А. С. Холево

КВАНТОВЫЕ СИСТЕМЫ, КАНАЛЫ, ИНФОРМАЦИЯ

ПРЕДИСЛОВИЕ

Квантовая теория информации, изучающая общие закономерности передачи, хранения и преобразования информации в системах, подчиняющихся законам квантовой механики, сформировалась как самостоятельная область исследований в 1990-е годы, однако ее зарождение относится к 1950-м гг., вслед за появлением основ теории информации и помехоустойчивой связи в трудах В. А. Котельникова и К. Шеннона. На начальном этапе, который охватывает период 1950-80 гг., основным вопросом было выяснение фундаментальных ограничений на возможности передачи и обработки информации, обусловленных квантово-механической природой ее носителя. Развитие информационных технологий в направлении микроминиатюризации, использования достижений квантовой оптики и электроники заставляет предположить, что в обозримой перспективе такие ограничения могут стать главным препятствием для дальнейшей экстраполяции существующих технологий и принципов обработки информации.

С другой стороны, появление в 1980-90 гг. идей квантового компьютинга, квантовой криптографии и новых коммуникационных протоколов позволило говорить не только об ограничениях, но и о новых возможностях, заключенных в использовании специфически квантовых ресурсов, таких как *сцепленность* квантовых состояний (entanglement), *квантовый параллелизм*, и *дополнительность* между измерением и возмущением[1]. Квантовая теория информации дает ключ к пониманию фундаментальных закономерностей, до недавних пор остававшихся вне поля зрения исследователей, а также стимулирует развитие экспериментальной физики, значительно расширяющее возможности манипулирования состояниями микросистем и потенциально важное для новых эффективных приложений. В настоящее время работы в области *квантовой информатики*, включающей квантовую теорию информации, ее экспериментальные основы и технологические разработки, ведутся во многих научно-исследовательских центрах развитых стран.

В теории информации центральную роль играют понятия канала связи и его пропускной способности, дающей предельную скорость безоши-

[1] Обсуждение физических аспектов см. в книгах Нильсена и Чанга [22], К. А. Валиева и А. А. Кокина [2], Б. Б. Кадомцева [9].

бочной передачи данных. Математический подход придает этим понятиям универсальную значимость: например, память компьютера (классического или квантового) может рассматриваться как канал из прошлого в будущее, тогда 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.

Математическая традиция в теории информации восходит к А. Н. Колмогорову и А. Я. Хинчину. Для математика, которому не безразлична естественно-научная сторона его исследований, теория информации является источником глубоких идей и новых трудных задач, имеющих достойную мотивацию. В равной, если не в большей мере, это относится и к квантовой теории информации, проблематика которой оказывается тесно связанной с некоммутативным анализом, асимптотической теорией конечномерных нормированных операторных пространств и алгебр, тонкими свойствами структур положительности и тензорного произведения, а также с методами случайных матриц. В 2002 г. в издательстве МЦНМО вышла книга автора [38], посвященная математическому изложению основ квантовой теории информации. В ней, в частности, были затронуты две открытые проблемы: А) аддитивность энтропийных характеристик квантового канала, связанных с его классической пропускной способностью; Б) теорема кодирования для квантовой пропускной способности. Прошедшие годы ознаменовались быстрым прогрессом, в частности, усилиями разных исследователей было получено полное решение проблемы Б). Квантовая пропускная способность оказалась тесно связанной с криптографическими характеристиками канала, такими как пропускная способность для секретной передачи классической информации и скорость распределения случайного ключа. Был выделен важный класс деградируемых каналов, для которых эти две характеристики совпадают и зачастую могут быть вычислены в явном виде. Значительный прогресс был достигнут и в решении проблемы аддитивности (в которой, однако, по-прежнему остается много неисследованных вопросов): неожиданный результат П. Шора показал глобальную эквивалентность ряда ранее разрозненных формулировок[2]. Проблема аддитивности обсуждалась на Международном математическом конгрессе 2006 г., а квантовая

[2] Недавно был анонсирован контрпример к аддитивности минимальной выходной энтропии, см. M. B. Hastings, e-print arXiv:0809.3972.

теория информации вошла в список основных тем на 5-м Европейском конгрессе математиков 2008 г.

Все это показывает, что назрела необходимость в появлении книги, которая не только вводила бы в круг основных понятий квантовой теории информации, но и отражала по крайней мере некоторые из новейших ее достижений. Такую цель и преследует настоящая монография, основное содержание которой составляет математическая теория квантовых каналов. Ее материал разбит на три части.

Часть I содержит введение в статистическую структуру квантовой теории, а также первые применения информационных идей в квантовой статистике. Значительное внимание уделено обсуждению ключевого понятия сцепленных состояний составных квантовых систем. В части II дается обзор основных понятий и некоторых результатов классической теории информации, квантовые аналоги которых составляют основной предмет настоящей книги, а также прямое доказательство основной теоремы кодирования для классически-квантового канала.

Часть III посвящена изучению квантовых каналов общего вида и их информационных характеристик. В гл. 6 подробно обсуждается абстрактное понятие канала, которое является основным для всего рассматриваемого круга вопросов. С точки зрения теории операторов – это вполне положительное отображение соответствующей алгебры, аналог марковского отображения в некоммутативной теории вероятностей. С точки зрения статистической механики канал дает общее конечное описание эволюции открытой квантовой системы, взаимодействующей с окружением. Разнообразные энтропийные характеристики каналов и их свойства исследуются в гл. 7. Гл. 8 посвящена классической пропускной способности квантового канала и уже упоминавшейся проблеме аддитивности энтропийных характеристик каналов относительно операции тензорного произведения. В гл. 9 рассматривается круг вопросов, связанных с исправлением ошибок при хранении и передаче квантовой информации и с соответствующей квантовой пропускной способностью. В гл. 10, 11 действие переносится из конечномерного в сепарабельное гильбертово пространство, которое в частности является ареной для квантовых гауссовских систем. Многие эксперименты по квантовой обработке информации реализованы именно в таких системах с “непрерывными переменными”, основанных на принципах квантовой оптики. С другой стороны, математическая сторона вопроса связана здесь с интересными и ранее не изучавшимися аспектами канонических коммутационных соотношений. Здесь возникает еще одна глубокая аналитическая проблема – “гипотеза об оптимальных гауссовских ансамблях”, сопоставимая по сложности с проблемой аддитивности и, видимо, глубоко связанная с ней.

В основу книги легли лекции, прочитанные автором в Московском государственном университете и Московском физико-техническом инсти-

туте. Автор благодарит М. Е. Широкова, А. А. Кузнецову и В. А. Панова за помощь при подготовке рукописи.

Оглавление

Часть I

1. Векторы и операторы

Мы будем иметь дело с квантовомеханическими системами, которые описываются конечномерными комплексными гильбертовыми пространствами. С одной стороны, уже в этом случае, причем наиболее наглядно, проявляются радикальные отличия квантовой статистики. С другой, системы с конечным числом уровней представляют основной интерес для таких приложений, как передача информации, квантовые вычисления и квантовая криптография (хотя в последнее время все чаще используются системы с непрерывными переменными, которые описываются бесконечномерными пространствами, см. главы 10, 11.)

1.1 Гильбертово пространство

Пусть $\mathcal{H}$ - d-мерное комплексное векторное пространство размерности $\dim \mathcal{H} = d < \infty$, со скалярным произведением $\langle \phi | \, \psi \rangle, \phi, \psi \in \mathcal{H}$, удовлетворяющее аксиомам унитарного пространства; следуя скорее физической, нежели математической традиции, мы считаем, что $\langle \phi | \, \psi \rangle$ линейно по второму аргументу ψ и антилинейно по первому ϕ. Также следуя физической литературе, мы называем такое пространство гильбертовым (хотя в математике в конечномерном случае обычно используется термин "унитарное пространство", а термин "гильбертово" применяют лишь к бесконечномерным пространствам).

Мы будем использовать обозначения Дирака: вектор $\psi \in \mathcal{H}$ (который следует представлять себе как вектор - столбец) будет обозначаться $|\psi\rangle$; соответственно, $\langle \phi |$ определяет линейную функцию на $\mathcal{H}$, задаваемую скалярным произведением:

$$\langle \phi | : \psi \to \langle \phi | \, \psi \rangle, \quad \psi \in \mathcal{H}.$$

Такие линейные функции также образуют гильбертово пространство – двойственное пространство $\mathcal{H}^*$ (его элементы следует представлять себе как векторы - строки $|\phi\rangle^*$, эрмитово - сопряженные к $|\phi\rangle$). Пространство $\mathcal{H}^*$ анти-изоморфно $\mathcal{H}$ в силу отображения $|\psi\rangle \leftrightarrow \langle \psi |$; скалярное произведение $\langle \phi | \, \psi \rangle$ удобно представлять себе как произведение "bra" и "ket" векторов $\langle \phi |$ и $|\psi\rangle$. Квадрат нормы как в $\mathcal{H}$, так и в $\mathcal{H}^*$ равен $\langle \psi | \, \psi \rangle = \|\psi\|^2$.

Такой способ записи векторов позволяет удобно задавать действие операторов. Например, “внешнее произведение” “ket” и “bra” векторов $A = |\psi\rangle\langle\phi|$ есть оператор ранга один, который действует на вектор $|\chi\rangle$ в соответствии с формулой $A|\chi\rangle = |\psi\rangle\langle\phi|\,\chi\rangle$.

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

$$|\psi\rangle = \sum_{i=1}^{d} |e_i\rangle\langle e_i|\,\psi\rangle, \tag{1.1}$$

что эквивалентно равенству

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

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

Задача 1.1.1 *Запишите матричное представление для операторов в $\mathcal{H}$, аналогичное представлению векторов (1.1).*

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

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

$$\beta(\phi,\psi) = 14\sum_{k=0}^{3}(-i)^k\beta(\phi + i^k\psi, \phi + i^k\psi), \tag{1.3}$$

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

1.2 Операторы

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

$$\langle\phi|A^*\psi\rangle = \langle A\phi|\psi\rangle \quad \phi,\psi \in \mathcal{H}. \tag{1.4}$$

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

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

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

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

Теорема 1.2.1 (Спектральное разложение) *Для любого эрмитова оператора A существует ортонормированный базис из собственных векторов, которым отвечают вещественные собственные значения a_i, так что*

$$A = \sum_{i=1}^{d} a_i |e_i\rangle\langle e_i|. \tag{1.5}$$

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

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

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

$$A = \sum_{a\in \mathrm{spec}(A)} aE_a. \tag{1.6}$$

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

$$E_a E_{a'} = \delta_{aa'} E_a, \qquad \sum_{a\in \mathrm{spec}(A)} E_a = I. \tag{1.7}$$

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

1.3 Положительность

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

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

Для любого эрмитова оператора A имеет место разложение

$$A = A_+ - A_-, \tag{1.8}$$

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

Теорема 1.3.1 (Полярное разложение) *Любой оператор A в $\mathcal{H}$ может быть представлен в виде*

$$A = U|A| = |A^*|U, \tag{1.9}$$

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

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

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

1.4 След и двойственность

След оператора T определяется соотношением

$$\mathrm{Tr}\, T = \sum_{i=1}^{d} \langle e_i|Te_i\rangle, \tag{1.10}$$

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

Задача 1.4.1 *Покажите, что это определение не зависит от выбора базиса и что*

$$\operatorname{Tr} A^* = \overline{\operatorname{Tr} A}, \quad \operatorname{Tr} AB = \operatorname{Tr} BA. \tag{1.11}$$

Покажите, что

$$\operatorname{Tr} |\psi\rangle\langle\varphi| A = \langle\varphi|\, A\psi\rangle. \tag{1.12}$$

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

$$\operatorname{Tr} AB \geq 0 \tag{1.13}$$

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

Выражение

$$\|T\|_1 = \operatorname{Tr} |T| \tag{1.14}$$

определяет *следовую норму* на комплексном линейном векторном пространстве всех операторов, действующих в $\mathcal{H}$. Для эрмитова оператора T

$$\|T\|_1 = \operatorname{Tr} T_+ + \operatorname{Tr} T_- = \sum_{i=1}^{d} |t_i|, \tag{1.15}$$

где t_i – собственные значения оператора T. Другой важной нормой является *операторная норма*

$$\|A\| \equiv \|A\|_\infty = \max_{\psi:\|\psi\|=1} \|A\psi\|, \tag{1.16}$$

которая для эрмитова оператора A равна

$$\|A\| = \max_{a\in\operatorname{spec}(A)} |a|. \tag{1.17}$$

Имеет место неравенство

$$|\operatorname{Tr} TA| \leq \|T\|_1 \|A\|, \tag{1.18}$$

где равенство достижимо как для любого фиксированного T, так и для любого фиксированного A. А именно, для фиксированного T следует положить $A = U^*$, где U – унитарный оператор из полярного разложения $T = |T|U$. Тогда,

$$\|T\|_1 = \max_U |\operatorname{Tr} TU| = \max_{\|A\|=1} |\operatorname{Tr} TA|. \tag{1.19}$$

Аналогично, для заданного A равенство достигается при $T = |\psi\rangle\langle\psi|U^*$, где U – унитарный оператор из полярного разложения $A = U|A|$, а ψ – нормированный собственный вектор оператора $|A|$ с наибольшим собственным значением. Отсюда следует, что

$$\|A\| = \max_{\|T\|_1=1} |\operatorname{Tr} TA|. \tag{1.20}$$

Эти факты лежат в основе важного соотношения двойственности. Пространство всех операторов в $\mathcal{H}$, снабженное следовой нормой, превращается в комплексное банахово пространство $\mathfrak{T}(\mathcal{H})$. Это же пространство, снабженное операторной нормой, является банаховой алгеброй $\mathfrak{B}(\mathcal{H})$. Эти пространства являются двойственными; это означает, что любая линейная функция на $\mathfrak{T}(\mathcal{H})$ представима в виде $T \to \operatorname{Tr} TA$ для некоторого $A \in \mathfrak{B}(\mathcal{H})$ и имеет норму (1.20), и, наоборот, любая линейная функция на $\mathfrak{B}(\mathcal{H})$ представима в виде $A \to \operatorname{Tr} TA$ для некоторого $T \in \mathfrak{T}(\mathcal{H})$, и имеет норму (1.19).

Индекс h в обозначениях пространств операторов будет использоваться для обозначения соответствующих вещественных банаховых пространств эрмитовых операторов. Вещественные банаховы пространства $\mathfrak{T}_h(\mathcal{H})$ и $\mathfrak{B}_h(\mathcal{H})$ также находятся в двойственности, задаваемой билинейной формой $T, A \to \operatorname{Tr} TA$. То, что эта форма принимает вещественные значения для эрмитовых T, A, следует из (1.11).

В конечномерном случае пространства $\mathfrak{T}(\mathcal{H})$ и $\mathfrak{B}(\mathcal{H})$ совпадают с множеством всех операторов в $\mathcal{H}$ и различаются только нормами. Однако в бесконечномерном случае они существенно различаются (так же, как пространства последовательностей l^1 и l^∞ над конечным и бесконечным множествами индексов).

Задача 1.4.2 *Покажите, что (комплексная) размерность пространства всех операторов в $\mathcal{H}$ равна d^2, тогда как (вещественная) размерность пространства всех эрмитовых операторов в $\mathcal{H}$ также равна d^2. Если $\mathcal{H}$ – вещественное d-мерное гильбертово пространство, то размерность пространства всех операторов в $\mathcal{H}$ равна d^2, а размерность подпространства всех симметрических операторов в $\mathcal{H}$ равна $d(d+1)/2$.*

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 $\operatorname{extr}(\mathfrak{S})$ множество всех крайних точек выпуклого множества $\mathfrak{S}$. Имеет место следующий общий результат:

Теорема 1.5.1 (Каратеодори) *Пусть $\mathfrak{S}$ – компактное выпуклое подмножество $\mathbb{R}^n$, тогда любая точка $S \in \mathfrak{S}$ может быть представ-*

лена в виде выпуклой комбинации не более чем $n+1$ крайних точек $S_j \in extr(\mathfrak{S})$:

$$S = \sum_{j=1}^{n+1} p_j S_j, \quad S_j \in \mathrm{extr}(\mathfrak{S}). \tag{1.21}$$

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

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

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

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

Задача 1.5.1 *Непрерывная выпуклая (в частности, аффинная) функция на компактном выпуклом множестве $\mathfrak{S}$ достигает своего максимума в крайней точке этого множества.*

1.6 Комментарии

Основы операторного формализма квантовой механики были заложены в классическом труде Дирака [8]. Из множества учебников по линейной алгебре отметим современный курс Кострикина и Манина [13], в котором учтены потребности квантовой механики, а также книгу Глазмана и Любича [6], нацеленную на активное освоение некоммутативного конечномерного функционального анализа через решение задач. По поводу основных понятий и фактов выпуклого анализа см. книги Магарил-Ильяева и Тихомирова [17], Рокафеллара [25].

2. Состояния, наблюдаемые, статистика

2.1 Структура статистических теорий

2.1.1 Классические системы

Классическая система характеризуется наличием *фазового пространства* Ω, точки которого ω описывают детерминированные состояния системы. Для простоты далее рассматривается случай конечного множества Ω. (Статистическим) *состоянием* называется распределение вероятностей $P = \{p_\omega\}$ на Ω. Совокупность всех распределений вероятностей представляет собой симплекс $\mathfrak{P}(\Omega)$, в котором каждая точка однозначно представляется в виде выпуклой комбинации крайних точек – *чистых состояний*, задаваемых вырожденными распределениями вероятностей. Простейшей системой является *бит* – система с двумя чистыми состояниями $0, 1$, для которой множество статистических состояний изоморфно единичному отрезку.

Всякая детерминированная наблюдаемая величина есть функция $X = \{x_\omega\}$ на фазовом пространстве Ω, задающая разбиение Ω на непересекающиеся подмножества Ω_x, в которых X принимает какое-то значение x. Индикаторы $E_x(\omega)$ этих подмножеств удовлетворяют условиям

$$E_x(\omega)E_y(\omega) = \delta_{x,y}E_x(\omega); \quad \sum_x E_x(\omega) = 1.$$

Помимо детерминированных наблюдаемых, которые далее будут называться *четкими*, возможны *нечеткие* наблюдаемые, которые принимают значение x с вероятностями $M(x|\omega)$ (наблюдения с ошибкой, либо рандомизованные). Набор условных вероятностей $M = \{M(x|\omega)\}$ характеризуется свойствами

$$M(x|\omega) \geq 0; \quad \sum_x M(x|\omega) = 1.$$

Для четких наблюдаемых вероятности $M(x|\omega) = E_x(\omega)$ равны 0 или 1, т. е. $M(x|\omega)^2 = M(x|\omega)$.

Распределение вероятностей произвольной наблюдаемой M в состоянии P дается формулой

$$\mu_P^M(x) = \sum_\omega p_\omega M(x|\omega). \tag{2.1}$$

Для плавного перехода к квантовым системам полезно ввести матричное представление классических величин. Рассмотрим гильбертово пространство, в котором фиксирован ортонормированный базис $\{|\omega\rangle; \omega \in \Omega\}$. Всякой числовой функции f_ω на Ω сопоставим оператор

$$f = \sum_\omega f_\omega |\omega\rangle\langle\omega|,$$

диагональный в этом базисе. Тогда состоянию сопоставляется оператор $P = \sum_\omega p_\omega |\omega\rangle\langle\omega|$, характеризуемый свойствами

$$P \geq 0; \quad \operatorname{Tr} P = 1. \tag{2.2}$$

Наблюдаемая задается разложением единицы, т. е. набором операторов $M = \{M_x\}$, где

$$M_x = \sum_\omega M(x|\omega)|\omega\rangle\langle\omega|,$$

удовлетворяющих условиям

$$M_x \geq 0; \quad \sum_x M_x = I. \tag{2.3}$$

При этом для четкой наблюдаемой операторы M_x являются попарно ортогональными проекторами.

Формула (2.1) для распределения вероятностей наблюдаемой переходит в

$$\mu_P^M(x) = \operatorname{Tr} P M_x. \tag{2.4}$$

2.1.2 Аксиомы статистического описания

В основе математической структуры квантовой теории лежит разделение статистического эксперимента на стадии приготовления и измерения, которое порождает двойственность между квантовыми состояниями и наблюдаемыми. Следующий набор аксиом суммирует общие черты статистического описания и применим как к классическим, так и квантовым системам.

Аксиома 1 *Пусть задано множество $\mathfrak{S}$, элементы которого называют состояниями, и множество $\mathfrak{M}$, элементы которого называют (конечнозначными) наблюдаемыми. Для произвольного $M \in \mathfrak{M}$ определено конечное множество $\mathcal{X}^M$ возможных исходов измерения наблюдаемой. Для любых $S \in \mathfrak{S}$ и $M \in \mathfrak{M}$ определено распределение вероятностей μ_S^M на $\mathcal{X}^M$, называемое распределением вероятностей наблюдаемой M в состоянии S.*

Состояние S интерпретируется как более или менее подробное описание приготовления *статистического ансамбля* независимых индивидуальных экземпляров рассматриваемой системы, а наблюдаемая M – как величина, которая может быть измерена определенным прибором для каждого экземпляра системы. Таким образом, аксиома 1 предполагает *воспроизводимость* экспериментов и *устойчивость частот* при независимых повторениях эксперимента. Если эти предположения не выполнены для некоторой ситуации (например, по причине ее уникальности), то говорить о статистическом описании вообще не имеет смысла.

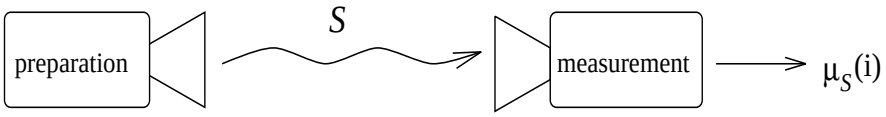


Рис. 2.1. Квантовое состояние, описываемое оператором плотности S, характеризует приготовление системы, тогда как статистика измерений описывается вероятностной мерой $\mu_S^M(x)$, где x обозначает возможные исходы измерения.

Вторая аксиома выражает возможность *смешивания* ансамблей.

