k.

Докажем свойство (1.16). Выбирая о.н.б., для которого первый вектор $|e_1\rangle = |\varphi\rangle/\|\varphi\|$, имеем

$$\operatorname{Tr} A|\psi\rangle\langle\varphi| = \langle e_1|A|\psi\rangle\langle\varphi||e_1\rangle = \langle\varphi|A|\psi\rangle.$$

Наконец, докажем (1.18). Рассмотрим спектральное разложение $A = \sum_{j=1}^{d} a_j|e_j\rangle\langle e_j|$. Тогда

$$\operatorname{Tr} AB = \operatorname{Tr} \sum_{j=1}^{d} a_j|e_j\rangle\langle e_j|\, B = \sum_{j=1}^{d} a_j\langle e_j|B|e_j\rangle.$$

Так как $A, B \geqslant 0$, то собственные значения a_j неотрицательны и $\langle e_j|B|e_j\rangle \geqslant 0$. Отсюда $\operatorname{Tr} AB \geqslant 0$. Если же $\operatorname{Tr} AB = 0$, то $\operatorname{Tr} C^*C = 0$, где $C = \sqrt{A}\sqrt{B}$. Следовательно, $C = 0$, а значит, $AB = \sqrt{A}C\sqrt{B} = 0$.

8. Ответ: $\mathsf{M}_S(X) = 1$, $\mathsf{D}_S(X) = 0$, $\mathsf{P}_S(X \models 1) = 1$.

9. Пусть S – точка компактного множества K, в которой непрерывная функция $\mathcal{F}$ достигает максимума (по теореме Вейерштрасса). По теореме 2 $S = \sum_j p_j S_j$, где S_j – крайние точки K. В силу выпуклости,

$$\mathcal{F}(S) \leqslant \sum_j p_j \mathcal{F}(S_j) \leqslant \max_j \mathcal{F}(S_j).$$

10. $P(S) = \sum_j s_j^2$, где s_j – собственные числа матрицы плотности S. Поскольку $\sum_j s_j = 1$, из неравенства Коши–Буняковского вытекает

$$1 = \left(\sum_{j=1}^d s_j\right)^2 \leqslant d\sum_{j=1}^d s_j^2,$$

причем равенство достигается тогда и только тогда, когда $s_j = 1/d$. С другой стороны, учитывая $0 \leqslant s_j \leqslant 1$, получаем $\sum_j s_j^2 \leqslant \sum_j s_j$, причем равенство достигается тогда и только тогда, когда одно из s_j равно 1.

Имеем $H(S) = \sum_{j=1}^d s_j \log \frac{1}{s_j}$, откуда, учитывая вогнутость функции log,

$$H(S) \leqslant \log\left(\sum_{j=1}^d s_j \frac{1}{s_j}\right) = \log d.$$

Для доказательства единственности точки максимума используйте строгую вогнутость функции log.

11. Используйте решение Задачи 3.

13. Ответы: $(0,0,1)$; $(0,0,-1)$; $(\pm 1,0,0)$; $(0,\pm 1,0)$.

15. Ответы:

$$\sqrt{\frac{1}{2}+\frac{1}{2\sqrt{2}}}\,|0\rangle + \sqrt{\frac{1}{2}-\frac{1}{2\sqrt{2}}}\,|1\rangle; \quad \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1-i}{2}\,|1\rangle; \quad \frac{1}{\sqrt{2}}\,|0\rangle + \frac{i}{\sqrt{2}}\,|1\rangle.$$

20. Усреднение по равномерному распределению вектора $\vec{a}$ на сфере с центром в 0 дает $\mathsf{M}a_x = \mathsf{M}a_y = \mathsf{M}a_z = 0$. В результате получается хаотическое состояние:

$$\mathsf{M}S(\vec{a}) = \frac{1}{2}(I + \sigma_x \mathsf{M}a_x + \sigma_y \mathsf{M}a_y + \sigma_z \mathsf{M}a_z) = \frac{1}{2}I.$$

23. Пусть X эрмитов оператор, такой что $XY = YX$ для всех эрмитовых операторов Y. Тогда $XY = YX$ для всех операторов Y, поскольку $Y = Y_1 + iY_2$, где Y_1, Y_2 эрмитовы

операторы. Пусть $Y = |e_j\rangle\langle e_k|$, где $|e_j\rangle$ – собственные векторы оператора X с собственными числами x_j. Тогда

$$x_j|e_j\rangle\langle e_k| = XY = YX = |e_j\rangle\langle e_k|x_k,$$

откуда $x_j = x_k$ для всех j, k.