Аксиома 2 *Для любых $S_1, S_2 \in \mathfrak{S}$ и любого числа $p, 0 < p < 1$, найдется $S \in \mathfrak{S}$, такое, что $\mu_S^M = p\mu_{S_1}^M + (1-p)\mu_{S_2}^M$ для всех $M \in \mathfrak{M}$. S называется смесью состояний S_1 и S_2 с весами $p, 1-p$.*

Следующая аксиома описывает возможность обработки информации, полученной при измерении наблюдаемой. Пусть $M_1, M_2 \in \mathfrak{M}$, а f – функция из $\mathcal{X}^{M_1}$ в $\mathcal{X}^{M_2}$ такая, что для всех $S \in \mathfrak{S}$

$$\mu_S^{M_2}(y) = \sum_{x:f(x)=y} \mu_S^{M_1}(x)$$

Тогда наблюдаемая M_2 называется *укрупнением* наблюдаемой M_1. В этом случае мы будем писать $M_2 = f \circ M_1$.

Аксиома 3 *Для всякой наблюдаемой $M_1 \in \mathfrak{M}$ и (с необходимостью конечнозначной) функции f на $\mathcal{X}^{M_1}$ найдется наблюдаемая $M_2 \in \mathfrak{M}$, такая, что $M_2 = f \circ M_1$.*

Пара непустых множеств $(\mathfrak{S}, \mathfrak{M})$, удовлетворяющих аксиомам 1 - 3 называется *статистической моделью* системы. Статистическая модель называется *отделимой*, если

Аксиома 4 *Из того, что $\mu_{S_1}^M = \mu_{S_2}^M$ для всех $M \in \mathfrak{M}$, следует, что $S_1 = S_2$, а из того, что $\mu_S^{M_1} = \mu_S^{M_2}$ для всех $S \in \mathfrak{S}$, следует, что $M_1 = M_2$.*

Для отделимых моделей как смешивание в $\mathfrak{S}$ так и укрупнение в $\mathfrak{M}$ определены однозначно. Таким образом, множество состояний $\mathfrak{S}$ приобретает выпуклую структуру, а множество наблюдаемых $\mathfrak{M}$ – структуру частичного порядка.

Наблюдаемые $M_1, \ldots, M_m$ называются *совместимыми*, если они являются укрупнениями некоторой наблюдаемой M, то есть $M_j = f_j \circ M$ для $j = 1, \ldots, m$. Результаты измерений совместимых наблюдаемых могут быть получены посредством обработки данных эксперимента с одним измерением. Статистические модели, в которых все наблюдаемые совместимы, являются по существу классическими.

Предложение 2.1.1 *Пусть $(\mathfrak{S}, \mathfrak{M})$ – отделимая статистическая модель, в которой существует максимальная наблюдаемая M^*, такая что все остальные наблюдаемые являются ее укрупнениями. Тогда найдется конечное множество Ω, взаимно однозначное аффинное отображение $S \to P_S$ множества состояний $\mathfrak{S}$ на выпуклое подмножество распределений вероятностей на Ω, и взаимно однозначное отображение $M \to f_M$ множества наблюдаемых $\mathfrak{M}$ на подмножество случайных величин на Ω, сохраняющее отношение укрупнения, такие что*

$$\mu_S^M(x) = \sum_{\omega: f_M(\omega) = x} P_S(\omega). \tag{2.5}$$

Доказательство. Возьмем за Ω множество исходов максимальной наблюдаемой M^*, и пусть $P_S(\omega)$ – ее распределение в состоянии S. Согласно предположению, для любой наблюдаемой M найдется функция f_M, такая что выполняется соотношение (2.5). Взаимная однозначность соответствий $S \to P_S$ и $M \to f_M$ вытекает из отделимости. Проверка их свойств оставляется в качестве задачи. □

Таким образом, возможные значения "максимальной наблюдаемой" составляют фазовое пространство Ω, и состояния представляются распределениями вероятностей P на пространстве Ω. В базовой классической модели, которую естественно назвать *моделью Колмогорова*, множеством состояний является симплекс всех распределений вероятностей $\mathfrak{P}\Omega$, а наблюдаемые описываются случайными величинами, т. е. (измеримыми) функциями на пространстве Ω. Доказанное предложение означает, что статистическая модель, в которой все наблюдаемые совместимы, вкладывается в модель Колмогорова.

Другой важной классической моделью является *модель Вальда*, отличающаяся от модели Колмогорова включением "нечетких" или "рандомизованных" наблюдаемых* (см. раздел 2.1.1). Чтобы получить для нее результат, аналогичный предложению 2.1.1, следует ввести стохастические

* Рандомизация была введена Вальдом в теории статистических решений, а также фон Нейманом в теории игр.

варианты понятий укрупнения и совместимости. Пусть $M_1, M_2 \in \mathfrak{M}$, и Π – переходная вероятность из $\mathcal{X}^{M_1}$ в $\mathcal{X}^{M_2}$, такая что для всех $S \in \mathfrak{S}$

$$\mu_S^{M_2}(y) = \sum_x \Pi(y|x)\mu_S^{M_1}(x)$$

Тогда наблюдаемая M_2 называется *стохастическим укрупнением* наблюдаемой M_1. Наблюдаемые $M_1, \ldots, M_m$ называются *стохастически совместимыми*, если все они являются стохастическими укрупнениями некоторой наблюдаемой M. В модели Вальда все наблюдаемые стохастически совместимы, и всякая отделимая модель, обладающая этим свойством, вкладывается в модель Вальда.

В статистической модели квантовой механики состояния и наблюдаемые описываются операторами в гильбертовом пространстве. Было предпринято много попыток построения аксиоматических схем, приводящих к формализму гильбертова пространства в квантовой механике. Однако это не является нашей целью и мы будем следовать по другому пути: мы возьмем за отправной пункт описание квантовых состояний как операторов плотности и следуя аксиомам статистической модели выведем из этого наиболее общую концепцию квантовой наблюдаемой (включающую как "четкие", так и "нечеткие" наблюдаемые).

2.1.3 Квантовые состояния

Состояние квантово-механической системы описывает статистический ансамбль независимых одинаково приготовленных экземпляров системы.

Определение 2.1.1 *Квантовое состояние задается оператором плотности, т. е. эрмитовым оператором S в гильбертовом пространстве системы $\mathcal{H}$, удовлетворяющим условиям*

$$S \geq 0, \quad \operatorname{Tr} S = 1.$$

Пусть $\mathfrak{S}(\mathcal{H})$ – множество всех операторов плотности. Это есть выпуклое подмножество вещественного линейного пространства всех эрмитовых операторов в $\mathcal{H}$. Выпуклая комбинация $S = \sum_j p_j S_j$ операторов плотности описывает *смесь* с весами p_j соответствующих статистических ансамблей, приготовленных в состояниях S_j. В квантовом статистическом ансамбле есть два вида стохастичности: во-первых, устранимая в принципе стохастичность, обусловленная флуктуациями классических параметров процедуры приготовления, и во-вторых, неуничтожимая никакими усилиями квантовая стохастичность, присутствующая в любом состоянии. Следующая теорема характеризует такие состояния без классической случайности.

Теорема 2.1.2 *Крайние точки множества квантовых состояний $\mathfrak{S}(\mathcal{H})$, называемые чистыми состояниями, суть одномерные проекторы и только они.*

Доказательство. Рассмотрим спектральное разложение эрмитова оператора S

$$S = \sum_{i=1}^{d} s_i |e_i\rangle\langle e_i|, \quad s_j \geq 0, \sum s_j = 1, \tag{2.6}$$

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

$$pS_1(I - S_1) + (1-p)S_2(I - S_2) + p(1-p)(S_1 - S_2)^2 = S - S^2 = 0. \tag{2.7}$$

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. □

С точки зрения теории вероятностей, одномерные проекторы являются некоммутативным аналогом вырожденных распределений, тогда как роль равномерного распределения играет *хаотическое состояние* с оператором плотности $S = \frac{1}{d} I$.

Задача 2.1.1 *Покажите, что если* $\dim \mathcal{H} = d$, *то* $\mathfrak{S}(\mathcal{H})$ *погружается в вещественное пространство размерности* $n = d^2 - 1$. *Если* $\mathcal{H}$ *– евклидово (вещественное гильбертово), то* $n = d(d+1)/2 - 1$.

Спектральное разложение (2.6) показывает, что любое квантовое состояние представимо в виде смеси не более чем d чистых состояний, где $d = \dim \mathcal{H}$; таким образом, для множества квантовых состояний теорема Каратеодори дает завышенное значение количества крайних точек, в действительности присутствующих в смеси (1.21). С другой стороны, эта теорема дает точное значение для симплекса распределений вероятностей на "фазовом пространстве" $\Omega = \{1, \ldots, n+1\}$, представляющем статистические состояния в классической системе. 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. раздел 3.1.4).

Простейшим, и в тоже время фундаментальным примером является *q-бит* – двухуровневая квантовая система, $\dim \mathcal{H} = 2$. Будем использовать канонический базис: $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$. В вещественном пространстве эрмитовых 2×2-матриц удобно ввести *базис Паули*:

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

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

$$S = \frac{1}{2}(I + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z) = \frac{1}{2}\begin{bmatrix} 1+a_z & a_x - ia_y \\ a_x + ia_y & 1-a_z \end{bmatrix}. \tag{2.8}$$

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

$$|\boldsymbol{a}|^2 \equiv a_x^2 + a_y^2 + a_z^2 \leq 1.$$

Таким образом, $\mathfrak{S}(\mathcal{H})$ как выпуклое множество изоморфно единичному шару в $\mathbb{R}^3$, который мы будем называть *шаром Блоха*.

Задача 2.1.2 *Покажите, что оператор плотности (2.8) имеет собственные значения* $\frac{1\pm|\boldsymbol{a}|}{2}$.

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

$$|\boldsymbol{a}\rangle = \begin{bmatrix} \cos\frac{\theta}{2}e^{-i\phi/2} \\ \sin\frac{\theta}{2}e^{i\phi/2} \end{bmatrix}. \tag{2.9}$$

Для электрона, являющегося частицей со спином $1/2$, вектор $\boldsymbol{a}$ описывает ансамбль (пучок частиц) со спином в направлении $\boldsymbol{a}$, приготовленный при помощи фильтра Штерна-Герлаха в направлении градиента магнитного поля. Смешанное состояние с $a_x = a_y = a_z = 0$, для которого все направления спина равновероятны, описывается оператором плотности $S = \frac{1}{2}I$, т. е. является хаотическим.

Другим важным примером двухуровневой системы является поляризация монохроматического фотона, визуализуемая в классических оптических экспериментах. В этом случае параметр $\frac{\theta}{2}$ есть угол линейной поляризации, в то время как $\frac{\phi}{2}$ характеризует круговую поляризацию. Для линейно поляризованного фотона $\phi = 0$, в частности, для вертикально поляризованного фотона $\theta = 0$, а для горизонтально поляризованного – $\theta = \pi$.

2.2 Квантовые наблюдаемые

2.2.1 Квантовые наблюдаемые: построение из аксиом

Рассмотрим статистическую модель, в которой пространство состояний есть выпуклое множество $\mathfrak{S}(\mathcal{H})$ всех операторов плотности в гильбертовом пространстве $\mathcal{H}$. Пусть M – некоторая наблюдаемая, тогда согласно аксиоме 2 вероятности исходов измерения $\mu_S^M(x)$ должны быть *аффинными* функциями состояния.

Теорема 2.2.1 *Предположим, что отображение $S \to \mu_S$ является аффинной функцией на $\mathfrak{S}(\mathcal{H})$, причем $0 \le \mu_S \le 1$. Тогда существует единственный оператор M в $\mathcal{H}$, $0 \le M \le I$ такой, что для любого $S \in \mathfrak{S}(\mathcal{H})$*

$$\mu_S = \operatorname{Tr} SM. \tag{2.10}$$

Набросок доказательства. Покажем, как можно расширить отображение $S \to \mu_S$ до линейной функции на пространстве $\mathfrak{T}(\mathcal{H})$ всех эрмитовых операторов в $\mathcal{H}$.

Прежде всего отметим, что любой эрмитов оператор A можно представить как разность двух положительных операторов, $A = A'_+ - A'_-$, например, в силу разложения (1.8). Нормируя на величины следов $t_\pm = \operatorname{Tr} A'_\pm$, мы всегда можем записать

$$A = t_+ S_+ - t_- S_-, \quad t_\pm \ge 0, \quad S_\pm \in \mathfrak{S}(\mathcal{H}). \tag{2.11}$$

Представление оператора A в виде линейной комбинации операторов плотности (2.11), конечно, не является единственным.

Задача 2.2.1 *Используя аффинность отображения $S \to \mu_S$, покажите, что значение $\mu_A = t_+\mu_{S_+} - t_-\mu_{S_-}$ определено единственным образом (т. е. зависит только от оператора A, но не от конкретного разложения $A = A'_+ - A'_-$), а функция $A \to \mu_A$ является вещественной и линейной.*

Далее, любой оператор A может быть представлен в виде $A = A_1 + iA_2$, где $A_1 = \frac{1}{2}(A + A^*), A_2 = \frac{1}{2i}(A - A^*)$ – эрмитовы операторы. Это представление единственно, и положив $\mu_A = \mu_{A_1} + i\mu_{A_2}$ мы получаем единственное комплексно линейное расширение функции μ_A на пространство всех линейных операторов в $\mathcal{H}$. Как отмечено в разделе 1.4, любая такая функция имеет вид

$$\mu_A = \operatorname{Tr} AM, \tag{2.12}$$

где M – (единственный) оператор в H. Взяв $A = |\psi\rangle\langle\psi|$, где ψ – единичный вектор, и используя (1.12), имеем, по предположению, $0 \le \langle\psi|M\psi\rangle \le 1$, следовательно, $0 \le M \le I$. □

Следствие 2.2.1 *Пусть $S \to \{\mu_S(x), x \in \mathcal{X}\}$ – аффинное отображения выпуклого множества состояний $\mathfrak{S}(\mathcal{H})$ в множество распределений вероятностей на конечном множестве $\mathcal{X}$. Тогда существует единственный набор эрмитовых операторов $\{M_x, x \in \mathcal{X}\}$ в $\mathcal{H}$, такой что $\sum_x M_x = I$ и*

$$\mu_S(x) = \operatorname{Tr} SM_x, \quad x \in \mathcal{X}. \tag{2.13}$$

Доказательство этого следствия предоставляется читателю в качестве **задачи**.

Набор операторов $\{M_x\}$ называется *разложением единицы* или *вероятностной операторно-значной мерой* на $\mathcal{X}$. Следствие 2.2.1 приводит к следующему определению:

Определение 2.2.1 *Квантовая наблюдаемая со значениями в множестве $\mathcal{X}$ описывается разложением единицы $\{M_x, x \in \mathcal{X}\}$. Вероятностное распределение μ_S^M наблюдаемой $M = \{M_x\}$ в состоянии S задается формулой(2.13).*

Принимая такое определение квантовой наблюдаемой, мы вводим *максимальную* статистическую модель, множество состояний которой совпадает с множеством всех операторов плотности в $\mathcal{H}$. В стандартных текстах по квантовой механике используется более узкое понятие квантовой наблюдаемой.

Определение 2.2.2 *Наблюдаемая называется четкой, если все операторы $M_x = E_x$ являются проекторами: $E_x^2 = E_x$.*

Задача 2.2.2 *Покажите, что условие $E_x^2 = E_x$ для всех исходов x эквивалентно тому, что $E_x E_y = \delta_{xy} E_x$ для всех x, y.*

Соответствующие разложения единицы называются *ортогональными*. Итак, четкие наблюдаемые описываются ортогональными разложениями единицы в $\mathcal{H}$. Наблюдаемые, значения которых являются вещественными числами, т. е. $\mathcal{X} \subset \mathbb{R}$, называются *вещественными*. Спектральное разложение

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

задает взаимно-однозначное соответствие между четкими вещественными наблюдаемыми $E = \{E_x\}$ и эрмитовыми операторами X в $\mathcal{H}$. Укрупнение для таких наблюдаемых принимает форму функционального исчисления, так что для любой функции $f : \mathbb{R} \to \mathbb{R}$ наблюдаемая $f \circ E$ описывается эрмитовым оператором $f(X)$ (**задача**). *Математическое ожидание* (среднее значение) наблюдаемой X в состоянии S задается *статистической формулой Борна-фон-Неймана*

$$\mathsf{E}_S(X) = \sum x \mu_S^E(x) = \operatorname{Tr} SX.$$

Соотношение (2.13) можно рассматривать как обобщение этой формулы.

Статистическая модель, в которой состояния описываются операторами плотности, а наблюдаемые – эрмитовыми (самосопряженными) операторами, была детально рассмотрена фон Нейманом.

2.2.2 Совместимость и дополнительность

Определение 2.2.3 *Коммутатором операторов X, Y называется оператор $[X, Y] = XY - YX$. Операторы X, Y коммутируют, если $[X, Y] = 0$.*

Теорема 2.2.2 *Пусть $E = \{E_x\}$ и $F = \{F_y\}$ – четкие наблюдаемые. Тогда следующие утверждения эквивалентны:*

(i) E, F совместимы;

(ii) Проекторы E_x и F_y коммутируют для всех x, y;

Если E, F вещественные наблюдаемые, то это эквивалентно условию:

(iii) Соответствующие эрмитовы операторы $X = \sum_x xE_x, Y = \sum_y yF_y$ коммутируют.

Доказательство. $(ii) \Rightarrow (i)$. Положим $M_{x,y} = E_xF_y = F_yE_x$. Так как произведение коммутирующих проекторов является проектором, то $M = \{M_{x,y}\}$ – (четкая) наблюдаемая. Кроме того, $E = f \circ M, F = g \circ M$, где $f(x,y) = x, g(x,y) = y$. Таким образом, E, F совместимы.

$(i) \Rightarrow (ii)$. Если E, F совместимы, то найдется наблюдаемая $M = \{M_z\}$, такая, что

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

для некоторых функций f, g (условимся считать $M_z = 0$, если z, удовлетворяющих условию в суммах, не существует).

Лемма 2.2.3 *Пусть $0 \le A \le P$, где P – проектор. Тогда $AP = A$ и, следовательно, $[A, P] = 0$.*

Доказательство. Имеем $(I - P)A(I - P) = 0$, откуда последовательно получаем $\sqrt{A}(I-P) = 0$, $A(I-P) = 0$, $A = AP$, и $PA = (AP)^* = AP$. □

Пусть теперь $f(z) = x$, тогда $E_x \ge M_z$; согласно лемме, $[E_x, M_z] = 0$. Если $f(z) \ne x$, то $I - E_x \ge M_z$, и вновь $[E_x, M_z] = 0$. Итак, E_x коммутирует со всеми M_z и, значит, с $F_y = \sum_{z:g(z)=y} M_z$.

$(ii) \Rightarrow (iii)$ очевидно. Обратно, пусть $[X, Y] = 0$, тогда $[f(X), g(Y)] = 0$ для произвольных многочленов f, g. Рассматривая многочлены, которые обращаются в нуль во всех точках соответствующего спектра, за исключением x (соотв. y), получаем $[E_x, F_y] = 0$. □

Если E, F совместимы, то обозначая $\hat{M}_{xy} = \sum_{z:f(z)=x,g(z)=y} M_z$, имеем

$$E_x = \sum_y \hat{M}_{xy}, \quad F_y = \sum_x \hat{M}_{xy}.$$

Наблюдаемая $\hat{M}$ описывает статистику совместного измерения E, F; их совместное распределение вероятностей в состоянии S определяется формулой

$$\mu_S^{EF}(x,y) = \operatorname{Tr} S\hat{M}_{xy}.$$

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

Задача 2.2.3 *Покажите, что единственной вещественной наблюдаемой, совместимой со всеми квантовыми наблюдаемыми, является постоянная, то есть наблюдаемая, кратная единичному оператору.*

Существование несовместимых наблюдаемых – это проявление квантового свойства *дополнительности*. Физические измерения над микрообъектами производятся при помощи макроскопических экспериментальных устройств (как, например, фильтр Штерна - Герлаха), предполагающих сложную и специфичную пространственно-временную организацию окружающей среды . Различные способы такой организации, соответствующие измерениям различных наблюдаемых, могут быть взаимно исключающими (несмотря на то, что относятся к одинаково приготовленному микрообъекту), то есть дополнительными. Дополнительность – это первое фундаментальное различие между квантовой и классической статистическими моделями.

Пример Рассмотрим математическую модель частицы со спином $1/2$. Спин является векторной величиной, компоненты которой вдоль осей x, y, z описываются эрмитовыми операторами $\sigma_x, \sigma_y, \sigma_z$. Пусть единичный вектор $\boldsymbol{a} = (a_x, a_y, a_z)$ задает направление в $\mathbb{R}^3$, тогда эрмитов оператор

$$\sigma(\boldsymbol{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z = \begin{bmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{bmatrix} \tag{2.14}$$

задает наблюдаемую "проекция спина на направление $\boldsymbol{a}$". Оператор $\sigma(\boldsymbol{a})$ имеет собственные значения ± 1 (спин вдоль и против направления $\boldsymbol{a}$) и спектральное разложение

$$\sigma(\boldsymbol{a}) = |\boldsymbol{a}\rangle\langle\boldsymbol{a}| - |-\boldsymbol{a}\rangle\langle-\boldsymbol{a}|. \tag{2.15}$$

Напомним, что направление $\boldsymbol{a}$ имеет углы Эйлера (θ, ϕ), а вектор $\boldsymbol{a}$, задаваемый соотношением (2.9), есть вектор чистого состояния в $\mathcal{H}$ с направлением спина $\boldsymbol{a}$. Соответствующий оператор плотности, являющийся проектором на вектор $|\boldsymbol{a}\rangle$, имеет вид

$$S(\boldsymbol{a}) = |\boldsymbol{a}\rangle\langle\boldsymbol{a}| = \frac{I + \sigma(\boldsymbol{a})}{2}.$$

Задача 2.2.4 *Докажите формулу*

$$\sigma(\boldsymbol{a}_1)\sigma(\boldsymbol{a}_2) = (\boldsymbol{a}_1 \cdot \boldsymbol{a}_2)I + i\sigma(\boldsymbol{a}_1 \times \boldsymbol{a}_2). \tag{2.16}$$

Это соотношение, записанное для базисных векторов, дает

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

$$\sigma_x\sigma_y = i\sigma_z, \quad \sigma_y\sigma_z = i\sigma_x, \quad \sigma_z\sigma_x = i\sigma_y. \tag{2.17}$$

Принимая во внимание, что $\operatorname{Tr}\sigma(\boldsymbol{a}) = 0$, соотношение (2.16) приводит к формуле для среднего значения $\sigma(\boldsymbol{b})$:

$$\operatorname{Tr} S(\boldsymbol{a})\sigma(\boldsymbol{b}) = \boldsymbol{a}\cdot\boldsymbol{b}.$$

Другим следствием формулы (2.16) является равенство

$$[\sigma(\boldsymbol{a}_1), \sigma(\boldsymbol{a}_2)] = 2i\sigma(\boldsymbol{a}_1 \times \boldsymbol{a}_2), \tag{2.18}$$

которое означает, что наблюдаемые $\sigma(\boldsymbol{a}_1), \sigma(\boldsymbol{a}_2)$ совместимы тогда и только тогда, когда $\boldsymbol{a}_1 = \pm\boldsymbol{a}_2$. В частности, компоненты спина $\sigma_x, \sigma_y, \sigma_z$ являются несовместимыми наблюдаемыми.

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

$$\mathsf{B}_S(X) = \left[\operatorname{Tr} SX_j^0 \circ X_k^0\right]_{j,k=1,\ldots,n}, \tag{2.19}$$

и кососимметричную *коммутационную матрицу*

$$\mathsf{C}_S(X) = \left[i\operatorname{Tr} S[X_j, X_k]\right]_{j,k=1,\ldots,n} = \left[i\operatorname{Tr} S[X_j^0, X_k^0]\right]_{j,k=1,\ldots,n}. \tag{2.20}$$

Имеем

$$\mathsf{B}_S(X) \geq \frac{i}{2}\mathsf{C}_S(X) \tag{2.21}$$

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

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

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

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

где $Z = \sum_{j=1}^{n} c_j X_j$.

Для двух наблюдаемых $X_1 = X$ и $X_2 = Y$ неравенство (2.21) эквивалентно соотношению неопределенностей Шредингера - Робертсона

$$\mathsf{D}_S(X)\mathsf{D}_S(Y) \geq \{\mathsf{E}_S(X - I\mathsf{E}_S(X)) \circ (Y - I\mathsf{E}_S(Y))\}^2 + \frac{1}{4}|\mathsf{E}_S[X,Y]|^2, \quad (2.22)$$

где

$$\mathsf{D}_S(X) = \operatorname{Tr} S(X - I\mathsf{E}_S(X))^2 \quad (2.23)$$

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

$$\mathsf{E}_S(X - I\mathsf{E}_S(X)) \circ (Y - I\mathsf{E}_S(Y)) \quad (2.24)$$

представляют собой *ковариацию* X, Y в состоянии S; в этом случае $[X, Y] = 0$ и (2.22) 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.

Задача 2.2.5 *Докажите некоммутативное неравенство Коши- Буняковского*

$$|\operatorname{Tr} SX^*Y|^2 \leq \operatorname{Tr} SX^*X \operatorname{Tr} SY^*Y, \quad (2.25)$$

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

2.2.4 Выпуклая структура множества наблюдаемых

Множество всех (конечнозначных) квантовых наблюдаемых в гильбертовом пространстве $\mathcal{H}$ обозначим $\mathfrak{M}(\mathcal{H})$. Для данного гильбертова пространства $\mathcal{H}$, множество $\mathfrak{S}(\mathcal{H})$ квантовых состояний и множество $\mathfrak{M}(\mathcal{H})$ квантовых наблюдаемых вместе с обобщенным статистическим правилом Борна - фон Неймана (2.13) по построению удовлетворяют аксиомам отделимой статистической модели. Кроме того, эта модель имеет дополнительную структуру в множестве наблюдаемых $\mathfrak{M}(\mathcal{H})$, аналогичную выпуклой структуре $\mathfrak{S}(\mathcal{H})$.

Пусть $\{M^j\}$ – конечный набор наблюдаемых с одним и тем же пространством 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

$$M_x = \sum_j p_j M_x^j; \quad x \in \mathcal{X}. \quad (2.26)$$

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.

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

Теорема 2.2.4 *Всякая четкая наблюдаемая $M \in \mathfrak{M}_{\mathcal{X}}$ есть крайняя точка выпуклого множества $\mathfrak{M}_{\mathcal{X}}$. Обратно, всякая крайняя точка множества $M \in \mathfrak{M}_{\mathcal{X}}$ 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 (2.7)

$$pM_x^1(I - M_x^1) + (1-p)M_x^2(I - M_x^2) + p(1-p)(M_x^1 - M_x^2)^2 = 0. \quad (2.27)$$

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

Теперь заметим, что всегда $M_x \leq I$ и, значит, $M_x^2 \leq M_x$. Зафиксировав некоторый исход $x \in \mathcal{X}$, имеем

$$M_x = \frac{1}{2}M_x^2 + \frac{1}{2}(2M_x - M_x^2); \quad (2.28)$$

$$M_{x'} = \frac{1}{2}M_{x'}(I + M_x) + \frac{1}{2}M_{x'}(I - M_x), \quad x' \neq x. \quad (2.29)$$

Если $[M_x, M_{x'}] \equiv 0$, то эти соотношения представляет M как выпуклую комбинацию (с равными весами) двух других наблюдаемых. Если M – крайняя точка, то они должны совпасть с M, в частности, $M_x^2 = M_x$, значит, M – четкая наблюдаемая. □

Крайние точки множества наблюдаемых будем называть *экстремальными* наблюдаемыми. Из теоремы 2.2.4 следует, что в классическом случае экстремальные наблюдаемые совпадают с четкими, давая им очень понятную характеризацию как наблюдаемых без стохастичности в процедуре измерения. В квантовой статистической модели все не так просто (и поэтому более интересно). Множество крайних точек квантовых наблюдаемых исчерпывается четкими наблюдаемыми только в случае двух исходов измерения (они играют особую роль в различных аксиоматических подходах; разные авторы называют такие наблюдаемые *предложениями, вопросами, эффектами*; 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 $P = M_0$. Переходя к наблюдаемым с более чем двумя значениями, рассмотрим следующую конструкцию.

Определение 2.2.4 *Пусть $\{|\psi_x\rangle\}$ – произвольный набор векторов (которые могут быть ненормированными) в $\mathcal{H}$, такой, что $\sum_x |\psi_x\rangle\langle\psi_x| = I$. Такой набор называется переполненной системой.*

Если $\{|\psi_x\rangle\}$ – переполненная система, то произвольный вектор $|\psi\rangle \in \mathcal{H}$ может быть представлен в виде

$$|\psi\rangle = \sum_x c_x |\psi_x\rangle, \quad \text{где} \quad c_x = \langle\psi_x|\psi\rangle, \tag{2.30}$$

где коэффициенты $\{c_x\}$ могут определяться неединственным образом, так как векторы $|\psi_x\rangle$ могут быть линейно зависимы. Для операторов имеет место аналогичное матричное представление. В d-мерном комплексном пространстве переполненная система с более чем d векторами всегда существует, и может быть построена из любой полной (возможно, линейно зависимой) системы следующим образом. Пусть $\{|\phi_x\rangle\}$ – полная система векторов в $\mathcal{H}$. Соответствующий *оператор Грама* определяется как

$$G = \sum_x |\phi_x\rangle\langle\phi_x|. \tag{2.31}$$

Полнота означает, что оператор G невырожден. Система векторов $|\psi_x\rangle = G^{-1/2}|\phi_x\rangle$ является переполненной, так как

$$\sum_x |\psi_x\rangle\langle\psi_x| = G^{-1/2} \underbrace{\sum_x |\phi_x\rangle\langle\phi_x|}_{=G} G^{-1/2} = I. \tag{2.32}$$

Переполненная система определяет квантовую наблюдаемую M с компонентами

$$M_x = |\psi_x\rangle\langle\psi_x|, \quad x \in \mathcal{X}. \tag{2.33}$$

В частности, для любого ортонормированного базиса $\{e_x\}$, наблюдаемая $M = \{|e_x\rangle\langle e_x|\}$ является четкой и, следовательно, экстремальной наблюдаемой.

Теорема 2.2.5 *Наблюдаемая (2.33) является экстремальной тогда и только тогда, когда операторы M_x линейно независимы.*

Доказательство. Пусть M – крайняя точка и предположим, что

$$\sum_x c_x |\psi_x\rangle\langle\psi_x| = 0. \tag{2.34}$$

Взяв достаточно малое $\epsilon > 0$, определим

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

Тогда $M^{\pm}$ являются наблюдаемыми и, по построению, $M = \frac{1}{2}M^+ + \frac{1}{2}M^-$. Но M – крайняя точка, значит, $M_x^+ = M_x^- = M_x$. Итак, из (2.34) следует $c_x = 0$, т. е. компоненты M линейно независимы.

Обратно, пусть

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

– разложение M, тогда $0 \leq pM_x^1 \leq |\psi_x\rangle\langle\psi_x|$. Используя лемму 2.2.3, получаем

$$M_x^1|\psi_x\rangle\langle\psi_x| = |\psi_x\rangle\langle\psi_x|M_x^1 = \langle\psi_x|\psi_x\rangle M_x^1,$$

откуда $M_x^1 = \lambda_x|\psi_x\rangle\langle\psi_x|$ 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. □

В силу результата задачи 1.4.2, размерность вещественного пространства эрмитовых операторов равна d^2. Итак, все переполненные системы с n линейно независимыми компонентами, $d < n \leq d^2$ являются нечеткими экстремальными наблюдаемыми. Конкретный пример такой наблюдаемой при $n = 2, d = 3$ будет рассмотрен в следующем разделе.

2.3 Оптимальное различение квантовых состояний

2.3.1 Постановка задачи

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

Пусть квантовая система приготовляется в одном из состояний S_x, $x = 1, \ldots, n$. Над системой можно производить произвольное измерение. Требуется найти оптимальную наблюдаемую, позволяющую наилучшим образом выяснить, в каком из этих априорно данных состояний находится система. Такая постановка задачи характерна для математической статистики и ее применений в теории связи (см. раздел 3.2.1). В высокоточных экспериментах и в квантовой оптике исследователи уже способны оперировать элементарными квантовыми системами, такими как одиночные ионы, атомы и фотоны. В обсуждаемых предложениях квантовых вычислений информация записывается в состояния элементарных квантовых ячеек - q-битов, а считывается при помощи квантовых измерений. Со статистической точки зрения, измерение дает оценку квантового состояния — либо всего состояния целиком, либо некоторых его параметров, при этом возникает вопрос наиболее точного оценивания.

Статистика измерения описывается квантовой наблюдаемой, т. е. разложением единицы $M = \{M_x\}$ в пространстве системы $\mathcal{H}$. Вероятность принять решение y, при условии, что система находилась в состоянии S_x, равна $p_M(y|x) = \operatorname{Tr} S_x M_y$. При этом вероятность того, что будет принято

правильное решение, равна $p_M(x|x)$. Примем дополнительное предположение, что состояния S_x появляются с вероятностями π_x (например, в случае равновероятных состояний $\pi_x = 1/n$.) Тогда средняя вероятность правильного решения

$$\mathcal{P}\{M\} = \sum_{x=1}^{n} \pi_x p_M(x|x),$$

и задача состоит в ее максимизации.

2.3.2 Оптимальные наблюдаемые

Вероятность правильного решения

$$\mathcal{P}\{M\} = \sum_{x=1}^{n} \operatorname{Tr} W_x M_x,$$

где $W_x = \pi_x S_x \geq 0$, является аффинной функцией,

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

определенной на выпуклом множестве наблюдаемых

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

Максимизация аффинной функции, заданной на выпуклом множестве – это типичная задача линейного программирования.

Теорема 2.3.1 *Средняя вероятность правильного решения $\mathcal{P}\{M\}$ достигает максимума в крайней точке множества $\mathfrak{M}_n$. Наблюдаемая M^0 оптимальна тогда и только тогда, когда найдется эрмитов оператор Λ^0 такой, что*

i. $(\Lambda^0 - W_y)M_y^0 = 0$;
ii. $\Lambda^0 \geq W_y$.

При этом имеет место соотношение двойственности

$$\max\{\mathcal{P}\{M\} : M \in \mathfrak{M}_n\} = \min\{\operatorname{Tr} \Lambda : \Lambda \geq W_y, y = 1, \ldots, n\}. \qquad (2.35)$$

Доказательство. Поскольку $\mathcal{P}\{M\}$ является непрерывной аффинной функцией на компактном выпуклом множестве $\mathfrak{M}_n$, первое утверждение вытекает из результата задачи 1.5.1.

Докажем достаточность условий i, ii. Пусть Λ^0 – эрмитов оператор, удовлетворяющий условиям i, ii. Используя условие i и свойство (1.13), получаем

$$\mathcal{P}\{M\} = \mathrm{Tr}\sum_x W_x M_x \le \mathrm{Tr}\sum_x \Lambda^0 M_x = \mathrm{Tr}\,\Lambda^0. \tag{2.36}$$

Из условия ii получаем $\Lambda^0 M_x^0 = W_x M_x^0$. Суммируя по x и беря след, получаем

$$\mathrm{Tr}\,\Lambda^0 = \mathrm{Tr}\,\Lambda^0 \sum_x M_x^0 = \mathrm{Tr}\sum_x W_x M_x^0 = \mathcal{P}\{M^0\}. \tag{2.37}$$

Итак

$$\mathcal{P}\{M\} \le \mathcal{P}\{M^0\}, \quad \text{for all}\, M. \tag{2.38}$$

. Отметим, что

$$\Lambda^0 = \sum_x W_x M_x^0 = \sum_x M_x^0 W_x.$$

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

$$\mathrm{Tr}\sum_y W_y X_y^2 - \mathrm{Tr}\,\Lambda(\sum_y X_y^2 - I), \tag{2.39}$$

где Λ эрмитов оператор, по всевозможным наборам эрмитовых операторов X_y. Здесь оператор Λ представляет совокупность вещественных множителей Лагранжа для данной задачи. Пусть X_y^0 оптимальный набор, положим $X_y = X_y^0 + \epsilon Y_y$, и рассмотрим (2.39) как функцию от ϵ. Рассматривая коэффициенты при ϵ и ϵ^2, получаем условия

$$\mathrm{Tr}[(W_y - \Lambda)X_y^0 + X_y^0(W_y - \Lambda)]Y_y = 0,$$

$$\mathrm{Tr}(W_y - \Lambda)Y_y^2 \le 0$$

для произвольных эрмитовых Y_y, т.е.

$$(W_y - \Lambda)X_y^0 + X_y^0(W_y - \Lambda) = 0, \quad \Lambda - W_y \ge 0.$$

Второе неравенство есть условие ii теоремы. Полагая $M_y^0 = (X_y^0)^2$, получаем из первого соотношения $\mathrm{Tr}(\Lambda - W_y)M_y^0 = 0$, что вместе со вторым неравенством влечет условие i. □

Задача 2.3.1 *Доказать, что операторный множитель Лагранжа Λ является единственным решением двойственной задачи в правой части (2.35).*

Проиллюстрируем значение и применение этих условий на примерах.

Пример 2.3.1

Рассмотрим классический случай, когда все операторы плотности S_x, а значит, и операторы W_x, коммутируют и, следовательно, найдется общий ортонормированный базис, в котором все они диагональны,

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

Тогда двойственная задача имеет решение

$$\Lambda^0 = \sum_\omega \max_x W_x(\omega)|\omega\rangle\langle\omega|,$$

где $\max\limits_x W_x(\omega)$ — верхняя огибающая функций $W_x(\omega)$; $x = 1, \ldots, n$. Одно из решений исходной задачи дается формулой

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

где $\mathbf{1}_{\Omega_x}$ обозначает индикатор подмножества Ω_x, причем непересекающиеся подмножества $\Omega_x \subset \{\omega : \Lambda^0(\omega) = W_x(\omega)\}$ образуют разбиение множества $\Omega = \{\omega\}$.

Это сводится к принципу *максимального правдоподобия* в классической статистике: решение x необходимо принимать для тех ω, для которых апостериорный выигрыш $W_x(\omega)$ максимален. 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 2.2.4).

Пример 2.3.2

Различение двух квантовых состояний. Пусть S_0, S_1 два оператора плотности, π_0, π_1 – их априорные вероятности. Наблюдаемая дается разложением единицы $\{M_0, M_1\}$, так что $M_0 + M_1 = I$. Этот случай можно свести к предыдущему примеру, полагая

$$W_0 = \pi_1 S_1 + W_0', \quad W_1 = \pi_1 S_1 + W_1',$$

где операторы $W_0' = \pi_0 S_0 - \pi_1 S_1, W_1' = 0$ коммутируют. Таким образом,

$$\mathcal{P}\{M\} = \pi_1 + \operatorname{Tr}\left(W_0' M_0 + W_1' M_1\right), \tag{2.40}$$

и

$$\Lambda' = \max\{\pi_0 S_0 - \pi_1 S_1, 0\} = (\pi_0 S_0 - \pi_1 S_1)_+ \tag{2.41}$$

Всякий оптимальный оператор M_0^0 имеет вид

$$M_0^0 = \mathbf{1}_{(0,\infty)}(\pi_0 S_0 - \pi_1 S_1) + X_0, \tag{2.42}$$

где первое слагаемое является проектором на собственное подпространство оператора $\pi_0 S_0 - \pi_1 S_1$, отвечающее положительным собственным значениям, а второе слагаемое имеет носитель на нулевом подпространстве оператора $\pi_0 S_0 - \pi_1 S_1$, причем $0 \leq X_0 \leq I$. Максимальная вероятность правильного решения равна

$$\mathcal{P}\{M\} = \pi_1 + \mathrm{Tr}\,(\pi_0 S_0 - \pi_1 S_1)_+ . \tag{2.43}$$

Используя соотношение

$$(\pi_0 S_0 - \pi_1 S_1)_+ = \frac{1}{2}|\pi_0 S_0 - \pi_1 S_1| + \frac{1}{2}(\pi_0 S_0 - \pi_1 S_1),$$

получаем

$$\max \mathcal{P}\{M\} = \frac{1}{2}(1 + \|\pi_0 S_0 - \pi_1 S_1\|_1). \tag{2.44}$$

В частности, при $\pi_0 = \pi_1 = \frac{1}{2}$ различимость состояний S_0 и S_1 определяется величиной ядерной нормы разности соответствующих операторов плотности.

Предложение 2.3.2 *Пусть $S_0 = |\psi_0\rangle\langle\psi_0|, S_1 = |\psi_1\rangle\langle\psi_1|$ – чистые состояния, тогда максимум величины $\mathcal{P}\{M\}$ равен*

$$\max_M \mathcal{P}\{M\} = \frac{1}{2}\left(1 + \sqrt{1 - 4\pi_0\pi_1|\langle\psi_0|\psi_1\rangle|^2}\right). \tag{2.45}$$

В частности, при $\pi_0 = \pi_1 = 1/2$,

$$\max_M \mathcal{P}\{M\} = \frac{1}{2}\left(1 + \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}\right). \tag{2.46}$$

Доказательство. Рассмотрим задачу на собственные значения для оператора ранга 2 $\pi_0|\psi_0\rangle\langle\psi_0| - \pi_1|\psi_1\rangle\langle\psi_1|$. Этот оператор может иметь нулевые собственные значения, отвечающие ортогональному дополнению к носителю $\mathcal{L}$, который натянут на векторы $|\psi_0\rangle, |\psi_1\rangle$. Существенная составляющая оптимальной наблюдаемой также сосредоточена на $\mathcal{L}$, поэтому можно ограничиться поиском собственных векторов вида

$$|\psi\rangle = c_0|\psi_0\rangle + c_1|\psi_1\rangle.$$

Подстановка в уравнение для собственных векторов дает

$$\pi_0(c_0 + \langle\psi_0|\psi_1\rangle c_1) = \lambda c_0; \tag{2.47}$$
$$-\pi_1(\langle\psi_1|\psi_0\rangle c_0 + c_1) = \lambda c_1, \tag{2.48}$$

откуда находим собственные значения

$$\lambda_{0,1} = \frac{1}{2}\left[\pi_0 - \pi_1 \pm \sqrt{1 - 4\pi_0\pi_1|\langle\psi_0|\psi_1\rangle|^2}\right]$$

с соответствующим базисом собственных векторов $|e_0\rangle, |e_1\rangle$. Следовательно

$$\|\pi_0|\psi_0\rangle\langle\psi_0| - \pi_1|\psi_1\rangle\langle\psi_1|\|_1 = \lambda_0 - \lambda_1 = \sqrt{1 - 4\pi_0\pi_1|\langle\psi_0|\psi_1\rangle|^2},$$

откуда следует (2.45). □

Оптимальная четкая наблюдаемая в подпространстве $\mathcal{L}$ имеет вид $\{|e_0\rangle\langle e_0|, |e_1\rangle\langle e_1|\}$. Если $\pi_0 = \pi_1 = 1/2$, то оптимальный базис расположен симметрично по отношению к $|\psi_0\rangle, |\psi_1\rangle$ (рис. 2.2).

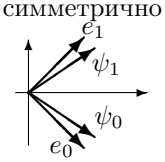


Рис. 2.2. *Различение двух чистых состояний*

Пример 2.3.3

На плоскости (рассматриваемой как вещественное подпространство двумерного комплексного пространства) рассмотрим "равноугольную" конфигурацию трех векторов (рис. 2.3)

$$|\psi_j\rangle = \begin{bmatrix} \cos\frac{2j\pi}{3} \\ \sin\frac{2j\pi}{3} \end{bmatrix}, \quad j = 0, 1, 2. \tag{2.49}$$

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

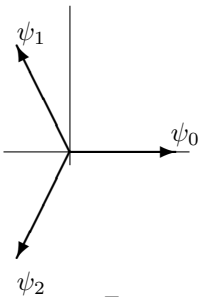


Рис. 2.3. *Три равноугольные состояния*

Имеем

$$S_j = \frac{1}{2}\left(I + \begin{bmatrix} \cos\frac{4j\pi}{3} & \sin\frac{4j\pi}{3} \\ \sin\frac{4j\pi}{3} & -\cos\frac{4j\pi}{3} \end{bmatrix}\right), \tag{2.50}$$

так что

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

в силу того, что $\sum_{j=0}^{2} e^{i\frac{4\pi j}{3}} = 0$.

Отсюда следует, что $M_k^0 = \frac{2}{3}S_k; k = 0, 1, 2$, является разложением единицы, отвечающим переполненной системе векторов $\sqrt{\frac{2}{3}}|\psi_k\rangle; k = 0, 1, 2$. Соответствующая наблюдаемая экстремальна, поскольку выполнено условие теоремы 2.2.5.