24. Пусть $XY - YX = cI$. Беря след, получаем $\operatorname{Tr} XY - \operatorname{Tr} YX = c \operatorname{Tr} I$, откуда $c = 0$ в силу (1.15).

28. Пусть $X^{(1)}$ обозначает первое измерение наблюдаемой X, а $X^{(2)}$ – второе. Тогда условная вероятность получить при втором измерении исход y, если при первом измерении получен исход x, равна

$$\mathsf{P}_S(X^{(2)} \models y \,|\, X^{(1)} \models x) = \operatorname{Tr} S_x E_y = p(x)^{-1} \operatorname{Tr} E_x S E_x E_y = \delta_{xy}.$$

29. Подставляя в формулу (1.61) одномерные проекторы $E^k_{x_k} = |e^k_{x_k}\rangle\langle e^k_{x_k}|$, получаем

$$\mathsf{P}_S(X^1 \models x^1, \ldots, X^n \models x^n) = p(x^1)p^1(x^2|x^1)\ldots p^{n-1}(x^n|x^{n-1}),$$

где $p(x^1) = \langle e^1_{x^1}|S|e^1_{x^1}\rangle$, $p^{k-1}(x^k|x^{k-1}) = |\langle e^k_{x^k}|e^{k-1}_{x^{k-1}}\rangle|^2$.

31. Используя критерий Сильвестра, получаем, что S является матрицей плотности тогда и только тогда, когда $a^2 + b^2 \leqslant 1$. При этом мера чистоты $P(S) = [3 + 2(a^2 + b^2)]/9 \leqslant 5/9$, так что состояние не может быть чистым.

Для нахождения распределения X и апостериорных состояний рассмотрите отдельно случаи, когда среди чисел x_1, x_2, x_3 нет совпадающих, ровно два совпадающих, все три совпадают.

33. Перейдите в систему координат, в которой вектор $\vec{a}$ совпадает с осью z.

36. Первое утверждение следует из теоремы вращения Эйлера. Для матрицы Адамара: $\alpha = i$, $\varphi = \pi$, $\vec{a} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$.

37. Согласно формуле (1.31), комплексному сопряжению $c_0 \to \overline{c}_0$, $c_1 \to \overline{c}_1$ в $\mathbb{C}^2$ соответствует отображение $a_x \to a_x$, $a_y \to -a_y$, $a_z \to a_z$, т.е. отражение R_{xz} в $\mathbb{R}^3$.

39. Ответ: $\frac{1}{\sqrt{6}}[-1, \sqrt{2}, 0, 1, -\sqrt{2}, 0]^\top$.

40. Ответ: $-\frac{i}{\sqrt{2}}\left(|10\rangle + |01\rangle\right)$.

41. Ответ:

$$\sigma_x \otimes \sigma_z = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \sigma_z \otimes \sigma_x = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{bmatrix},$$

$$\sigma_y \otimes \sigma_x = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \sigma_y \otimes \sigma_y = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

43. Ответ:

$$S_{12} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}; \quad S_1 = S_2 = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2}I.$$

45. Вектор несцепленный: $|\psi\rangle = \frac{1}{2}(|0\rangle - |1\rangle) \otimes (|0\rangle - |1\rangle)$.

62. Из условия $M_x = M_x^2$ следует $\sqrt{M_x}(I - M_x)\sqrt{M_x} = 0$. Поскольку $\sum_x M_x = I$, то $0 \leqslant M_y \leqslant I - M_x$ для $x \neq y$. Поэтому $\sqrt{M_x}M_y\sqrt{M_x} = 0$, откуда $\sqrt{M_y}\sqrt{M_x} = 0$, а значит $M_yM_x = 0$.

63. Идемпотентность матрицы $P = [\langle\psi_j|\psi_k|]\rangle_{j,k=1,\ldots,n}$ означает, что

$$\sum_{k=1}^{n} \langle\psi_j|\psi_k\rangle\langle\psi_k|\psi_l\rangle = \langle\psi_j|\psi_l\rangle, \quad j,l = 1,\ldots,n,$$

или

$$\langle\psi_j|\Bigg(\sum_{k=1}^{n} |\psi_k\rangle\langle\psi_k| - I\Bigg)|\psi_l\rangle = 0. \tag{7.2}$$

При выполнении условия полноты это влечет уравнение переполненности (4.6). Обратно, из переполненности следует полнота и соотношение (7.2).