Покажем, что в случае равновероятных состояний, $\pi_j = 1/3$, $\{M_k^0\}$ является оптимальной наблюдаемой. Проверим условия теоремы 2.3.1. Поскольку $S_j^2 = S_j$, то

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

так что выполнено условие ii: $I/3 = \Lambda^0 \geq S_j/3$. Условие i также выполнено, поскольку

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

Итак, $\max \mathcal{P}\{M\} = \operatorname{Tr} \Lambda^0 = 2/3$. Найдем теперь максимум по всевозможным четким наблюдаемым с тремя исходами. Нетривиальное ортогональное разложение единицы с тремя компонентами в двумерном пространстве имеет вид $M_0 = |e_0\rangle\langle e_0|$, $M_1 = |e_1\rangle\langle e_1|$, $M_2 = 0$, где $|e_0\rangle, |e_1\rangle$, – произвольный базис. Поэтому задача сводится к оптимальному различению двух равновероятных состояний S_0, S_1. Применяя соотношение (2.46) для случая $\langle \psi_0 | \psi_1 \rangle = -\frac{1}{2}$, получаем

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

Таким образом, использование в квантовой статистике нечетких наблюдаемых может привести к выигрышу при различении состояний исходной системы по сравнению с четкими наблюдаемыми. Подчеркнем, что в аналогичном классическом случае никакая рандомизация не может улучшить качество процедуры различения состояний. С геометрической точки зрения, причина, конечно, состоит в том, что в квантовом случае не все экстремальные наблюдаемые (среди которых и находится оптимальная) являются четкими.

2.4 Комментарии

1. Из текстов, посвященных математическим основаниям квантовой механики, выделим монографии фон Неймана [21], Макки [16], Сигала [27], Фаддеева и Якубовского [29]. Обзор аксиоматических подходов имеется в статье Вайтмана [154], а также в Замечаниях к гл. VIII книги Рида и Саймона [24]. Желанной целью любой аксиоматики является вывод формализма гильбертова пространства; такой вывод в рамках операционального подхода (см. ниже) в конечномерном варианте осуществлен в работе Араки [53]. Впрочем, это является отдельной темой; для наших целей достаточно просто постулировать множество состояний $\mathfrak{S} = \mathfrak{S}(\mathcal{H})$. Описание смешанных состояний операторами (матрицами) плотности было

предложено независимо фон Нейманом, Вейлем и Ландау (см. примечание на с. 240 в книге [21]).

Аксиоматический подход, в основу которого положена двойственность между парой частично упорядоченных пространств, порождаемых выпуклым множеством квантовых состояний и порядковым интервалом 0-1 наблюдаемых (тестов) разрабатывался школой Людвига [128], [84], см. также Дэвис [75] (подобный подход часто называют *операциональным*). Подробное исследование статистической структуры квантовой теории на основе операционального подхода предпринято в книгах Холево [37], [36], где можно найти также подробную библиографию.

В предложении 2.1.1 нам пришлось предположить существование максимальной наблюдаемой только потому, что для простоты мы с самого начала ограничились конечными множествами исходов. При надлежащем обобщение существование максимальной наблюдаемой (с вообще говоря бесконечным пространством Ω) может быть выведено из совместимости всех наблюдаемых модели. По существу, об этом говорит известная теорема Колмогорова о построении вероятностного пространства для случайного процесса, заданного системой согласованных (т. е. совместимых) конечномерных распределений. Подробное исследование отношения стохастической совместимости, включающее доказательство того факта, что отделимая модель, все наблюдаемые которой стохастически совместимы, вкладывается в модель Вальда, имеется в работе Холево [99].

Изложение основных понятий для нужд квантовой теории информации имеется в книгах Нильсена и Чанга [22], Хайаши [92]. В работе Фукса [85] описана информационно-теоретическая точка зрения на основания квантовой теории.

2. Излагаемая здесь концепция квантовой наблюдаемой разрабатывалась Людвигом [128], Дэвисом и Льюисом [75], Холево [37]. По поводу выпуклой структуры множества наблюдаемых см. Краус [117].

Спин квантовой частицы органически связан с представлениями группы вращений трехмерного пространства [3], [16]. Результаты, связанные с теоремой 2.2.5 см. в работе Дэвиса [74].

3. Задача различения двух чистых квантовых состояний была впервые рассмотрена в 1968 г. в работе П. А. Бакута и С. С. Щурова, см. книгу [14]. Подобные задачи естественно возникают при обнаружении слабых источников света, а также в квантовой оптике. Общая теория обнаружения и оценивания для квантовых состояний была разработана в 1970-е гг. в трудах Хелстрома, Юна, Холево, Стратоновича, Белавкина, см. монографии [30], [37]. Новый интерес к теории оценивания квантовых состояний появился в связи с идеями квантовых вычислений: всякое такое вычисление завершается измерением параметров конечного состояния квантового компьютера, которое должно быть максимально точным. Результаты современной теории оценивания квантовых состо-

яний, включая асимптотический подход, излагаются в книгах Хайаши [92], Петца [133]. Байесовская задача является задачей линейного программирования (см. [17], [25]) и теорема 2.3.1 может быть доказана на основе соответствующей теоремы двойственности. Пример 2.3.3 с тремя равноугольными состояниями предложен Холево [33].

3. Составные квантовые системы и сцепленность

3.1 Составные системы

3.1.1 Тензорное произведение

Своеобразие и новые возможности квантовой теории информации в значительной мере обусловлены свойствами составных квантовых систем. Пусть $\mathcal{H}_j$, $j = 1, 2$ – гильбертовы пространства двух квантовых систем со скалярными произведениями $\langle \cdot | \cdot \rangle_j$. Составная система описывается тензорным произведением гильбертовых пространств, которое может быть построено следующим образом.

Согласно разделу 1.1, элемент $\psi \in \mathcal{H}$ определяет антилинейную функцию $\psi(\phi) = \langle \phi | \psi \rangle$ аргумента $\phi \in \mathcal{H}$; для двух элементов $\psi_j \in \mathcal{H}_j; j = 1, 2$, обозначим $\psi_1 \otimes \psi_2$ би-антилинейную функцию аргументов $\phi_1 \in \mathcal{H}_1, \phi_2 \in \mathcal{H}_2$, определенную соотношением

$$(\psi_1 \otimes \psi_2)(\phi_1, \phi_2) = \langle \phi_1 | \psi_1 \rangle_1 \langle \phi_2 | \psi_2 \rangle_2.$$

Рассмотрим векторное пространство $\mathcal{L}$ конечных линейных комбинаций таких функций $\sum_j c_j \psi_1^j \otimes \psi_2^j$. Введем скалярное произведение на $\mathcal{L}$, полагая

$$\langle \varphi_1 \otimes \varphi_2 | \psi_1 \otimes \psi_2 \rangle = \langle \varphi_1 | \psi_1 \rangle_1 \langle \varphi_2 | \psi_2 \rangle_2$$

на порождающих элементах и далее продолжая по линейности на $\mathcal{L}$.

Задача 3.1.1 *Покажите, что такое линейное продолжение возможно, оно единственно и удовлетворяет всем свойствам скалярного произведения.*

Пространство $\mathcal{L}$ с определенным выше скалярным произведением называется *тензорным произведением* $\mathcal{H}_1 \otimes \mathcal{H}_2$ гильбертовых пространств $\mathcal{H}_1, \mathcal{H}_2$.

Задача 3.1.2 *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* $\dim \mathcal{H}_1 \otimes \mathcal{H}_2 = \dim \mathcal{H}_1 \times \dim \mathcal{H}_2$.

Отсюда вытекает, что всякий вектор $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ однозначно записывается в виде

$$|\psi\rangle = \sum_{j,k} c_{jk} |e_1^j\rangle \otimes |e_2^k\rangle.$$

Производя суммирование по j и обозначая $|\psi_k\rangle = \sum_{j=1}^{d_1} c_{jk}|e_1^j\rangle \in \mathcal{H}_1$, получаем

$$|\psi\rangle = \sum_{k=1}^{d_2} |\psi_k\rangle \otimes |e_2^k\rangle,$$

так что в общем случае $\mathcal{H}_1 \otimes \mathcal{H}_2$ изоморфно прямой ортогональной сумме $d_2 = \dim \mathcal{H}_2$ слагаемых $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_1$.

Для операторов X_j в пространствах $\mathcal{H}_j$ зададим их тензорное произведение, полагая

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

и продолжая по линейности. Фиксируем базис в $\mathcal{H}_2$, так что $\mathcal{H}_1 \otimes \mathcal{H}_2$ представляется как $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_1$. Тогда $X_1 \otimes X_2$ можно представить в виде блочной матрицы $[X_1 x_2^{jk}]$, где $[x_2^{jk}]$ – матрица оператора X_2 в базисе $\{e_2^k\}$. Произвольный оператор X в $\mathcal{H}_1 \otimes \mathcal{H}_2$ задается блочной матрицей $[X_{jk}]$, элементы которой суть операторы в $\mathcal{H}_1$.

Напомним, что в разделе 1.4 $\mathfrak{B}_h(\mathcal{H})$ обозначалось вещественное линейное пространство всех эрмитовых операторов в $\mathcal{H}$, т. е. четких вещественных наблюдаемых. В силу результата задачи 1.4.2, в случае комплексных гильбертовых пространств $\mathcal{H}_1\,\mathcal{H}_2$

$$\dim \mathfrak{B}_h(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim \mathfrak{B}_h(\mathcal{H}_1) \cdot \dim \mathfrak{B}_h(\mathcal{H}_2)$$

тогда как в случае вещественных $\mathcal{H}_1\,\mathcal{H}_2$ для размерностей пространств симметричных операторов имеет место неравенство $>$.

Задача 3.1.3 *Если S_j – операторы плотности в $\mathcal{H}_j; j = 1, 2$, то $S_1 \otimes S_2$ – оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2$.*

Пусть T – оператор в $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. *Частичный след* оператора T (относительно $\mathcal{H}_2$), обозначаемый $\mathrm{Tr}_{\mathcal{H}_2} T$, определяется как оператор в $\mathcal{H}_1$, ассоциированный с формой

$$\langle\phi|(\mathrm{Tr}_{\mathcal{H}_2} T)|\psi\rangle = \sum_k \langle\phi \otimes e_2^k|T|\psi \otimes e_2^k\rangle, \quad \phi, \psi \in \mathcal{H}_1.$$

Задача 3.1.4 *Покажите, что определение корректно, т. е. не зависит от выбора ортонормированного базиса $\{e_2^k\}$. Если $T = T_1 \otimes T_2$, то $\mathrm{Tr}_{\mathcal{H}_2}(T_1 \otimes T_2) = (\mathrm{Tr}\, T_2)T_1$.*

Пусть $\mathcal{H}_1 \otimes \mathcal{H}_2$ реализовано как $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_1$, так что $T = [T_{jk}]$, тогда $\mathrm{Tr}_{\mathcal{H}_2} T = \sum_{j=1}^{d_2} T_{jj}$ и $\mathrm{Tr}_{\mathcal{H}_1} T = [\mathrm{Tr} T_{jk}]$.

3.1.2 Расширение Наймарка

Принципиальная связь между ортогональными и неортогональными разложениями единицы устанавливается следующей теоремой.

Теорема 3.1.1 (Наймарк) *Пусть $\{M_x\}$ – разложение единицы в $\mathcal{H}$ с m компонентами, $\dim \mathcal{H} = d$. Тогда существует гильбертово пространство $\mathcal{K}$ размерности $\dim \mathcal{K} \leq md$, изометрическое отображение $V : \mathcal{H} \to \mathcal{K}$ и ортогональное разложение единицы $\{E_x\}$ в $\mathcal{K}$, такие что*

$$M_x = V^* E_x V, \quad x \in \mathcal{X}. \tag{3.1}$$

Изометрическое отображение (изометрия) – это линейное отображение V, сохраняющее скалярное произведение векторов в гильбертовых пространствах. Для любых $|\phi\rangle, |\psi\rangle \in \mathcal{H}$ выполняется $\langle V\phi|\, V\psi\rangle = \langle \phi|\, \psi\rangle$, так что изометрия характеризуется равенством $V^*V = I$. Изометрическое вложение V позволяет отождествить $\mathcal{H}$ с подпространством $V\mathcal{H}$ пространства $\mathcal{K}$ и считать, что $\mathcal{H} \subset \mathcal{K}$. Тогда M_x можно рассматривать просто как ограничение E_x на $\mathcal{H}$:

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

Доказательство. Построение $\mathcal{K}$ будет осуществлено в два этапа. Сначала определим векторы $|\Psi\rangle \in \mathcal{H}_m$ формулой

$$|\Psi\rangle = \begin{bmatrix} |\psi_1\rangle \\ \vdots \\ |\psi_m\rangle \end{bmatrix}, \qquad |\psi_i\rangle \in \mathcal{H}. \tag{3.2}$$

Определим также псевдоскалярное произведение, т. е. форму, удовлетворяющую всем свойствам скалярного произведения, кроме невырожденности, соотношением

$$\langle \Psi'|\Psi\rangle = \sum_x \langle \psi'_x | M_x | \psi_x \rangle. \tag{3.3}$$

Соответствующая квадратичная форма может быть вырождена.

Для того, чтобы обеспечить невырожденность, определим $\mathcal{K}$ как фактор-пространство $\mathcal{K} = \mathcal{H}_m/\mathcal{H}_0$, где $\mathcal{H}_0 = \{\Psi_0 \in \mathcal{H}_m : \langle \Psi_0|\Psi_0\rangle = 0\}$, т. е. объединим векторы, разности которых имеют нулевую норму. Теперь отображение $V : \mathcal{H} \to \mathcal{H}_m$, определяемое как

$$V|\psi\rangle = \begin{bmatrix} |\psi\rangle \\ \vdots \\ |\psi\rangle \end{bmatrix}, \tag{3.4}$$

является изометрией, так как

$$\langle\psi|V^*V|\psi\rangle = \sum_{j=1}^{m}\langle\psi|M_x|\psi\rangle = \langle\psi|\psi\rangle. \tag{3.5}$$

В качестве ортогонального разложения единицы $\{E_x\}$ в $\mathcal{K}$ рассмотрим $E_x|\Psi\rangle = [0,\ldots,|\psi_x\rangle,\ldots,0]^\top$, где единственная ненулевая компонента находится на x-м месте. Таким образом,

$$\langle V\phi|E_x|V\psi\rangle = \langle\phi|M_x|\psi\rangle\ , \quad \phi,\psi\in\mathcal{H},$$

что завершает доказательство. □

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

Следствие 3.1.1 *Пусть $\{M_x\}$ – наблюдаемая в $\mathcal{H}$, тогда найдется гильбертово пространство $\mathcal{H}_0$, единичный вектор $\psi_0\in\mathcal{H}_0$ и четкая наблюдаемая $\{E_x\}$ в $\mathcal{H}\otimes\mathcal{H}_0$, такие, что*

$$M_x = \mathrm{Tr}_{\mathcal{H}_0}(I\otimes|\psi_0\rangle\langle\psi_0|)E_x. \tag{3.6}$$

Доказательство. В соответствии с теоремой Наймарка, $M_x = V^*E_xV$, где $V:\mathcal{H}\to\mathcal{K}$ – изометрическое вложение. Отождествим $\mathcal{H}$ с подпространством $\mathcal{K}$. Расширяя, если необходимо, пространство $\mathcal{K}$, можно считать, что $\dim\mathcal{K} = \dim\mathcal{H}\cdot d_0$, и значит

$$\mathcal{K}\simeq\mathcal{H}\oplus\cdots\oplus\mathcal{H}\simeq\mathcal{H}\otimes\mathcal{H}_0,$$

где $\mathcal{H}_0 = \ell^2_{d_0}$ – координатное гильбертово пространство размерности d_0, причем $\mathcal{H}$ отождествляется с первым слагаемым в прямой сумме, или с подпространством $\mathcal{H}\otimes|\psi_0\rangle$, где $|\psi_0\rangle = [1,0,\ldots,0]^\top$. Тогда

$$(I\otimes|\psi_0\rangle\langle\psi_0|)E_x = \begin{bmatrix} M_x & \cdots \\ 0 & 0\end{bmatrix},$$

так что (3.6) действительно выполнено. □

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

В классической статистике рандомизация, т. е. добавление "рулетки", хотя и может оказаться полезным приемом (например, в теории игр), не увеличивает информации о состоянии наблюдаемой системы. Из результатов раздела 2.3.2 следует, что в квантовой статистике это не всегда верно: парадоксальным образом, квантовая рандомизация позволяет извлечь больше информации о состояниях наблюдаемой системы, нежели содержится в четких наблюдаемых, не использующих независимой вспомогательной системы.

3.1.3 Разложение Шмидта и очищение

Рассмотрим состояние S_{12} составной системы в гильбертовом пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Определение 3.1.1 *Чистое состояние S_{12} называется сцепленным, если оно не представимо в виде тензорного произведения $S_1 \otimes S_2$.*

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

$$|\psi_{12}\rangle = \frac{1}{\sqrt{d}} \sum_{j=1}^{d} |e_j^1\rangle \otimes |e_j^2\rangle \tag{3.7}$$

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 $\mathcal{H}_{1,2}$. Смысл этого термина будет разъяснен позднее в разделе 7.4, где понятие сцепленности будет также обобщено на произвольные состояния составной системы.

Мы будем часто использовать следующий результат:

Теорема 3.1.2 (Разложение Шмидта) *Пусть $S_{12} = |\psi\rangle\langle\psi|$ – чистое состояние в гильбертовом пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2$, и пусть $S_1 = \mathrm{Tr}_{\mathcal{H}_2} S_{12}$, $S_2 = \mathrm{Tr}_{\mathcal{H}_1} S_{12}$ – частичные состояния. Тогда S_1 и S_2 имеют одни и те же ненулевые собственные значения λ_j. Более того,*

$$|\psi\rangle = \sum_j \sqrt{\lambda_j} |e_j^1\rangle \otimes |e_j^2\rangle, \tag{3.8}$$

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

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

$$|\psi\rangle = \sum_j |e_j^1\rangle \otimes |h_j^2\rangle, \tag{3.9}$$

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

$$\sum_{j,k}\langle h_j^2|h_k^2\rangle|e_k^1\rangle\langle e_j^1|\sum_j \lambda_j|e_j^1\rangle\langle e_j^1| \equiv S_1, \tag{3.10}$$

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 3.1.3 (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 $|\psi\rangle\langle\psi| \in \mathcal{H}_1 \otimes \mathcal{H}_2$, такие, что $S_1 = \mathrm{Tr}_{\mathcal{H}_2}|\psi\rangle\langle\psi|$.*

Для любого чистого состояния $|\psi'\rangle\langle\psi'| \in \mathcal{H}_1 \otimes \mathcal{H}_2$, 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 (3.8) с произвольным базисом $\{e_j^2\}$ в гильбертовом пространстве $\mathcal{H}_2$, изоморфном $\mathcal{H}_1$. Любой другой вектор $|\psi'\rangle$ имеет разложение (3.8) с другим базисом в $\mathcal{H}_2$. Остается заметить, что любые два базиса в гильбертовом пространстве связаны унитарным преобразованием. □

Следствие 3.1.2 *Пусть S_{12} – состояние в $\mathcal{H}_1 \otimes \mathcal{H}_2$, такое, что частичное состояние S_1 является чистым. Тогда $S_{12} = S_1 \otimes S_2$.*

Доказательство. Для чистого состояния S_{12} утверждение очевидным образом следует из теоремы 3.1.2; для произвольного состояния S_{12} мы можем применить аналогичное рассуждение к его очищению. □

Из теоремы 3.1.2 следует, что очищение в существенном единственно: любые два очищения с пространствами $\mathcal{H}_2$, $\mathcal{H}_2'$, такими что $\dim \mathcal{H}_2 \leq \dim \mathcal{H}_2'$, связаны изометрическим вложением $\mathcal{H}_2$ в $\mathcal{H}_2'$.

Замечательно, что существует одно универсальное гильбертово пространство, содержащее очищения всех состояний в $\mathcal{H}$. Рассмотрим гильбертово пространство $L^2(\mathcal{H})$ всех операторов $X, Y, ...$ в $\mathcal{H}$, снабженное скалярным произведением

$$(X, Y) = \mathrm{Tr}X^*Y.$$

Тогда линейное соответствие

$$|\psi\rangle\langle\varphi| \leftrightarrow |\psi\rangle \otimes \langle\varphi|, \tag{3.11}$$

где $\langle\varphi| \in \mathcal{H}^*$ (двойственное пространство линейных функций на $\mathcal{H}$), единственным образом продолжается до изометрии между гильбертовыми пространствами $L^2(\mathcal{H})$ и $\mathcal{H} \otimes \mathcal{H}^*$. Обозначая L_A (соответственно R_B) оператор левого умножения на A (соответственно правого умножения на B), мы получим для $X = |\psi\rangle\langle\varphi|$

$$L_A R_B X \equiv AXB \leftrightarrow |A\psi\rangle \otimes \langle B^*\varphi|.$$

Для произвольного оператора плотности S в $\mathcal{H}$ рассмотрим единичный вектор $\sqrt{S}W \in L^2(\mathcal{H})$, где W – произвольный унитарный оператор в $\mathcal{H}$. Тогда

$$\left(\sqrt{S}W, L_A R_B \sqrt{S}W\right) = \mathrm{Tr}\sqrt{S}A\sqrt{S}WBW^*,$$

Следовательно, полагая $B = I$, получаем

$$\mathrm{Tr}SA = \left(\sqrt{S}W, (A \otimes I_{\mathcal{H}^*})\sqrt{S}W\right).$$

Поэтому $\sqrt{S}W$ есть очищение S в пространстве $L^2(\mathcal{H})$, отождествленном с $\mathcal{H} \otimes \mathcal{H}^*$ через соответствие (3.11); более того, все очищения S получаются таким образом.

3.1.4 Парадокс Эйнштейна-Подольского-Розена. Неравенство Белла

Существование сцепленности, наряду с дополнительностью, является принципиальным структурным различием между классическим и квантовым описанием системам.

Чистое состояние классической системы задается точкой фазового пространства (практически – набором конкретных значений параметров, описывающих “внутренние свойства” системы). Если рассматриваемая классическая система состоит из нескольких подсистем, то всякое ее чистое состояние с необходимостью является произведением чистых состояний подсистем. Фиксируя значения параметров составной системы, мы тем самым фиксируем значения параметров всех ее подсистем. При этом нет необходимости прибегать к статистическому описанию.

В случае квантовых систем дело обстоит иначе. Как показывает следствие 3.1.2, для любого чистого сцепленного состояния S_{12} составной системы частичные состояния $S_{1,2}$ с необходимостью являются смешанными, т. е. требуют статистического описания.

Такой тип поведения составных систем совершенно необычен с классической точки зрения и, как будет показано далее в этом разделе, вообще несовместим с классическим способом описания (который в зарубежной литературе часто называется “реализмом”).

Ключевой пример необычного (с классической точки зрения) поведения составной квантовой системы рассмотрели Эйнштейн, Подольский и Розен (ЭПР) в 1935 году. В 50-х годах Бом представил пример в более рельефной форме, использующей спиновые степени свободы, а в 60-х Белл внес бóльшую логическую ясность, предложив фундаментальное неравенство, которое выполняется для составных классических систем, удовлетворяющих естественному дополнительному условию "локальности" (разделимости), но нарушается для квантовых систем, и в принципе может быть проверено экспериментально.

Рассмотрим две частицы со спином $1/2$, каждая из которых описывается гильбертовым пространством $\mathcal{H}$ с $\dim \mathcal{H} = 2$. В результате некоторого взаимодействия (или реакции распада) в начальный момент возникает совместное состояние их спинов, которое задается вектором

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

где базисные векторы

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

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

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

Каждое слагаемое здесь описывает состояние с разнонаправленными спинами, а $|\psi\rangle$ есть их суперпозиция, которую невозможно представить в виде тензорного произведения векторов состояний, относящихся к разным частицам. Такое состояние, называемое в физике "синглетом", является образцом сцепленного состояния двух квантовых систем.

Предположим, что частицы разлетаются на некоторое макроскопическое расстояние, при этом их спиновое состояние – синглет – сохраняется. Рассмотрим эксперимент, в котором в двух удаленных друг от друга лабораториях A и B над этими разлетевшимися частицами производятся одновременные измерения: наблюдаемой спина $\sigma(\boldsymbol{a})$ для одной частицы и $\sigma(\boldsymbol{b})$ для другой (рис. 3.1.4). Операторы $X = \sigma(\boldsymbol{a}) \otimes I, Y = I \otimes \sigma(\boldsymbol{b})$ коммутируют, следовательно соответствующие наблюдаемые совместимы и их ковариация дается выражением (2.24).

Задача 3.1.5 *Используя выражения для матричных элементов,*

$$\langle\uparrow|\sigma(\boldsymbol{a})|\uparrow\rangle = a_z, \quad \langle\downarrow|\sigma(\boldsymbol{a})|\uparrow\rangle = a_x + ia_y, \quad \langle\downarrow|\sigma(\boldsymbol{a})|\downarrow\rangle = -a_z, \quad (3.12)$$

вытекающие из (2.14), покажите, что в синглетном состоянии среднее значение и дисперсия каждой наблюдаемой равны

$$\mathsf{E}\sigma(\boldsymbol{a}) = 0, \quad \mathsf{D}\sigma(\boldsymbol{a}) = 1$$

а ковариация между спинами задается формулой

$$\langle\psi|\sigma(\boldsymbol{a}) \otimes \sigma(\boldsymbol{b})|\psi\rangle = -\boldsymbol{a} \cdot \boldsymbol{b}. \tag{3.13}$$

Отсюда следует, что если $\boldsymbol{b} = \boldsymbol{a}$, то коэффициент корреляции равен -1, и следовательно между исходами a, b измерений имеется детерминированная связь: $a = -b$. Из формул (3.12) и соотношения $\sigma(\boldsymbol{a})^2 = I$ следует

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

откуда

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

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

Таким образом, приходится выбирать между следующими альтернативами:

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

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

Задача 3.1.6 *Для любого направления* $\boldsymbol{a}$

$$|\psi\rangle = \frac{\epsilon}{\sqrt{2}}\Big[|\boldsymbol{a}\rangle \otimes |-\boldsymbol{a}\rangle - |-\boldsymbol{a}\rangle \otimes |\boldsymbol{a}\rangle\Big], \tag{3.14}$$

где ϵ *– несущественный множитель, по модулю равный единице.*

Более основательное рассмотрение мысленного эксперимента с синглетным состоянием приводит к более конкретному и глубокому выводу: если пытаться описывать корреляции двух спинов классически и в соответствии с принципом локальности, то оказывается невозможным достичь такого характера и уровня коррелированности, который соответствует предсказаниям квантовой теории. "Локальный реализм" и предсказываемая квантовой статистикой корреляция (3.13) несовместимы.

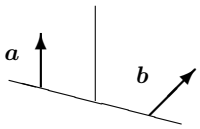


Рис. 3.1. *Направления спинов.*

Это доказывается с помощью **неравенства Белла-Клаузера-Хорна-Шимони**:

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

$$|\mathsf{E}X_1Y_1 + \mathsf{E}X_1Y_2 + \mathsf{E}X_2Y_1 - \mathsf{E}X_2Y_2| \le 2, \quad (3.15)$$

где E *–математическое ожидание, соответствующее распределению* P.

Доказательство получается усреднением по распределению P элементарного неравенства

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

которое в свою очередь следует из

$$|X_1Y_1 + X_1Y_2 + X_2Y_1 - X_2Y_2| \le |Y_1 + Y_2| + |Y_1 - Y_2| \le 2\max\{|Y_1|, |Y_2|\}.$$

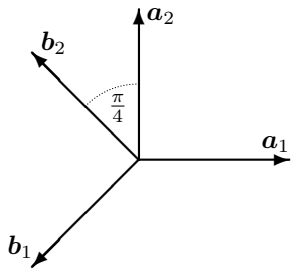


Рис. 3.2. *Выбор направлений* $\boldsymbol{a}_j$ *и* $\boldsymbol{b}_k$.

Вернемся теперь к системе из двух q-битов и рассмотрим четыре различных эксперимента, в которых наблюдаемые спина $\sigma(\boldsymbol{a}_j), (j=1,2)$ измеряются для первого q-бита, а $\sigma(\boldsymbol{b}_k), (k=1,2)$ – для второго; направления $\boldsymbol{a}_j, \boldsymbol{b}_k, (j,k=1,2)$ будут выбраны позже. Во всех четырех экспериментах система приготавливается в синглетном состоянии. Предположим, что существует "локальное" классическое представление для результатов всех этих четырех экспериментов. Это означает, что найдется вероятностное пространство Ω, распределение вероятностей P на Ω, описывающее статистический ансамбль в синглетном состоянии и случайные величины $X_j, Y_k; j,k=1,2$, принимающие значения ± 1, которые дают классическое описание наблюдаемых спина $\sigma(\boldsymbol{a}_j), \sigma(\boldsymbol{b}_k); j,k=1,2$. Тогда квантовые корреляции (3.13) должны удовлетворять неравенству (3.15). Условие "локальности" (а лучше сказать, *разделимости* подсистем) выражается в том, что случайные величины, описывающие спин первой системы (X_1 в случае первых двух корреляций и X_2 в двух других случаях) предполагаются одними и теми же для экспериментов, различающихся выбором измеряемого направления спина (Y_1 или Y_2) во второй системе (и наоборот), что и позволяет применить неравенство (3.15). Отказ от этого означал бы, что данное

классическое описание допускает влияние выбора измерения во второй системе на внутренние характеристики первой системы (термин "локальность" возникает в связи с тем, что системы предполагаются пространственно удаленными друг от друга).

Теперь выберем конкретные направления спинов, как показано на рисунке 3.2. Подставляя корреляции, даваемые формулой (3.13) в левую часть неравенства (3.15), получаем значение $2\sqrt{2}$, нарушающее это неравенство. Следовательно, исходное предположение о возможности классического "локального" описания является неверным. Условие разделимости кажется настолько естественным, что оно даже трудно уловимо, однако именно оно запрещает мгновенное влияние измерения, проводящегося в одной системе, на результаты измерений в другой системе. Если же от него отказаться, то интересующие нас четыре физические корреляции могут быть любыми величинами из отрезка $[-1, 1]$ и левая часть неравенства (3.15) будет ограничена лишь величиной 4, что не входит в противоречие с квантовой теорией.

Таким образом, возникает дилемма: либо квантовая теория дает неверные предсказания для корреляций, либо составная система их двух q-битов не имеет классического "локального" описания. Был проделан целый ряд экспериментов, начиная со знаменитых экспериментов Аспека (1981-1982), результаты которых свидетельствуют в пользу квантовой теории.

3.2 Квантовая система как носитель информации

3.2.1 Передача классической информации

Если носитель информации является классической системой с конечным числом состояний d, то максимальное количество двоичных единиц – бит, которое может быть записано либо передано с помощью такого носителя, равно, очевидно $\log d$ *. Если передача происходит без ошибок, то говорят об идеальном канале связи. В общем случае классический канал описывается условными вероятностями $p(y|x)$ получить сообщение y на выходе, если на вход послано сообщение x. Для идеального канала $p(y|x) = \delta_{xy}$; антиподом идеального канала является канал, для которого вероятности $p(y|x) = p(y)$ не зависят от посланного сообщения x и, следовательно, информация вообще не передается. Более подробно о классических каналах будет рассказано в гл. 4.

Рассмотрим теперь, как происходит передача классической информации посредством квантового носителя, описываемого гильбертовым пространством $\mathcal{H}$. Состояние носителя S приготовляется некоторыми макро-

* Далее $\log = \log_2$ обозначает двоичный логарифм. Чтобы избежать несущественных оговорок, мы будем игнорировать тот факт, что $\log d$ не обязательно является целым числом.

скопическими устройствами. Изменяя параметры этих устройств, экспериментатор изменяет состояние, и таким образом получает возможность "записывать" классические сообщения, содержащиеся в значениях параметров, в квантовом состоянии. Пусть имеется n разных сообщений и S_x – квантовое состояние, отвечающее сообщению с номером $x; x = 1, \ldots, n$. Отображение $x \to S_x$ описывает конечный результат физического процесса, порождающего состояние S_x. Подробное описание этого процесса не входит в задачу теории информации, которую интересуют лишь эти конечные состояния S_x.

Для того чтобы извлечь информацию о сообщении, содержащуюся в квантовом состоянии, необходимо произвести некоторое измерение. Если на выходе такого канала измеряется наблюдаемая $M = \{M_y\}$, то условная вероятность получить на выходе сообщение y, если на вход было послано сообщение x, равна

$$P_M(y|x) = \operatorname{Tr} S_x M_y. \tag{3.16}$$

Таким образом, для фиксированного измерения получается обычный классический канал связи. Предположим, что приготовление состояний и измерение удается организовать так, что $S_x = |e_x\rangle\langle e_x|, M_y = |e_y\rangle\langle e_y|$, тогда $P_M(y|x) = \delta_{xy}$, т. е. возникает идеальный классический канал, способный передать $\log d$ бит, где $d = \dim \mathcal{H}$. Далее в главе 5 будет установлено, что эта величина является верхней границей для количества классической информации, которое вообще может передано с помощью данного квантового носителя. Отсюда следуют важные выводы:

1) то обстоятельство, что гильбертово пространство содержит бесконечно много различных векторов чистых состояний, не помогает передать неограниченное количество информации; причина кроется в том, что чем больше состояний используется для передачи, тем они становятся ближе друг к другу, и, следовательно, более неразличимыми;

2) размерность гильбертова пространства является мерой максимального информационного ресурса квантовой системы.

3.2.2 Сцепленность и локальные операции

В этом разделе потребуются элементарные сведения об эволюциях квантовой системы. В дальнейшем, в главе 6 этот вопрос будет рассмотрен углубленно и с общих позиций. Пока же достаточно знать, что обратимые эволюции квантовой системы описываются унитарными операторами U: в результате такой эволюции вектор исходного чистого состояния ψ преобразуется в вектор $U\psi$. Соответственно, оператор плотности (смешанное состояние) S преобразуется в USU^*.

Рассмотрим теперь следующий вопрос. Нелокальный, с классической точки зрения, характер ЭПР-корреляций наводит на мысль попытаться использовать их для мгновенной передачи сообщений. Покажем, что

этого невозможно достичь, находясь в рамках квантовой механики (с точки зрения которой ЭПР-корреляции не противоречат локальности). Рассмотрим двух участников A и B, располагающих квантовыми системами 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. Участник A (передатчик) использует сообщения x, которые желательно передать участнику B (приемнику), для выполнения некоторых унитарных операций U_x (см. ниже) в своем пространстве $\mathcal{H}_A$ (такие операции называются *локальными*, т. е. действующими нетривиально только в одной из двух систем). 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, участник B может произвести локальное измерение произвольной наблюдаемой в $\mathcal{H}_B$, что соответствует разложению единицы вида $M = \{I_A \otimes M_y\}$ в пространстве составной системы $\mathcal{H}_A \otimes \mathcal{H}_B$. Результирующая переходная вероятность (3.16)

$$P_M(y|x) = \operatorname{Tr}(U_x \otimes I_B)S_{AB}(U_x \otimes I_B)^*(I_A \otimes M_y) = \operatorname{Tr} S_B M_y,$$

где $S_B = \operatorname{Tr}_{\mathcal{H}_A} S_{AB}$ – частичное состояние системы B, не зависит от x. Это означает, что локальные операции участника A никак не влияют на состояние B, а значит, информация действительно не передается.

Рассмотрим теперь другую ситуацию, в которой участник A производит локальное измерение наблюдаемой M_x^A и посылает исход своего измерения участнику B. В этом случае *апостериорное состояние* $S_B(x)$ участника B будет зависеть от исхода x. Чтобы найти апостериорные состояния, предположим, что B производит свое локальное измерение M_y^B и рассмотрим совместное распределение вероятностей исходов двух измерений

$$p_{x,y} = \operatorname{Tr} S_{AB}(M_x^A \otimes M_y^B).$$

Вводя условные вероятности $p(y|x) = \frac{p_{x,y}}{p_x}$ (в предположении, что $p_x > 0$), имеем

$$p_x = \operatorname{Tr} S_{AB}(M_x^A \otimes I_B) = \operatorname{Tr} S_A M_x^A,$$
$$p(y|x) = \operatorname{Tr} S_B(x) M_x^B,$$

где

$$S_B(x) = p_x^{-1} \operatorname{Tr}_{\mathcal{H}_A} S_{AB}(M_x^A \otimes I_B)$$

– апостериорные состояния.

Изменение состояния участника B от S_B к $S_B(x)$ означает переход от полного статистического ансамбля, который имеется в распоряжении B к подансамблю, который характеризуется значением x исхода измерения участника A. Если этот исход не учитывается (например, A вовсе не сообщает его B), то состояние B не изменяется

$$S_B = \sum_x p_x S_B(x).$$

Задача 3.2.1 *Пусть* $S_{AB} = |\psi\rangle\langle\psi|$ *– чистое сцепленное состояние,*

$$|\psi\rangle = \sum_x |e_x^A\rangle \otimes |\psi_x^B\rangle,$$

где e_x^A *– ортонормированный базис в* $\mathcal{H}_A$, $\sum_x \|\psi_x^B\|^2 = 1$. *Предположим, что* A *измеряет наблюдаемую* $M_x^A = |e_x^A\rangle\langle e_x^A|$ *и посылает исход* x *участнику* B. *Тогда* $p_x = \|\psi_x^B\|^2$, *и апостериорные состояния* B *суть*

$$S_B(x) = p_x^{-1}|\psi_x^B\rangle\langle\psi_x^B|. \tag{3.17}$$

Возвращаясь к эксперименту ЭПР, получаем из формулы (3.14), что если A измеряет наблюдаемую $\sigma(\boldsymbol{a})$ и посылает исход своего измерения ± 1 участнику B, то его апостериорным состоянием (т. е. состоянием соответствующего подансамбля) оказывается $|\mp\boldsymbol{a}\rangle\langle\mp\boldsymbol{a}|$. Если же связь между A и B отсутствует, то состояние B в результате измерения A не изменяется, оставаясь хаотическим. Таким образом, квантовая теория вполне согласуется с принципом локальности.

3.2.3 Сверхплотное кодирование

Хотя сцепленные состояния сами по себе не позволяют передавать информацию, оказывается, что наличие такого состояния позволяет увеличить максимальное количество классической информации, передаваемой от A к B, вдвое, если между системами имеется идеальный квантовый канал связи, позволяющий безошибочно передать любое квантовое состояние (например, посредством пересылки самого физического агента). Таким образом, сцепленное состояние выступает как “катализатор” при передаче классической информации через квантовый канал связи, и с этот точки зрения, также представляет собой особого рода информационный ресурс.

Для простоты рассмотрим системы A и B, каждая из которых представляет собой q-бит, между которыми имеется идеальный квантовый канал связи. Из обсуждения в предыдущем разделе вытекает, что максимальное количество классической информации, которое может быть передано от A к B, равно одному биту, и получается при кодировании бита в два ортогональных вектора, например,

$$0 \to |0\rangle, \quad 1 \to |1\rangle.$$

Протокол “сверхплотного кодирования”, позволяющий удвоить количество передаваемой информации, имеет в своей основе простой математический факт: все векторы *базиса Белла*

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

в системе из двух q-битов AB (мы используем канонический базис $|0\rangle, |1\rangle$ в пространстве одного q-бита и опускаем нормировочный множитель $1/\sqrt{2}$) могут быть получены из любого из этих векторов действием "локальных" унитарных операторов, т. е. операторов, действующих нетривиально только в пространстве q-бита A (или B), например

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

Таким образом, если система AB изначально находится в сцепленном состоянии $|e_+\rangle$, передатчик A может закодировать 2 бита классической информации в 4 состояния базиса Белла, производя только свои локальные операции, а затем переслать свой q-бит приемнику B по идеальному квантовому каналу. Тогда, производя измерение в базисе Белла, B получает 2 бита классической информации.

3.2.4 Телепортация квантовых состояний

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 представляет особого рода информационный ресурс, содержащий сведения о статистике всевозможных измерений над данной квантовой системой. 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 произвольного состояния (no cloning). Устройства, воспроизводящие в принципе произвольное классическое сообщение, хорошо известны. Однако существование прибора, который выполнял бы аналогичную задачу для квантовой информации, противоречит основному динамическому принципу квантовой механики. В самом деле, 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 размножал бы произвольное квантовое состояние.

Кратко остановимся на вопросе, каким вообще образом может быть передано квантовое состояние. Очевидно, что это может быть пересылка самого физического агента, т. е. системы, приготовленный в том или

ином состоянии. Одно из ярких достижений квантовой теории информации состоит в открытии принципиальной возможности способа, при котором сама система физически не пересылается, а передается лишь классическая информация. При этом существенным дополнительным ресурсом, который вновь играет роль "катализатора" передачи, является сцепленность между передатчиком и приемником. Заметим, что свести передачу произвольного квантового состояния только к передаче классической информации, не используя дополнительного квантового ресурса, невозможно: поскольку классическая информация копируема, это означало бы возможность копирования и квантовой информации. Дадим описание этого способа, получившего название *телепортация* квантового состояния.

Пусть имеются два участника A и B, играющие роль, соответственно, передатчика и приемника. В простейшей версии системы A и B являются двухуровневыми (q-битами).

i. Перед началом передачи система AB приготовляется в состоянии $|00\rangle + |11\rangle$. (Напомним, что мы всюду опускаем нормировочный множитель $1/\sqrt{2}$).
ii. Третий участник C пересылает A произвольное чистое состояние

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

которое должно быть передано участнику B. На этом шаге совокупность трех систем CAB описывается состоянием

$$(a|0\rangle + b|1\rangle) \otimes (|00\rangle + |11\rangle).$$

iii. Затем участник A производит некоторое обратимое преобразование состояния системы CA;
iv. Участник A производит определенное измерение (с 4 исходами, что составляет 2 бита классической информации) и посылает результат измерения B по классическому каналу связи. Преобразование и измерение будут описаны ниже.
v. 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"):

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

при которой состояние первого q-бита сохраняется, а состояние второго q-бита не изменяется, либо изменяется на противоположное, в зависимости от состояния первого q-бита. При этом базис переходит в базис,

следовательно, в 4-х мерном пространстве CA этому преобразованию соответствует унитарный оператор. Затем к q-биту A применяется операция Адамара H с унитарной матрицей

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

Тогда

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

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

Начальное состояние всей системы CAB есть

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

После действия CNOT на CA получаем

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

Потом H действует на C

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

Выделяя состояние системы CA, получаем

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

Теперь производится измерение наблюдаемой в системе CA, соответствующей базису $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Согласно формуле (3.17), апостериорное состояние системы B, в зависимости от полученного исхода измерения, описывается одним из векторов

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

Исход измерения посылается от A к B по классическому (идеальному) каналу связи. В зависимости от полученного сообщения B применяет к своему состоянию один из унитарных операторов Паули $\sigma_0 = I, \sigma_x, \sigma_z, \sigma_y$, в каждом случае переводящих вектор состояния B в $a|0\rangle + b|1\rangle$. Таким образом, B переходит в состояние, в котором первоначально находилась система C, при этом состояние C необратимо изменяется (иначе оказалось бы возможным копирование квантовой информации).

3.3 Комментарии

1. Подробнее о тензорном произведении гильбертовых пространств см. гл. II.4 [24], где можно найти решение задач этого раздела. Общую формулировку теоремы Наймарка см. в [23], где она доказана для общего разложения единицы (вероятностной операторно-значной меры, см. определение 10.7.2) в бесконечномерном гильбертовом пространстве.

Аналог разложения (3.8) в сепарабельном гильбертовом пространстве восходит к работе Шмидта 1907 г., посвященной интегральным уравнениям. Прием очищения смешанного состояния путем введения вспомогательной системы, широко применяемый в квантовой теории информации, использовался еще в работе Линдблада [126].

Термин "сцепленность" (Verschränktheit (нем.), entanglement (англ.)) был введен Шредингером, который первым обратил внимание на необычные свойства сцепленных состояний. В российской физической литературе используется термин "запутанность" или "перепутанность". Тщательный логический анализ эксперимента ЭПР проведен Беллом [57], который в частности установил первое из неравенств типа (3.15) и применил его к анализу парадокса ЭПР.

2. Протокол сверхплотного кодирования предложен Беннетом и Виснером [64].

Протокол квантовой телепортации предложен в знаменитой публикации Беннета, Брассара, Крепо, Джоза, Переса и Вуттерса [58]. Возможность телепортации состояния поляризации фотона была продемонстрирована экспериментально Цайлингером в 1997 г.