65. Из условия 1) теоремы 14

$$\max\{\mathcal{P}\{M\}\colon M \in \mathfrak{M}_n\} = \operatorname{Tr}\sum_k W_kM_k^0 = \operatorname{Tr}\sum_k \Lambda^0M_k^0 = \operatorname{Tr}\Lambda^0.$$

Пусть Λ любой эрмитов оператор, такой что $\Lambda \geqslant W_k$, $k = 1,\ldots,n$, тогда

$$\operatorname{Tr}\Lambda = \operatorname{Tr}\sum_k \Lambda M_k^0 \geqslant \operatorname{Tr}\sum_k W_kM_k^0 = \operatorname{Tr}\Lambda^0,$$

поэтому Λ^0 является решением двойственной задачи. При этом, если $\operatorname{Tr}\Lambda = \operatorname{Tr}\Lambda^0$, то $\operatorname{Tr}\sum_k(\Lambda - W_k)M_k^0 = 0$, что вместе с условием $\Lambda \geqslant W_k$, $k = 1,\ldots,n$, влечет $(\Lambda - W_k)M_k^0 = 0$. Используя условие 1), получаем $(\Lambda - \Lambda^0)M_k^0 = 0$, откуда, суммируя по k, имеем $\Lambda - \Lambda^0 = 0$.

66. Соотношения $M_0 = X$, $M_1 = I - X$ устанавливают взаимно-однозначное аффинное соответствие между точками выпуклого множества $\{0 \leqslant X \leqslant I\}$ и множества наблюдаемых с двумя значениями. Из теоремы 10 следует, что крайними точками последнего являются четкие наблюдаемые и только они, что соответствует $X^2 = X$.

70. В силу вогнутости функции $t \to \log t$ имеем

$$H(X) = \sum_{x\in\mathcal{X}} p_x \log\frac{1}{p_x} \leqslant \log\sum_{x\in\mathcal{X}} p_x\frac{1}{p_x} = \log|\mathcal{X}|.$$

71. Код, использующий двоичные последовательности длины $n(H(X)-\delta)$, имеет асимптотически неисчезающую вероятность ошибки, стремящуюся к единице при $n\to\infty$. В самом деле, пусть C – совокупность слов, которые были использованы для взаимно-однозначного кодирования в $n(H(X)-\delta)$ двоичных последовательностей, тогда как все остальные слова кодируются какой-то другой последовательностью. Тогда вероятность ошибки равна $1-\mathsf{P}(C)$, где

$$\mathsf{P}(C) = \mathsf{P}(C\cap T^{n,\delta/2}) + \mathsf{P}(C\cap\overline{T^{n,\delta/2}}) \tag{7.3}$$

$$\leqslant |C|2^{-n(H(X)-\delta/2)} + \mathsf{P}(\overline{T^{n,\delta/2}}) \tag{7.4}$$

$$\leqslant 2^{-n\delta/2}+\epsilon, \tag{7.5}$$

что может быть сделано сколь угодно малым для достаточно больших n.

72. Используя вогнутость функции $t\to -t\log t$, получаем

$$\begin{aligned} H(X|Y) &= -\sum_x\left[\sum_y p_y p(x|y)\log p(x|y)\right] \\ &\leqslant -\sum_x\left[\sum_y p_y p(x|y)\right]\log\left[\sum_{y'} p_{y'}p(x|y')\right] = -\sum_x p_x\log p_x = H(X). \end{aligned}$$

74. Пусть $\{|e_k\rangle\}$ – о.н.б. из общих собственных векторов операторов p_xS_x (не ограничивая общности, можно считать $p_x>0$ для всех x). Обозначая X случайную величину с распределением p_x, а Y случайную величину со значениями k и условным распределением $p(k|x)=\{\langle e_k|S_x|e_k\rangle\}$, имеем

$$I_1(p,M) = H(Y)-H(Y|X),$$

причем $H(Y)=H\big(\sum_x p_xS_x\big)$, $H(Y|X)=\sum_x p_xH\,(S_x)$.

77. Отображение (6.3) является композицией двух отображений: унитарной эволюции и частичного следа, полная положительность которых устанавливается непосредственно.