Часть II

4. Классическая энтропия и информация

4.1 Энтропия случайной величины и сжатие данных

Пусть X дискретная случайная величина, принимающая значения в конечном множестве $\mathcal{X}$ и пусть $P = \{p_x; x \in \mathcal{X}\}$ ее распределение вероятностей. *Энтропия* $H(X) = H(P)$ определяется соотношением

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

где

$$\eta(t) = \begin{cases} -t \log t, \, t > 0, \\ \\ 0, \qquad t = 0, \end{cases} . \tag{4.2}$$

Энтропия $H(X)$, как мы увидим далее, является мерой неопределенности, изменчивости или информационного содержания случайной величины X. В дальнейшем, не ограничивая общности, можно предполагать, что $p_x > 0$ для всех x.

Пусть d количество элементов в множестве $\mathcal{X}$. Имеют место неравенства

$$0 \leq H(P) \leq \log d, \tag{4.3}$$

причем минимальное значение принимается на вырожденных распределениях $\{1, 0, \ldots, 0\}, \ldots, \{0, \ldots, 0, 1\}$, а максимальное — на равномерном распределении $\{\frac{1}{d}, \ldots, \frac{1}{d}\}$. Первое неравенство вытекает из неотрицательности функции $\eta(t)$ на отрезке $[0, 1]$, а второе — из вогнутости функции $t \to \log t$:

$$H(P) = \sum_{x=1}^{d} p_x \log \frac{1}{p_x} \leq \log \sum_{x=1}^{d} p_x \frac{1}{p_x} = \log d.$$

Дадим операциональную интерпретацию энтропии $H(X)$ как меры информационного содержания случайной величины X. Рассмотрим "случайный источник", который порождает последовательность независимых одинаково распределенных случайных величин с распределением P. Последовательность $w = (x_1, \ldots, x_n)$ букв алфавита $\mathcal{X}$ называется *словом*

длины n. Общее количество таких слов равно $|\mathcal{X}|^n = 2^{n\log|\mathcal{X}|}$, 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 распределение P случайной величины X в общем случае не является равномерным, можно предложить намного лучший способ кодирования. Возможность сжатия данных тесно связана со следующим свойством *асимптотической равнораспределенности*.

Теорема 4.1.1 *Если $X_1,\dots,X_n,\dots$ независимые и одинаково распределенные случайные величины с распределением $P=\{p_x\}$, то*

$$-\frac{1}{n}\sum_{i=1}^{n}\log p_{X_i} \longrightarrow H(X) \qquad \textit{по вероятности.} \tag{4.4}$$

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

$$\mathsf{P}\left\{\Big|-\frac{1}{n}\sum_{i=1}^{n}\log p_{X_i}-H(X)\Big|<\delta\right\} > 1-\epsilon. \tag{4.5}$$

Это прямо вытекает из *закона больших чисел*, примененного к последовательности независимых одинаково распределенных случайных величин $-\log p_{X_i}; i=1,\dots,n$, математическое ожидание которых равно

$$\mathsf{E}(-\log p_{X_i}) = \sum_x p_x(-\log p_x) = H(X).$$

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

$$p_w = p_{x_1}\cdot\ldots\cdot p_{x_n} = 2^{-n\left(-\frac{1}{n}\sum_{i=1}^{n}\log p_{x_i}\right)} \tag{4.6}$$

мы теперь можем использовать соотношение (4.5), чтобы ввести понятие *типичного слова*:

Определение 4.1.1 *Слово w, имеющее вероятность p_w, называется δ-типичным, если*

$$2^{-n(H(X)+\delta)} < p_w < 2^{-n(H(X)-\delta)}, \tag{4.7}$$

другими словами, для него выполняется событие в формуле (4.5).

Множество всех δ-типичных слов длины n будет обозначаться $T^{n,\delta}$. Из (4.5), (4.7) вытекают следующие свойства множества δ-типичных слов (где δ,ϵ – фиксированные положительные числа):

Задача 4.1.1 *i. Существует не более $2^{n(H(X)+\delta)}$ типичных слов.*
ii. Для достаточно больших n множество не-типичных слов имеет вероятность $\mathsf{P}\left(\overline{T^{n,\delta}}\right)\leq\epsilon$.

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

Теперь можно осуществить эффективное *сжатие данных*, используя все двоичные последовательности длины $n(H(X)+\delta)$, чтобы закодировать все δ-типичные слова, и отбрасывая не-типичные (или кодируя их одним и тем же добавочным символом). Вероятность ошибки при таком кодировании будет меньше или равна ϵ.

Обратно, любой код, использующий двоичные последовательности длины $n(H(X)-\delta)$, имеет асимптотически неисчезающую вероятность ошибки, стремящуюся к единице при $n \to \infty$. В самом деле, пусть C совокупность слов, которые были использованы для 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

$$\mathsf{P}(C) = \mathsf{P}(C \cap T^{n,\delta/2}) + \mathsf{P}(C \cap \overline{T^{n,\delta/2}}) \tag{4.8}$$

$$\leq |C| 2^{-n(H(X)-\delta/2)} + \mathsf{P}(\overline{T^{n,\delta/2}}) \tag{4.9}$$

$$\leq 2^{-n\delta/2} + \epsilon, \tag{4.10}$$

что может быть сделано сколь угодно малым для достаточно больших n.

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 $H(X) = \max_X H(X) = \log|\mathcal{X}|$ и сжатие невозможно.

Задача 4.1.2 *Пусть $\mathcal{X}$ конечное подмножество $\mathbb{R}$. В условиях теоремы 4.1.1 имеет место следующее неравенство для вероятностей больших уклонений:*

$$\mathsf{P}\left\{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathsf{E}X) \geq \delta\right\} \leq \exp\left\{-n \sup_{s>0}\left[s\delta - \mu(s)\right]\right\}, \tag{4.11}$$

где

$$\mu(s) = \ln \mathsf{E}\mathrm{e}^{s(X-\mathsf{E}X)} = o(s); \qquad s \to 0.$$

Отсюда, в частности, вытекает закон больших чисел с экспоненциальной скоростью убывания "хвостов" распределения, поскольку $s\delta - \mu(s) > 0$ для малых $s > 0$. Указание: доказать и использовать следующую версию неравенства Маркова

$$\mathsf{P}\left\{X \geq x\right\} \leq \exp(-sx)\,\mathsf{E}\mathrm{e}^{sX}$$

для любых $x \in \mathbb{R}$ и $s > 0$.

4.2 Условная энтропия, относительная энтропия и информация Шеннона

Если X, Y случайные величины на одном вероятностном пространстве, имеющие совместное распределение $\{p_{x,y}\}$, то можно определить их *совместную энтропию* $H(X,Y)$ аналогично соотношению (4.1). *Условная энтропия* $H(Y|X)$ определяется как

$$\begin{aligned} H(Y|X) &= \sum_x p_x H(Y|X=x) \\ &= -\sum_x p_x \sum_y p(y|x) \log p(y|x) \\ &= -\sum_{x,y} p_{x,y} \log p_{x,y} + \sum_x p_x \log p_x \\ &= H(X,Y) - H(X). \end{aligned} \tag{4.12}$$

Существует полезное общее правило, что всякое линейное соотношение для (условных) энтропий продолжает оставаться верным, если в каждый член подставляется одно и тоже дополнительное условие, например (4.12) влечет

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

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 $x \in \mathcal{X}$. Входное распределение вероятностей $P = \{p_x\}$ трансформируется каналом в совместное распределение $p_{x,y} = p(y|x)p_x$ и порождает выходное распределение $P' = \{p'_y\}$, где $p'_y = \sum_x p(y|x)p_x$. *Шенноновское количество информации* (о случайной величине X, содержащееся в Y) определяется соотношением:

$$I(X;Y) = H(Y) - H(Y|X). \tag{4.14}$$

где энтропия $H(Y)$ может быть интерпретирована как информационное содержание выхода, а условная энтропия $H(Y|X)$ – как его бесполезная составляющая, обусловленная *шумом* в канале связи. Подставляя (4.12) в это выражение для шенноновской информации, мы видим, что оно симметрично по X и Y, и поэтому может быть также названо *взаимной информацией*

$$I(X;Y) = H(X) + H(Y) - H(X,Y). \tag{4.15}$$

Также $I(X;Y) = H(X) - H(X|Y)$, где теперь $H(X)$ может быть интерпретировано как информационное содержание источника, а $H(X|Y)$ – как *потеря* информации в канале связи с шумом. Явное выражение для

шенноновской информации через входное распределение и переходные вероятности канала имеет вид

$$I(X,Y) = \sum_{xy} p_x p(y|x) \log \left(\frac{p(y|x)}{\sum_{x'} p(y|x') p_{x'}} \right). \tag{4.16}$$

В классической статистике весьма полезной характеристикой является относительная энтропия двух распределений $P = \{p_x\}$, $Q = \{q_x\}$ (информационное количество Кульбака-Лейблера-Санова):

$$H(P;\, Q) = \begin{cases} \sum\limits_{x:p_x>0} p_x \log \frac{p_x}{q_x}, \text{ если } \{x : p_x > 0\} \subseteq \{x : q_x > 0\}; \\ +\infty \qquad\qquad\qquad\quad \text{в противном случае.} \end{cases}$$

Используя неравенство $\ln t \leq t - 1$, получаем

$$H(P;\, Q) \geq -\log \mathrm{e} \sum_{x:p_x>0} p_x \left(\frac{q_x}{p_x} - 1 \right) \geq 0, \tag{4.17}$$

причем равенство имеет место тогда и только тогда, когда $P = Q$.

Величина $H(P;\, Q)$ играет важную роль как "асимметричное расстояние" между распределениями вероятностей. Она обладает следующим свойством монотонности. Рассмотрим канал $r(y|x)$ и два распределения $P = \{p_x\}$, $Q = \{q_x\}$ на входе, которые трансформируются в выходные распределения $P' = \{p'_y\}$, $Q' = \{q'_y\}$,

$$p'_y = \sum_x r(y|x) p_x, \quad q'_y = \sum_x r(y|x) q_x.$$

Тогда

$$H(P';\, Q') \leq H(P;\, Q).$$

В самом деле,

$$\begin{aligned} H(P;\, Q) &= \sum_{xy} p_x r(y|x) \log \frac{r(y|x) p_x}{r(y|x) q_x} \\ &= \sum_{xy} p_{x,y} \left(\log \frac{p'_y}{q'_y} + \log \frac{p(x|y)}{q(x|y)} \right) \\ &= H(P';\, Q') + \sum_y p'_y H(p(x|y); q(x|y)) \\ &\geq H(P';\, Q'). \end{aligned}$$

Задача 4.2.1 *Докажите, что* $I(X;Y) = H(p_{x,y};\, p_x \cdot p_y)$. *Поэтому, в силу (4.17),* $I(X;Y) \geq 0$, *причем* $I(X;Y) = 0$ *тогда и только тогда, когда* X *и* Y *независимые случайные величины:* $p_{x,y} = p_x \cdot p_y$.

Отсюда непосредственно вытекает монотонность условной энтропии

$$0 \leq H(Y|X) \leq H(Y),$$

а также субаддитивность энтропии:

$$H(X,Y) \leq H(X) + H(Y). \tag{4.18}$$

Отметим также выражение для информации в терминах относительной энтропии

$$I(X,Y) = \sum_x p_x H(P'_x; P'), \tag{4.19}$$

где мы обозначили $P'_x = \{p(y|x)\}$ для каждого фиксированного $x \in \mathcal{X}$, а $P' = \sum_x p_x P'_x$ есть безусловное распределение Y.

4.3 Шенноновская пропускная способность канала с шумом

Рассмотрим канал $X \to Y$, определяемый переходной вероятностью $p(y|x)$. Важнейшей характеристикой канала является *шенноновская пропускная способность*

$$C_{Shan} = \max_X I(X;Y), \tag{4.20}$$

где максимум берется по всевозможным распределениям $P = \{p_x\}$ на входе X. Далее будет показано, что она совпадает с операционально определенной информационной пропускной способностью канала.

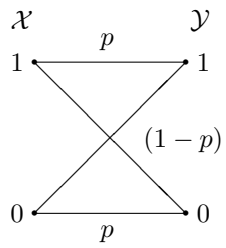


Рис. 4.1. *Двоичный симметричный канал.*

В качестве примера вычисления C_{Shan} рассмотрим *двоичный симметричный канал* (см. рис. 4.3). В этом случае алфавиты $\mathcal{X}$ и $\mathcal{Y}$ состоят из двух букв $0, 1$ которые передаются без ошибки с вероятностью $1-p$ и изменяются на другую с вероятностью p. Вводя *двоичную энтропию*

$$h_2(p) = -p \log p - (1-p) \log(1-p), \tag{4.21}$$

взаимную информацию можно записать как $I(X;Y) = H(Y) - H(Y|X) = H(Y) - h_2(p)$, что ограничено сверху величиной

$$C_{Shan} = \log 2 - h_2(p) = 1 - h_2(p), \tag{4.22}$$

достигающейся для равномерного распределения: $p_0 = p_1 = \frac{1}{2}$.

В общем случае $I(X;Y)$ является вогнутой функцией входного распределения вероятностей $P = \{p_x\}$. Поэтому для нахождения оптимального входного распределения P, максимизирующего $I(X;Y)$, могут быть применены условия Куна-Таккера из выпуклого анализа.

Задача 4.3.1 *Докажите, что энтропия $H(P)$ является вогнутой функцией распределения P. Указание: используйте вогнутость функции $t \to \eta(t)$. Поэтому выходная энтропия $H(Y) = H(P')$ также вогнута, тогда как условная энтропия $H(Y|X)$ аффинна. Представление (4.14) таким образом влечет вогнутость функции $I(X;Y)$.*

Задача 4.3.2 *(Условия Куна-Таккера) Пусть $F(P)$ – вогнутая функция распределения вероятностей P, непрерывно дифференцируемая внутри симплекса распределений на $\mathcal{X}$, так что частные производные имеют пределы, возможно бесконечные, на границе симплекса. Для того чтобы распределение P^0 было точкой абсолютного максимума для $F(P)$, необходимо и достаточно, чтобы нашлось λ, такое что*

$$\frac{\partial F(P^0)}{\partial p_x} \begin{cases} = \lambda, & p_x^0 > 0; \\ \leq \lambda & p_x^0 = 0. \end{cases}$$

В случае шенноновской информации (4.16) имеем

$$\frac{\partial F(P)}{\partial p_x} = H(P_x; P') - \log \mathrm{e},$$

что в сочетании с (4.19) приводит к следующему условию "максимального расстояния" для оптимального входного распределения:

$$H(P_x^0; (P^0)') \begin{cases} = \mu, & p_x^0 > 0; \\ \leq \mu, & p_x^0 = 0, \end{cases} \tag{4.23}$$

причем с необходимостью $\mu = C_{Shan}$.

Предположим, что имеются два канала $\{p_j(y_j|x_j)\}; j = 1, 2$, и рассмотрим составной канал, описываемый переходной вероятностью $\{p_1(y_1|x_1)p_2(y_2|x_2)\}$.

Предложение 4.3.1 *Шенноновская пропускная способность составного канала равна*

$$(C_{Shan})_{12} = (C_{Shan})_1 + (C_{Shan})_2 . \tag{4.24}$$

Таким образом, шенноновская пропускная способность аддитивна для независимых каналов.

Доказательство. Пусть P^j – оптимальное входное распределение для $j-$го канала, тогда оно удовлетворяет условию (4.23) с константами $\mu_j = (C_{Shan})_j$. Используя тот факт, что

$$H\left(P^1 \times P^2; Q^1 \times Q^2\right) = H\left(P^1; Q^1\right) + H\left(P^2; Q^2\right),$$

получаем, что распределение $P = P^1 \times P^2$ удовлетворяет условию (4.23) для составного канала с константой $\mu = (C_{Shan})_1 + (C_{Shan})_2$. □

4.4 Теорема кодирования для канала с шумом

Для данного канала $p(y|x)$ можно рассмотреть составной канал *без памяти*

$$p(y^n|x^n) = p(y_1|x_1) \cdot \ldots \cdot p(y_n|x_n), \tag{4.25}$$

который побуквенно передает слова длины n, используя независимо n раз исходный канал:

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

Обозначим X^n, Y^n случайные величины на входе и выходе составного канала.

Аналогично (4.20) рассмотрим величину

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

для составного канала. Используя свойство (4.24), получаем, что последовательность $\{C_n\}$ *аддитивна* для канала без памяти, откуда

$$C_n = nC_{Shan}. \tag{4.26}$$

Чтобы уменьшить влияние шума, используется кодирование сообщений на входе и, соответственно, декодирование на выходе составного канала. Процесс передачи информации тогда изображается следующей диаграммой:

$$i \quad \longrightarrow x^n \longrightarrow y^n \longrightarrow j, \tag{4.27}$$

где i (соответственно j) обозначает переданное (принятое) сообщения. Можно просто считать $i, j \in \{1, \ldots, N\}$ номерами соответствующих сообщений, так как интерес представляет количество N передаваемых сообщений.

Цель состоит в том, чтобы выбрать кодирование и декодирование, которые максимизировали бы *скорость передачи* $R = \frac{\log N}{n}$, равную числу бит на один передаваемый символ, при условии малости вероятности ошибки. Дадим точные определения.

Определение 4.4.1 *Код (W, V) размера N для составного канала $p(y^n|x^n)$ состоит из кодирования, задаваемого совокупностью кодовых слов $W = \{w^{(1)}, \ldots, w^{(N)}\}$ длины n и декодирования, задаваемого совокупностью $V = \{V^{(0)}, V^{(1)}, \ldots, V^{(N)}\}$ непересекающихся подмножеств $\mathcal{Y}^n$. Подмножества $V^{(1)}, \ldots, V^{(N)}$ интерпретируются как области принятия решения: если сообщение на выходе $y^n \in V^{(j)}; j = 1, \ldots, N$, то принимается решение, что было послано слово $w^{(j)}$; если же $y^n \in V^{(0)} = \overline{\cup_{j=1}^N V^{(j)}}$, то никакого определенного решения не принимается.*

Таким образом, *максимальная вероятность ошибки* такого кода есть

$$P_e(W,V) = \max_{1\le j\le N}\left(1 - p(V^{(j)}|w^{(j)})\right), \tag{4.28}$$

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

$$\overline{P}_e(W,V) = \frac{1}{N}\sum_{j=1}^{N}\left(1 - p(V^{(j)}|w^{(j)})\right) \le P_e(W,V), \tag{4.29}$$

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

Лемма 4.4.1 *Для любого кода (W,V) размера $2N$ найдется подкод $(\tilde{W},\tilde{V})$ размера N, который имеет максимальную вероятность ошибки $P_e(\tilde{W},\tilde{V}) \le 2\overline{P}_e(W,V)$.*

Доказательство. Обозначим $\epsilon = \overline{P}_e(W,V)$ и предположим, что N-подкода с требуемым свойством не существует, тогда найдутся по крайней мере $N+1$ кодовых слов с вероятностью ошибки $> 2\epsilon$. Тогда средняя вероятность ошибки $2N$-кода ограничена снизу величиной

$$\overline{P}_e(W,V) > \frac{1}{2N}2\epsilon(N+1) > \epsilon,$$

что противоречит предположению. □

Обозначим $p_e(n,N)$ (соответственно $\bar{p}_e(n,N)$) максимальную (соотв. среднюю) вероятность ошибки, минимизированную по всевозможным кодам длины n и размера N, тогда

$$\frac{1}{2}p_e(n,N) \le \bar{p}_e(n,2N) \le p_e(n,2N) \tag{4.30}$$

Определение 4.4.2 *Число $R \ge 0$ называется достижимой скоростью передачи, если*

$$\lim_{n\to\infty} p_e(n,2^{nR}) = 0.$$

Точная верхняя грань всех достижимых скоростей передачи называется информационной пропускной способностью C канала $p(y|x)$.

Теорема 4.4.2 (Теорема кодирования для канала с шумом) *Для канала без памяти*

$$C = C_{Shan}.$$

Таким образом, операционально (и асимптотически) определенная информационная пропускная способность оказывается равной удобно вычислимой "однобуквенной" характеристике канала, определенной соотношением (4.20). Неравенство $C \leq C_{Shan}$ будет следовать из утверждения

$$\lim_{n\to\infty} p_e(n, 2^{nR}) = 0, \text{ если } R < C_{Shan},$$

называемого *прямой теоремой кодирования*, тогда как *слабое обращение* состоит в том, что

$$\liminf_{n\to\infty} p_e(n, 2^{nR}) > 0, \text{ если } R > C_{Shan}, \tag{4.31}$$

откуда следует $C \geq C_{Shan}$. На самом деле, имеет место более сильное утверждение

$$\lim_{n\to\infty} p_e(n, 2^{nR}) = 1, \text{ если } R > C_{Shan}.$$

Доказательство. Слабое обращение. Благодаря неравенству (4.30), достаточно доказать аналог утверждения (4.31) для вероятности ошибки $\overline{p}_e(n, N)$.

Лемма 4.4.3 *(Неравенство Фано) Пусть X, Y случайные величины и $\hat{X} = \hat{X}(Y)$ оценка случайной величины X по наблюдениям Y с вероятностью ошибки $p_e = \mathsf{P}(\hat{X}(Y) \neq X)$, тогда*

$$H(X|Y) \leq h_2(p_e) + p_e \log(|\mathcal{X}| - 1) \leq 1 + p_e \log |\mathcal{X}|. \tag{4.32}$$

Доказательство. Пусть E индикатор ошибки оценивания,

$$E = \begin{cases} 0, & \text{если } \hat{X}(Y) = X; \\ 1, & \text{в противном случае.} \end{cases} \tag{4.33}$$

Поскольку E является функцией $(X, \hat{X})$ и поэтому имеет определенное значение при фиксированных значениях, имеем $H(X|Y) = H(E, X|Y)$. Однако из (4.13) следует

$$H(E, X|Y) = H(E|Y) + H(X|E, Y). \tag{4.34}$$

Здесь $H(E|Y) \leq H(E) = h_2(p_e)$ и

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

где был использован тот факт, что $H(X|E = 0, Y)$ также равно нулю, поскольку $E = 0$ означает, что мы знаем X, если известно Y. Условие $E = 1$ то же самое, что $X \neq \hat{X}(Y)$, и оставляет возможными не более $|\mathcal{X}| - 1$ значений X, так что энтропия ограничена сверху величиной $\log(|\mathcal{X}| - 1)$. □

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

$$\begin{aligned} C_n &\geq I(X^n; Y^n) = H(X^n) - H(X^n|Y^n) \\ &\geq \log N \left(1 - \underbrace{\mathsf{P}\{\hat{X}^n \neq X^n\}}_{=\overline{P}_e(W,V)} \right) - 1 \end{aligned} \tag{4.35}$$

согласно неравенству Фано. Подставляя $N = 2^{nR}$, получаем

$$\overline{P}_e(W, V) \geq 1 - \frac{C_n}{nR} - \frac{1}{nR}. \tag{4.36}$$

Беря минимум по всевозможным кодам и используя аддитивность (4.26), получаем в пределе $n \to \infty$

$$\liminf_{n\to\infty} \overline{p}_e(n, 2^{nR}) \geq 1 - \frac{C_{Shan}}{R} > 0$$

для $R > C_{Shan}$.

Доказательство прямого утверждения. Основная идея, принадлежащая Шеннону, состоит в использовании *случайного кодирования*. Рассмотрим N слов $w^{(1)}, \ldots, w^{(N)}$, выбираемых случайным образом независимо с распределением вероятностей

$$\mathsf{P}\{w^{(i)} = (x_1, \ldots, x_n)\} = p_{x_1} \cdot \ldots \cdot p_{x_n}, \tag{4.37}$$

где однобуквенное распределение $\{p_x\}$ выбрано так, что оно максимизирует $I(X; Y)$. Заметим, что имеется примерно $2^{nH(Y)}$ типичных слов 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

$$N \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)} = 2^{nC_{Shan}}. \tag{4.38}$$

Чтобы придать этому рассуждению точный смысл, назовем выходное слово y^n *условно типичным* для входного слова $w = x^n$, если

$$2^{-n(H(Y|X)+\delta)} < p(y^n|w) < 2^{-n(H(Y|X)-\delta)}. \tag{4.39}$$

Обозначим $T_w^{n,\delta} \subset \mathcal{Y}^n$ подмножество всех условно типичных слов для данного входного слова w.

Задача 4.4.1 *Предполагая, что w распределено случайно с распределением (4.37), покажите, используя закон больших чисел, что* $\mathsf{P}\{Y^n \in \overline{T_w^{n,\delta}}\} \le \varepsilon$ *для любого* $\varepsilon > 0$ *и достаточно больших* n.

Для данного кодирования $W = \{w^{(1)}, \dots, w^{(N)}\}$ построим специальное субоптимальное декодирование. Множества $T_{w^{(j)}}^{n,\delta}$ могут пересекаться для разных j, и чтобы получить непересекающиеся области принятия решений, мы положим

$$V^{(j)} = T_{w^{(j)}}^{n,\delta} \cap \overline{\left(\cup_{k:k\neq j} T_{w^{(k)}}^{n,\delta}\right)}. \tag{4.40}$$

Тогда если, например, было передано слово $w^{(1)}$, ошибка возникает тогда и только тогда, когда

$$y^n \in \overline{V^{(1)}} = \overline{T_{w^{(1)}}^{n,\delta}} \cup \left(\cup_{k=2}^{N} T_{w^{(k)}}^{n,\delta}\right). \tag{4.41}$$

Теперь предположим, что кодовые слова $w^{(1)}, \dots, w^{(N)}$ выбираются случайным образом, как описано выше, т. е. независимо с распределением (4.37), и оценим математическое ожидание средней вероятности ошибки $\overline{P}_e(W, V)$. Это даст нам требуемую оценку, поскольку очевидно, что

$$\bar{p}_e(n, N) \le \mathsf{E}\overline{P}_e(W, V).$$

Обозначим $\mathsf{P}_w\{B\}$ вероятность $\mathsf{P}\{Y^n \in B | X^n = w\}$, где B некоторое, возможно случайное, подмножество $\mathcal{Y}^n$. В силу полной симметрии между кодовыми словами

$$\mathsf{E}\overline{P}_e(W, V) = \mathsf{E}\,\mathsf{P}_{w^{(1)}}\{\overline{V^{(1)}}\} \tag{4.42}$$
$$= \mathsf{E}\,\mathsf{P}_{w^{(1)}}\{\overline{V^{(1)}} \cap T^{n,\delta}(Y)\} + \mathsf{E}\,\mathsf{P}_{w^{(1)}}\{\overline{V^{(1)}} \cap \overline{T^{n,\delta}(Y)}\},$$

где $T^{n,\delta}(Y)$ есть множество выходных δ-типичных последовательностей y^n, определяемое аналогично определению 4.1.1 с заменой $H(X)$ на $H(Y)$. Тогда, учитывая (4.41),

$$\mathsf{E}\overline{P}_e(W, V) \le \mathsf{P}\{Y^n \in \overline{T_{w^{(1)}}^{n,\delta}}\} + \sum_{k=2}^{N} \mathsf{E}\,\mathsf{P}_{w^{(1)}}\{T_{w^{(k)}}^{n,\delta} \cap T^{n,\delta}(Y)\} + \mathsf{P}\{\overline{T^{n,\delta}(Y)}\}. \tag{4.43}$$

Первый и третий члены $\le \varepsilon$ для достаточно больших n в силу соответствующих свойств (условно) нетипичных слов. Каждое слагаемое в средней сумме оценивается как

$$\sum_{w^{(1)}} \sum_{w^{(k)}} \sum_{y^n \in T_{w^{(k)}}^{n,\delta} \cap T^{n,\delta}(Y)} p(y^n|w^{(1)}) p_{w^{(1)}} p_{w^{(k)}}$$

$$\le 2^{n(H(Y|X)+\delta)} \sum_{w^{(1)}} \sum_{w^{(k)}} \sum_{y^n \in T^{n,\delta}(Y)} p(y^n|w^{(1)}) p_{w^{(1)}} p(y^n|w^{(k)}) p_{w^{(k)}}$$

$$= 2^{n(H(Y|X)+\delta)} \sum_{y^n \in T^{n,\delta}(Y)} (p_{y^n})^2 \leq 2^{n(H(Y|X)-H(Y)+2\delta)}.$$

Здесь первое неравенство получается введением множителя $2^{n(H(Y|X)+\delta)} p(y^n|w^{(k)}) > 1$ на множестве $T^{n,\delta}_{w^{(k)}}$, а последнее – из второго неравенства в определении типичного слова y^n. Окончательно,

$$\mathsf{E}\overline{P}_e(W,V) \leq 2\varepsilon + (N-1)2^{n(H(Y|X)-H(Y)+2\delta)} \leq 2\varepsilon + N2^{-n(C_{Shan}-2\delta)},$$

что можно сделать меньше 3ε если $N = 2^{nR}$ с $R < C_{Shan}$ и δ достаточно мало. □

Следует отметить, что метод случайного кодирования, позволивший доказать существования асимптотически оптимальных кодов, а также раскрывающий природу таких кодов, мало пригоден для практического применения. Его реализация уже при достаточно умеренных значениях n требует огромного (экспоненциально растущего) объема вычислений. Нахождению практически приемлемых методов кодирования/декодирования посвящен специальный раздел теории информации – теория кодирования, широко использующая методы современной алгебры.

4.5 Канал с перехватом

Другой важнейшей проблемой современной теории информации является изучение систем с многими пользователями, ярким примером которой может служить интернет. Специфическим примером системы с недружественными участниками является канал с перехватом, на котором мы здесь кратко остановимся, поскольку это понадобится в гл. 9, посвященной передаче квантовой информации.

Пусть $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ – три конечных алфавита, причем алфавит $\mathcal{X}$ ассоциируется с передатчиком, $\mathcal{Y}$ – с приемником, а $\mathcal{Z}$ – с перехватчиком данных. *Каналом с перехватом* называется пара каналов $p(y|x); q(z|x)$ от передатчика к приемнику и к перехватчику, соответственно. Передатчик посылает слова $w = (x_1, \ldots, x_n)$ длины n по составным каналам $p(y^n|x^n), q(y^n|x^n)$, и цель состоит в асимптотически безошибочной передаче максимального количества сообщений приемнику при условии, что перехватчик получит асимптотически исчезающее количество информации.

Определение 4.5.1 *Код $(\mathcal{M}, r, V)$ размера N для составного канала с перехватом состоит из множества сообщений $\mathcal{M} = \{1, \ldots, N\}$, переходной вероятности $r(x^n|m)$ из $\mathcal{M}$ в $\mathcal{X}^n$, задающей рандомизованное кодирование сообщений $m \in \mathcal{M}$ в слова x^n, и обычного декодирования V для приемника, задаваемого разбиением множества $\mathcal{Y}^n$ на N непересекающихся подмножеств $V^{(0)}, V^{(1)}, \ldots, V^{(N)} \subset \mathcal{Y}^n$.*

Максимальная вероятность ошибки такого кода равна

$$P_e(\mathcal{M}, r, V) = \max_{1 \le m \le N} \left(1 - p(V^{(m)}|m)\right), \tag{4.44}$$

где $p(V^{(m)}|m) = \sum_{x^n} \mathsf{P}(Y^n \in V^{(m)}|x^n) r(x^n|m)$ – вероятность правильного решения для приемника. Мы будем называть R *достижимой скоростью* для канала с перехватом, если найдется последовательность кодов $\left(\mathcal{M}^{(n)}, r^{(n)}, V^{(n)}\right)$ размера $N = 2^{nR}$, таких что

$$\lim_{n\to\infty} P_e\left(\mathcal{M}^{(n)}, r^{(n)}, V^{(n)}\right) = 0$$

и

$$\lim_{n\to\infty} I(M^n; Z^n) = 0,$$

где M^n – случайная величина, равномерно распределенная на множестве сообщений $\mathcal{M}^{(n)}$. Последнее условие выражает требование, что количество информации перехватчика о посланном сообщении стремится к нулю. Точная верхняя грань множества достижимых скоростей называется *секретной классической пропускной способностью* C_p канала с перехватом.

Теорема кодирования для канала с перехватом дает следующее выражение:

$$C_p = \max\left[I(M;Y) - I(M;Z)\right], \tag{4.45}$$

где максимум берется по всевозможным случайным величинам M, Y, Z, таким что последовательность $M, X, (Y, Z)$ образует цепь Маркова, причем пары X, Y и X, Z связаны, соответственно, каналами $p(y|x)$ и $q(z|x)$.

Задача 4.5.1 *Докажите слабое обращение теоремы кодирования (неравенство $\le$ в (4.45)), используя неравенство Фано.*

Доказательство прямого утверждения основано на следующей идее. Фиксируем $\delta, \varepsilon > 0$ и некоторое распределение на алфавите $\mathcal{X}$. Рассмотрим новое случайное кодирование $W^{(n)}$ с N независимыми словами, равномерно распределенными на множестве $T^{n,\delta}$ $\delta-$типичных слов длины n, а также субоптимальное декодирование $V^{(n)}$, построенное как в (4.40). Небольшая модификация доказательства теоремы 4.4.2 позволяет доказать, что при $N = 2^{n[I(X;Y)-\delta]}$ с высокой вероятностью выполняется

$$\bar{P}_e(W^{(n)}, V^{(n)}) \le \varepsilon. \tag{4.46}$$

Чтобы сделать код секретным, передатчик должен пожертвовать $n\left[I\left(X;Z\right) + \delta/2\right]$ битами информации, дополнительно рандомизуя входные сообщения. Положим

$$N_{\mathcal{Z}} = 2^{n[I(X;Z)+\delta/2]},\ N_{\mathcal{Y}} = 2^{n[I(X;Y)-I(X;Z)-3\delta/2]},$$

так что $N_{\mathcal{Z}}N_{\mathcal{Y}} = N$; представим набор кодовых слов $W^{(n)}$ в виде прямоугольной таблицы с $N_{\mathcal{Y}}$ строками и $N_{\mathcal{Z}}$ столбцами. Тогда

$$W^{(n)} = \left\{w^{mj}; m = 1, \ldots, N_{\mathcal{Y}}; j = 1, \ldots, N_{\mathcal{Z}}\right\}.$$

Пусть теперь для каждого значения m передатчик выбирает значение j случайным образом с равными вероятностями. Такая рандомизация приводит к тому, что почти вся переданная информация оказывается скрытой для перехватчика: для каждого значения m набор кодовых слов

$$\left\{w^{mj}; j = 1, \ldots, N_{\mathcal{Z}}\right\}$$

с высокой вероятностью содержит почти максимально возможную информацию $I(X;Z)$, при условии, что перехватчик применяет оптимальное декодирование. Поэтому взаимная информация между наборами кодовых слов с различными значениями m должна быть близка к нулю, так что рандомизация внутри каждого набора уничтожает почти всю информацию, передаваемую перехватчику. Эти рассуждения вкупе с (4.46) показывают, что скорость $I(X;Y) - I(X;Z) - \delta$ является достижимой. Они остаются справедливыми и для любой последовательности M, X, YZ, удовлетворяющей условиям теоремы кодирования, поэтому $I(M;Y) - I(M;Z) - \delta$ также является достижимой скоростью.

4.6 Гауссовский канал

Рассмотрим непрерывный аналог канала без памяти (4.25), у которого в качестве входного и выходного алфавитов выступает вещественная прямая $\mathbb{R}$. Канал задается переходной плотностью вероятности

$$p(y_i|x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - x_i)^2}{2\sigma^2}\right]; \quad i = 1, \ldots, n.$$

Эквивалентно, можно считать, что выходные случайные величины Y_i получаются сложением входного сигнала x_i с независимыми одинаково распределенными переменными шума Z_i с $\mathsf{E}Z_i = 0$ и $\mathsf{E}Z_i^2 = \sigma^2$

$$Y_i = x_i + Z_i; \quad i = 1, \ldots, n.$$

Легко видеть, что естественно определенная информационная пропускная способность такого канала бесконечна, если не ввести каких-либо ограничений на сигнал. Обычно рассматривается квадратичное ограничение

$$\sum_{i=1}^{n} x_i^2 \leq ns^2, \tag{4.47}$$

мотивируемое конечностью мощности входного сигнала. При этом ограничении пропускная способность оказывается равной

$$C = \frac{1}{2}\log\left(1 + \frac{s^2}{\sigma^2}\right). \tag{4.48}$$

Эта формула также принадлежит Шеннону и также имеет простое эвристическое истолкование. Неравенства

$$\frac{1}{n}\sum_{i=1}^{n} Z_i^2 \le \sigma^2 + \varepsilon; \quad \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \le s^2 + \sigma^2 + \varepsilon$$

имеют место с высокой вероятностью для произвольного $\varepsilon > 0$ и достаточно больших n.

Задача 4.6.1 *Докажите это утверждение, используя неравенство Чебышева и (4.47).*

Первое из этих неравенств означает, что для любого входного слова $w = (x_1, \ldots, x_n)$ выходной вектор $Y^n = (Y_1, \ldots, Y_n)$ лежит в $n-$мерном шаре радиуса $\sqrt{n(\sigma^2 + \varepsilon)}$ с центром w, тогда как второе – что выходные векторы для всевозможных входов, удовлетворяющих ограничению (4.47), лежат в $n-$мерном шаре радиуса $\sqrt{n(s^2 + \sigma^2 + \varepsilon)}$. Отношение этих объемов дает приблизительную оценку числа непересекающихся шаров первого типа в большом шаре:

$$N = \left[\frac{\sqrt{n(s^2 + \sigma^2 + \varepsilon)}}{\sqrt{n(\sigma^2 + \varepsilon)}}\right]^n = 2^{nR},$$

где $R = \frac{1}{2}\log\left(1 + \frac{s^2}{\sigma^2 + \varepsilon}\right)$ может быть сделано сколь угодно близким к (4.48).

4.7 Комментарии

1. Имеется много прекрасных книг по теории информации. Здесь мы во многом следуем книге Ковера и Томаса [72], которая дает ясное и неформальное введение в предмет.

Метод типов (типичных последовательностей) был систематически разработан Чисаром и Кернером [45].

2. Идея использования показательной функции в неравенстве Маркова с последующей оценкой больших уклонений восходит к Бернштейну и была развита Крамером; ее важность для теории информации подчеркивалась Черновым, см. например [72], лемма 12.9.1, а также [5].

3. Доказательство утверждения задачи 4.3.2 и его приложения в теории информации обсуждаются в книге Галлагера [5], теоремы 4.4.1, 4.5.1.

4. Первоначальные идеи теоремы кодирования были изложены в пионерской работе Шеннона [46]. Наше доказательство отличается от данного в книге [72] тем, что использует условную, а не совместную типичность. Именно такой подход допускает некоммутативное обобщение, см. раздел 5.6.

5. Теория многотерминальных систем излагается в книгах Чисара и Кернера [45], Ковера и Томаса [72]. Доказательство теоремы кодирования для канала с перехватом см. в [45].

6. Строгое доказательство формулы (4.48) см., например, в [72]. Квантовый аналог канала с аддитивным гауссовским шумом рассматривается в разделе 11.5.1.

5. Квантовая теорема кодирования

5.1 Пропускная способность классически-квантового канала связи

Как было выяснено в разделе 3.2.1, простейшая модель квантового канала предполагает, что есть классический параметр x, пробегающий (конечный) входной алфавит $\mathcal{X}$ и отображение $x \to S_x$ в квантовые состояния на выходе канала. Мы будем называть такую модель классически-квантовым (c-q) каналом. Если на выходе такого канала измеряется наблюдаемая $M = \{M_y\}$, то условная вероятность получить исход y, при условии, что был послан сигнал x, дается формулой

$$P(y|x) = \operatorname{Tr} S_x M_y. \tag{5.1}$$

Таким образом, для фиксированного измерения мы получаем обычный классический канал связи. Это приводит к вопросу о максимальном количестве классической информации, которое может быть почти без помех передано по данному квантовому каналу и о соответствующей предельной характеристике – классической пропускной способности. Этот вопрос будет детально рассмотрен в настоящей главе.

Рассмотрим соответствующий составной c-q канал, который отображает слово $w = (x_1, \ldots, x_n)$ в состояние-произведение $S_w = S_{x_1} \otimes \ldots \otimes S_{x_n}$ в пространстве $\mathcal{H}^{\otimes n}$. Процесс передачи классической информации описывается диаграммой

$$i \quad \longrightarrow w \longrightarrow S_w \longrightarrow j \tag{5.2}$$

Предположение о том, что слово w отображается в тензорное произведение состояний S_{x_j}, соответствует определению канала без памяти в классическом случае. На выходе канала находится приемник, который производит измерение некоторой наблюдаемой $M = \{M_j^{(n)}\}$ в пространстве $\mathcal{H}^{\otimes n}$ (получив исход измерения j, считаем, что было послано сообщение j). В итоге приемник выдает ответ о принятом решении; таким образом, разложение единицы в пространстве $\mathcal{H}^{\otimes n}$ описывает статистику всей решающей процедуры, которая включает в себя физическое измерение и последующую классическую обработку его результатов. Выбор

наблюдаемой M формально аналогичен выбору решающей процедуры в классическом случае, но как мы увидим, играет здесь гораздо более важную роль.

Определение 5.1.1 *Код (W, M) размера N для составного c-q канала $x^n \to S_{x^n}$ состоит из (классического) кодирования W, задаваемого набором N кодовых слов $w^{(i)}; i = 1, \ldots, N$, длины n и (квантового) декодирования, задаваемого наблюдаемой $M = \{M_j; j = 0, 1, \ldots, N\}$ в $\mathcal{H}^{\otimes n}$.*

Исход 0 означает уклонение от принятия решения. Вероятность декодировать входное сообщение i как сообщение j равна

$$p_{WM}(j|i) = \operatorname{Tr} S_{w^{(i)}} M_j, \quad j = 0, 1, \ldots, N. \tag{5.3}$$

Вероятность правильного решения тогда равна $p_{WM}(i|i) = \operatorname{Tr} S_{w^{(i)}} M_i$, а максимальная (по всем входным сообщениям) вероятность ошибки кода (W, M) есть

$$P_e(W, M) = \max_{1 \le j \le N} \left(1 - p_{WM}(j|j)\right). \tag{5.4}$$

Средняя ошибка кода равна

$$\overline{P}_e(W, M) = \frac{1}{N} \sum_{j=1}^{N} [1 - p_{WM}(j|j)]. \tag{5.5}$$

Будем обозначать далее

$$p_e(n, N) = \min_{W,M} P_e(W, M), \quad \bar{p}_e(n, N) = \min_{W,M} \overline{P}_e(W, M), \tag{5.6}$$

соответственно, максимальную и среднюю ошибку, минимизированные по всем кодам (W, M) размера N, использующим слова длины n.

Определение 5.1.2 *Классической пропускной способностью C c-q канала $x \to S_x$ называется точная верхняя грань скоростей передачи R, которые достижимы в том смысле, что $\lim_{n\to\infty} p_e(n, 2^{nR}) = 0$ или, что равносильно (в силу неравенств (4.30)), $\lim_{n\to\infty} \bar{p}_e(n, 2^{nR}) = 0$.*

5.2 Формулировка теоремы кодирования

Понятие энтропии квантового состояния является естественным обобщением энтропии распределения вероятностей. Пусть S – оператор плотности в d-мерном гильбертовом пространстве $\mathcal{H}$,

$$S = \sum_{j=1}^{d} s_j |e_j\rangle\langle e_j|$$

– его спектральное представление. Собственные числа s_j образуют распределение вероятностей. *Энтропия фон Неймана* оператора плотности S определяется соотношением

$$H(S) = \sum_{j=1}^{d} \eta(s_j) = \operatorname{Tr} \eta(S). \tag{5.7}$$

Далее мы будем использовать более наглядное выражение $H(S) = -\operatorname{Tr} S \log S$, хотя логарифм не определен для вырожденных операторов плотности. Как и в случае распределений, это не приведет к недоразумениям. Из (4.3) следует, что

$$0 \leq H(S) \leq \log d,$$

причем минимум достигается на чистых состояниях, а максимум – на хаотическом состоянии $\bar{S} = \frac{1}{d} I$. Как и в классическом случае, энтропия является мерой неопределенности, а также информационного содержания состояния (последнее утверждение будет подробно обосновано в разделе 5.5).

Легко проверяются следующие свойства квантовой энтропии:

Задача 5.2.1 *i. Унитарная инвариантность: $H(VSV^*) = H(S)$, где V – унитарный оператор. Более того, это равенство справедливо для любого оператора V, изометричного на носителе S;*
ii. Аддитивность: $H(S_1 \otimes S_2) = H(S_1) + H(S_2)$.