78. Рассмотрим блочную матрицу $[X_{ij}]_{i,j=1,2}\geqslant 0$, где

$$X_{11}=\begin{bmatrix}1&0\\0&0\end{bmatrix},\quad X_{12}=\begin{bmatrix}0&1\\0&0\end{bmatrix},\quad X_{21}=\begin{bmatrix}0&0\\1&0\end{bmatrix},\quad X_{22}=\begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

Тогда матрица $\left[X_{ij}^{\top}\right]_{i,j=1,2}$ не является положительно определенной (ее определитель равен -1).

79. Рассмотрим классически-квантовый канал (6.4) с состояниями S_j. Используя спектральное разложение, получаем $S_j=\sum_k|\psi_{jk}\rangle\langle\psi_{jk}|$. Тогда (6.4) принимает вид представления Крауса

$$\Phi[S]=\sum_{jk}V_{jk}SV_{jk}^*,$$

где $V_{jk}=|\psi_{jk}\rangle\langle e_j|$.

Для квантово-классического канала (6.5), используя спектральное разложение компонент наблюдаемой $M_j = \sum_k |\varphi_{jk}\rangle\langle\varphi_{jk}|$, имеем

$$\Phi[S] = \sum_{jk} V_{jk} S V_{jk}^*,$$

где $V_{jk} = |e_j\rangle\langle\varphi_{jk}|$.

80. Подставим выражение (6.8) для S_{12} в (6.9):

$$\begin{aligned}\mathrm{Tr}_2\ S_{12}(I_1 \otimes S^\top) &= \mathrm{Tr}_2 \sum_{j,k=1}^{d} \Phi[|e_j\rangle\langle e_k|] \otimes |e_j\rangle\langle e_k| S^\top \\ &= \sum_{j,k=1}^{d} \Phi[|e_j\rangle\langle e_k|]\langle e_j|S|e_k\rangle \\ &= \Phi\Bigg[\sum_{j,k=1}^{d} |e_j\rangle\langle e_j|S|e_k\rangle\langle e_k|\Bigg] = \Phi[S].\end{aligned}$$

81. Напомним, что матрицы Паули $\{\sigma_\gamma; \gamma = 0, x, y, z\}$ образуют базис в вещественном пространстве эрмитовых 2×2-матриц. Из (6.13) с диагональной матрицей T, при $\vec{b} = 0$ получаем

$$\Phi\,[I] = I, \quad \Phi\,[\sigma_\gamma] = \lambda_\gamma \sigma_\gamma, \quad \gamma = x, y, z. \tag{7.6}$$

Подставляя $\sigma_\gamma; \gamma = 0, x, y, z$, в (6.14), получаем также соотношения (7.6), где λ_γ и μ_γ связаны соотношениями (6.15).

Список литературы

[1] К. А. Валиев, А. А. Кокин, *Квантовые компьютеры*: *надежды и реальность*, 2-е изд., ИКИ, М., 2004.

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

[3] С. Я. Килин, Д. Б. Хорошко, А. П. Низовцев, *Квантовая криптография*: *идеи и практика*, Беларуская навука, 2007.

[4] А. Ю. Китаев, "Квантовые вычисления: алгоритмы и исправление ошибок", *УМН*, **52**:6 (1997), 53–112.

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

[6] А. И. Кострикин, Ю. И. Манин, *Линейная алгебра и геометрия*, Наука, М., 1986.

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

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

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

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

[11] К. Хелстром, *Квантовая теория проверки гипотез и оценивания*, Мир, М., 1978.

[12] А. С. Холево, *Вероятностные и статистические аспекты квантовой теории*, 2-е изд., ИКИ, М.-Ижевск, 2003; `http://www.mcnmo.ru/free-books/holevo-qprob.pdf`.

[13] А. С. Холево, *Статистическая структура квантовой теории*, ИКИ, М.-Ижевск, 2003.

[14] А. С. Холево, *Квантовые системы*, *каналы*, *информация*, МЦНМО, М., 2010; `http://www.mcnmo.ru/free-books/holevo-quantum.pdf`.

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

[16] J. Watrous, *The theory of quantum information*, `http://cs.uwaterloo.ca/~watrous/TQI/TQI.pdf`.

Научное издание

Лекционные курсы НОЦ

Выпуск 30

Александр Семенович Холево

Математические основы квантовой информатики

Тираж 150 экз.
Отпечатано в ФГУП «Издательство «Наука» (Типография «Наука»)
121009, Москва, Шубинский пер., 6
http://www.mi.ras.ru/noc/ e-mail: journals@mi.ras.ru