В ранних работах по квантовым оптическим каналам связи для оценки пропускной способности канала $x \to S_x$ использовалась величина

$$\chi\left(\{\pi_x\}; \{S_x\}\right) = H\Big(\sum_x \pi_x S_x\Big) - \sum_x \pi_x H(S_x), \tag{5.8}$$

которая может рассматриваться как формальный некоммутативный аналог шенноновской информации, причем первое слагаемое играет роль выходной энтропии, а второе – условной энтропии квантового выхода относительно классического входного сигнала. Замечательно, что величина (5.8) действительно оказывается непосредственно связанной с классической пропускной способностью c-q канала, определенной выше.

В ситуации, когда состояния $\{S_x\}$ фиксированы, мы будем обозначать величину (5.8) просто как $\chi(\pi)$. Основной целью следующих разделов будет доказательство теоремы:

Теорема 5.2.1 (Квантовая теорема кодирования) *Определенная выше классическая пропускная способность C c-q канала $x \to S_x$ равна величине*

$$C_\chi = \max_\pi \chi(\pi).$$

Заметим, что квантовая энтропия непрерывна на компактном множестве квантовых состояний, поэтому величина χ непрерывна как функция распределения π и максимум в этой формуле действительно достигается. Введем обозначение

$$\bar{S}_\pi = \sum_x \pi_x S_x \tag{5.9}$$

для среднего выходного квантового состояния. В общем случае

$$\chi(\pi) \leq H\left(\bar{S}_\pi\right),$$

причем равенство имеет место для канала с чистыми состояниями $S_x = |\psi_x\rangle\langle\psi_x|$. Поскольку максимально возможное значение энтропии равно $\log \dim \mathcal{H}$, из этой теоремы вытекает абсолютная верхняя граница в терминах размерности выходной квантовой системы

$$C \leq \log \dim \mathcal{H}. \tag{5.10}$$

Таким образом, несмотря на то, что в гильбертовом пространстве имеется бесконечно много разных чистых состояний, это обстоятельство не может быть использовано для передачи неограниченного количества информации. Грубо говоря, чем гуще расположены векторы, тем труднее становится их различить. Верхняя граница и максимум информации достигаются, если выходные состояния являются ортогональными и равновероятными. Заметим, что такие состояния, как правило, не могут быть получены на выходе реального канала связи. Замечательно, однако, что как показывает следующий пример, ортогональность выходных состояний не является необходимой для достижения пропускной способности идеального канала.

Пример Канал с тремя чистыми состояниями ψ_0, ψ_1, ψ_2 в двумерном пространстве (см. рис. 2.3). Для равномерного распределения, $\pi_x = \frac{1}{3}$, среднее выходное состояние – хаотическое:

$$\bar{S}_\pi = \sum_{x=0}^{2} \pi_x |\psi_x\rangle\langle\psi_x| = \frac{1}{2} I, \tag{5.11}$$

откуда следует, что $C_\chi = \log 2 = 1$ бит – пропускная способность идеального канала.

Пример Двоичный канал с двумя чистыми состояниями ψ_0, ψ_1 (см. рис. 2.2). В этом случае

$$C_\chi = h_2\left(\frac{1+\epsilon}{2}\right), \tag{5.12}$$

где $\epsilon = |\langle\psi_0|\psi_1\rangle|$. В самом деле, энтропия $H(\bar{S}_\pi)$ максимизируется равномерным распределением $\pi_0 = \pi_1 = \frac{1}{2}$, поскольку в силу вогнутости энтропии (см. далее следствие 7.2.2)

$$H\left(\frac{1}{2}|\psi_0\rangle\langle\psi_0| + \frac{1}{2}|\psi_1\rangle\langle\psi_1|\right) = H\left(\frac{S+S'}{2}\right) \geq \frac{1}{2}\left(H(S) + H(S')\right),$$

где

$$S = \pi_0|\psi_0\rangle\langle\psi_0| + \pi_1|\psi_1\rangle\langle\psi_1|, \quad S' = \pi_1|\psi_0\rangle\langle\psi_0| + \pi_0|\psi_1\rangle\langle\psi_1|,$$

и $H(S') = H(S)$ в силу симметрии относительно биссектрисы угла между векторами. Аналогично решению задачи 2.3.2, получаем, что собственные значения оператора плотности $\frac{1}{2}|\psi_0\rangle\langle\psi_0| + \frac{1}{2}|\psi_1\rangle\langle\psi_1|$ равны $\frac{1\pm\epsilon}{2}$, откуда следует соотношение (5.12).

5.3 Верхняя граница

Рассмотрим произвольный c-q канал $x \to S_x$ и наблюдаемую $M = \{M_y\}$ на выходе канала. Отметим, что алфавиты $\mathcal{X}$ и $\mathcal{Y}$ могут не совпадать. Переменные x и y связаны классическим каналом $p_M(y|x) = \operatorname{Tr} S_x M_y$. Обозначим

$$\mathcal{I}_1(\pi, M) = \sum_{xy} \pi_x p_M(y|x) \log\left(\frac{p_M(y|x)}{\sum_{x'} p_M(y|x')\pi_{x'}}\right) \tag{5.13}$$

шенноновскую информацию между переменными x и y, соответствующую некоторому входному распределению $\pi = \{\pi_x\}$.

Задача 5.3.1 *Для данных π и $\mathcal{Y}$ количество информации $\mathcal{I}_1(\pi, M)$ является непрерывной и выпуклой функцией от $p_M(y|x)$ и, следовательно, от M. Указание: используйте соответствующее свойство шенноновской информации $I(X;Y)$.*

Для данного распределения π на входном алфавите рассмотрим величину

$$\sup_{\pi, M} \mathcal{I}_1(\pi, M).$$

Здесь супремум берется по всевозможным наблюдаемым на выходе канала $x \to S_x$.

Предложение 5.3.1 *Супремум количества информации $\mathcal{I}_1(\pi, M)$ достигается на наблюдаемой M^0 вида*

$$M^0_y = |\phi_y\rangle\langle\phi_y|; \quad y = 1, \ldots, m, \tag{5.14}$$

где $m \leq d^2$ в случае комплексного $\mathcal{H}$ ($m \leq \frac{d(d+1)}{2}$ в случае вещественного $\mathcal{H}$).

Доказательство. Для любого $k = 2, 3, \ldots$ множество $\mathfrak{M}_k$ наблюдаемых с k исходами компактно, и непрерывный выпуклый функционал $\mathcal{I}_1(\pi, M)$ достигает максимума на множестве $\mathfrak{M}_k$. Предположим, что $\hat{M}^0$ максимизирует $\mathcal{I}_1(\pi, M)$ на $\mathfrak{M}_k$. Отбрасывая, если необходимо, нулевые компоненты, получаем наблюдаемую $\tilde{M}^0 \in \mathfrak{M}_l, l \leq k$, для которой $\mathcal{I}_1(\pi, \tilde{M}^0) = \mathcal{I}_1(\pi, \hat{M}^0)$. Производя спектральное разложение компонент наблюдаемой $\tilde{M}^0$ в виде сумм операторов ранга один, мы можем построить новую наблюдаемую M^0 вида (5.14) с $m \geq l$, для которой $\tilde{M}^0$ является укрупнением:

$$\underbrace{M_1, M_2, \ldots}_{=\tilde{M}_1}, \underbrace{\ldots}_{=\tilde{M}_2}, \ldots, \underbrace{\ldots, M_m}_{=\tilde{M}_l}.$$

Тогда $\mathcal{I}_1(\pi, M^0) \geq \mathcal{I}_1(\pi, \tilde{M}^0)$. Это есть на самом деле утверждение о классической шенноновской информации: $I(X; f(Y)) \leq I(X; Y)$, или, эквивалентно, $H(X|f(Y)) \geq H(X|Y)$, что следует из монотонности классической условной энтропии, поскольку $H(X|Y) = H(X|f(Y), Y)$.

Функция $M \to \mathcal{I}(\pi, M)$ выпукла, поэтому мы можем предположить, что максимизирующая наблюдаемая $M^0 = \{M_y^0\}$ является крайней точкой множества $\mathfrak{M}_m$. Тогда, согласно теореме 2.2.5, операторы M_y^0 линейно независимы и из задачи 1.4.2 следует оценка $m \leq d^2$. Поскольку k было произвольным,отсюда следует, что $\sup_M \mathcal{I}_1(\pi, M)$ достигается на наблюдаемой из компактного выпуклого множества $\mathfrak{M}_{d^2}$. □

Теорема 5.3.2 *Для произвольного распределения* π

$$\max_M \mathcal{I}_1(\pi, M) \leq \chi(\pi), \tag{5.15}$$

причем равенство достигается тогда и только тогда, когда операторы $\pi_x S_x; x \in \mathcal{X}$ *коммутируют.*

В главе 7 мы получим неравенство (5.15) как следствие весьма общего свойства монотонности относительной энтропии. Здесь же мы дадим набросок оригинального доказательства, которое основано на сравнении свойств выпуклости классической и квантовой энтропий, использует достаточно элементарные факты и позволяет получить необходимое и достаточное условие достижения равенства. Это условие будет использовано в следующем разделе для установления неклассического свойства супераддитивности шенноновской информации для c-q канала.

Доказательство. Прежде всего докажем теорему в случае двух состояний S_0, S_1. Обозначим

$$\chi(t) = H((1-t)S_0 + tS_1) - (1-t)H(S_0) - tH(S_1), \quad t \in [0, 1]. \tag{5.16}$$

Положим также $S_t = (1-t)S_0 + tS_1$, $D = S_1 - S_0$ и пусть $S_t = \sum_k s_k E_k$ – спектральное разложение оператора S_t.

Задача 5.3.2 *Используя интегральную формулу Коши, получить представление*

$$S_t \log S_t = \frac{1}{2\pi i} \oint (Iz - S_t)^{-1} z \log z dz,$$

где – ветвь функции $z \log z$, аналитическая в правой полуплоскости, а интеграл берется по замкнутому контуру в правой полуплоскости, охватывающему отрезок $[\varepsilon, 1], \varepsilon > 0$.

Дифференцируя дважды по параметру t, получаем

$$[S_t \log S_t]'' = \left[\frac{1}{2\pi i} \oint \left[(Iz - S_t)^{-1} \right]'' z \log z dz \right],$$

причем

$$\left[(Iz - S_t)^{-1} \right]'' = 2 (Iz - S_t)^{-1} D (Iz - S_t)^{-1} D (Iz - S_t)^{-1}$$

$$= 2 \sum_{k,j,l} \frac{E_k D E_j D E_l}{(z - s_k)(z - s_j)(z - s_l)}.$$

Отсюда следует, что

$$\chi''(t) = -\sum_{k,j} (\operatorname{Tr} E_k D E_j D) f(s_k, s_j), \quad t > 0, \tag{5.17}$$

где

$$f(a,b) = \frac{1}{2\pi i} \oint \left[\frac{1}{(z-a)^2 (z-b)} + \frac{1}{(z-b)^2 (z-a)} \right] z \log z dz \tag{5.18}$$

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

Используя элементарное неравенство

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

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

$$\chi''(t) \leq -\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j}, \tag{5.19}$$

причем равенство достигается тогда и только тогда, когда $[D, S_t] = 0$, т. е. $[S_0, S_1] = 0$.

Задача 5.3.3 *Покажите, что оператор*

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

является решением уравнения

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

причем

$$\sum_{k,j} \operatorname{Tr} E_k D E_j D \frac{2}{s_k + s_j} = \operatorname{Tr} D L_t = \operatorname{Tr} S_t L_t^2. \tag{5.20}$$

Оператор L_t является некоммутативным аналогом логарифмической производной семейства S_t, а (5.20)– аналогом информационного количества Фишера в математической статистике.

Из (5.19), (5.20) вытекает, что

$$\chi''(t) \leq -\operatorname{Tr} D L_t = -\operatorname{Tr} S_t L_t^2, \tag{5.21}$$

в частности $\chi''(t) < 0$, так что $\chi(t)$ – вогнутая функция на отрезке $[0,1]$, причем $\chi(0) = \chi(1) = 0$.

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 и $J_M(t) = \mathcal{I}_1(\pi, M)$, где $\pi = \{1-t, t\}$. Положим также $D(y) = P_1(y) - P_0(y)$. Применяя результаты задач 5.3.2, 5.3.3 к диагональной матрице $\operatorname{diag}[P_t(y)]$ в роли состояния S_t, получаем

$$J_M''(t) = -\sum_y \frac{D(y)^2}{P_t(y)} = -\operatorname{Tr} D\Lambda_t, \tag{5.22}$$

где

$$\Lambda_t = \sum_y M_y \frac{D(y)}{P_t(y)}$$

Отсюда также вытекает, что $J_M''(t) < 0$, так что $J_M(t)$ – вогнутая функция на отрезке $[0,1]$, причем $J_M(0) = J_M(1) = 0$.

Заметим, что

$$\operatorname{Tr} D\Lambda_t = \sum_y \frac{D(y)^2}{P_t(y)} \geq \operatorname{Tr} S_t \Lambda_t^2, \tag{5.23}$$

поскольку

$$\Lambda_t^2 \leq \sum_y M_y \left[\frac{D(y)}{P_t(y)}\right]^2$$

в силу следующей разновидности неравенства Коши-Буняковского:

Задача 5.3.4 *Для любых вещественных* c_y

$$\left(\sum_y c_y M_y\right)^2 \le \sum_y c_y^2 M_y.$$

Наконец покажем, что $\operatorname{Tr} D\Lambda_t \le \operatorname{Tr} DL_t$, откуда согласно (5.21), (5.22) вытекает, что $J''_M(t) \ge \chi''(t)$, а значит $J_M(t) \le \chi(t), 0 < t < 1$, причем равенство имеет место тогда и только тогда, когда $[S_0, S_1] = 0$.

В самом деле, используя (5.23), получаем

$$\operatorname{Tr} DL_t = \operatorname{Tr} S_t L_t^2 = \operatorname{Tr} S_t \left[\Lambda_t + (L_t - \Lambda_t)\right]^2 \ge \operatorname{Tr} S_t \Lambda_t^2 + 2 \operatorname{Tr} S_t (L_t - \Lambda_t) \circ \Lambda_t$$

$$= -\operatorname{Tr} S_t \Lambda_t^2 + 2 \operatorname{Tr} (S_t \circ L_t) \Lambda_t = 2 \operatorname{Tr} D\Lambda_t - \operatorname{Tr} S_t \Lambda_t^2$$

$$= \operatorname{Tr} D\Lambda_t + \left[\operatorname{Tr} D\Lambda_t - \operatorname{Tr} S_t \Lambda_t^2\right] \ge \operatorname{Tr} D\Lambda_t.$$

Случай нескольких состояний $S_x; x = 0, 1, \ldots, k$, с распределением $\pi = \{\pi_x; x = 0, 1, \ldots, k\}$ сводится к случаю двух состояний при помощи следующего преобразования

Задача 5.3.5

$$\chi(\pi) = \sum_{m=1}^{k} (\pi_0 + \cdots + \pi_m) \chi_m(t_m),$$

где

$$\chi_m(t) = H((1-t)S_0^m + tS_m) - (1-t)H(S_0^m) - tH(S_m),$$

$$t_m = \frac{\pi_m}{\pi_0 + \cdots + \pi_m},$$

$$S_0^m = \sum_{j=0}^{m-1} \frac{\pi_j S_j}{\pi_0 + \cdots + \pi_{m-1}},$$

и аналогичного преобразования для $J_1(\pi, M)$. □

5.4 Доказательство слабого обращения

Неравенство (5.15) является основным инструментом в доказательстве слабого обращения квантовой теоремы кодирования, из которого следует, что $C \le C_\chi$.

Теорема 5.4.1 (Слабое обращение)

$$\liminf_{n\to\infty} \bar{p}_e(n, 2^{nR}) > 0 \quad \textit{при} \quad R > C_\chi. \tag{5.24}$$

Доказательство. Рассмотрим составной c-q канал $w \to S_w$, где $w = x^n$ обозначает слово длины n. Пусть $\mathcal{I}_n(\pi^{(n)}, M^{(n)})$ – шенноновская информация, определенная для этого составного канала аналогично $\mathcal{I}_1(\pi, M)$, где $\pi^{(n)} = \{\pi_w\}$ – распределение на словах длины n, а $M^{(n)}$ – наблюдаемая в $\mathcal{H}^{\otimes n}$. Тогда (5.15) влечет

$$\mathcal{I}_n(\pi^{(n)}, M^{(n)}) \leq C_\chi^{(n)}, \tag{5.25}$$

где

$$C_\chi^{(n)} = \max_{\pi^{(n)}} \chi\left(\{\pi_w\}; \{S_w\}\right) = \max_{\pi^{(n)}} \left[H\left(\sum_w \pi_w S_w\right) - \sum_w \pi_w H\left(S_w\right) \right]. \tag{5.26}$$

Лемма 5.4.2 *Последовательность $C_\chi^{(n)}$ аддитивна, т. е. $C_\chi^{(n)} = nC_\chi$.*

Доказательство. Неравенство $C_\chi^{(n)} \geq nC_\chi$ следует из аддитивности энтропии в случае, когда входные распределения $\pi^{(n)}$ являются произведениями. Доказательство обратного неравенства $C_\chi^{(n)} \leq nC_\chi$ следует из субаддитивности энтропии фон Неймана по отношению к тензорным произведениям (см. следствие 7.1.1), что влечет

$$\chi_n(\pi^{(n)}) \leq \sum_{k=1}^{n} \chi(\pi^{(n,k)}), \tag{5.27}$$

где $\pi^{(n,k)}$ есть k-е маргинальное распределение π на $\mathcal{X}$. □

Полагая

$$C_n = \max_{\pi^{(n)}, M^{(n)}} \mathcal{I}_n(\pi^{(n)}, M^{(n)}), \tag{5.28}$$

мы, таким образом, имеем $C_n \leq nC_\chi$, и применяя классическое неравенство Фано, получаем аналогично (4.36)

$$\overline{P}_e(W, M) \geq 1 - \frac{C_n}{nR} - \frac{1}{nR} \geq 1 - \frac{C_\chi}{R} - \frac{1}{nR}, \tag{5.29}$$

откуда следует (5.24).

Величина C_n равна максимальной шенноновской информации, достижимой при использовании n копий c-q канала, когда на выходе $\mathcal{H}^{\otimes n}$ допускаются измерения произвольных квантовых наблюдаемых. Следующий результат проясняет ее отношение к классической пропускной способности.

Предложение 5.4.3 *Классическая пропускная способность C канала $x \to S_x$ равна*

$$\sup_n \frac{1}{n} C_n = \lim_n \frac{1}{n} C_n. \tag{5.30}$$

Доказательство. Равенство в (5.30) вытекает из супераддитивности последовательности $\{C_n\}$:

Задача 5.4.1 *Покажите, что последовательность C_n супераддитивна, $C_{n+m} \geq C_n + C_m$, беря распределения вида $\pi^{(n+m)} = \pi^{(n)} \times \pi^{(m)}$.*

Неравенство $C \leq \sup_n \frac{1}{n} C_n$ получается, если перейти к пределу при $n \to \infty$ в первом соотношении (5.29). Противоположное неравенство $C \geq \sup_n \frac{1}{n} C_n$ вытекает из прямого утверждения классической теоремы кодирования Шеннона, примененной к составным каналам с измерением на выходе. В самом деле, покажем, что всякое значение $R < \sup_n \frac{1}{n} C_n$ является достижимой скоростью. Поскольку $n_0 R < C_{n_0}$ для некоторого n_0, мы можем выбрать наблюдаемую $M^{(n_0)}$ в $\mathcal{H}^{\otimes n_0}$, такую что

$$n_0 R < \max_{\pi^{(n_0)}} \mathcal{I}_n(\pi^{(n_0)}, M^{(n_0)}) \equiv C(M^{(n_0)}).$$

Величина $C(M^{(n_0)})$ является шенноновской пропускной способностью классического канала

$$p_M^{(n_0)}(j|x^{n_0}) = \operatorname{Tr} S_{x^{n_0}} M_j^{(n_0)}. \tag{5.31}$$

Поэтому минимальная вероятность ошибки для этого канала удовлетворяет соотношению $\tilde{p}_e(n, 2^{n(n_0 R)}) \to 0$ при $n \to \infty$. Очевидно, $\bar{p}_e(nn_0, 2^{n(n_0 R)}) \leq \tilde{p}_e(n, 2^{n(n_0 R)})$, поскольку канал (5.31) отвечает некоторому специальному блочному декодированию для исходного c-q канала. Таким образом, $\bar{p}_e(nn_0, 2^{n(n_0 R)}) \to 0$ при $n \to \infty$. Для произвольного n' можно найти n такое что $nn_0 \leq n' \leq (n+1)n_0$. Тогда

$$\bar{p}_e(n', 2^{n' R}) \leq \bar{p}_e(nn_0, 2^{(n+1)n_0 R}) \leq \bar{p}_e(nn_0, 2^{n(n_0 R')}) \to 0, \tag{5.32}$$

если R' выбрано так, что $R(1 + 1/n) \leq R' < \sup_n \frac{1}{n} C_n$ для достаточно больших n. □

Величина

$$C_1 = \max_{\pi, M} \mathcal{I}_1(\pi, M)$$

которую мы назовем *шенноновской пропускной способностью* c-q канала $x \to S_x$, представляет особый интерес: аналогично предложению 5.4.3, можно показать, что она равна пропускной способности c-q канала $x \to S_x$ с дополнительным ограничением, что на выходе составного канала допускаются только несцепленные декодирования. Здесь мы называем декодирование $M^{(n)} = \{M_j^{(n)}\}$ в $\mathcal{H}^{\otimes n}$ *несцепленным*, если

$$M_j^{(n)} = \sum_{y^n} p(j|y^n) M_{y_1}^1 \otimes \cdots \otimes M_{y_n}^n, \tag{5.33}$$

где $\{M_{y_k}^k\}$ – квантовая наблюдаемая для k-го экземпляра пространства $\mathcal{H}$, а $p(j|y^n)$ – условная вероятность, описывающая классическое преобразование результатов измерения несцепленной наблюдаемой $\{M_{y_1}^1 \otimes \cdots \otimes M_{y_n}^n\}$ на выходе.

Задача 5.4.2 *Обозначая C_u супремум достижимых скоростей для кодов (W, M) с дополнительным ограничением (5.33) на декодирования M, покажите, что $C_u = C_1$. Указание: используйте аналог первого неравенства в (5.29), чтобы доказать, что $C_u \leq C_1$. Для доказательства обратного неравенства используйте прямое утверждение классической теоремы кодирования для канала $p_M(y|x) = \operatorname{Tr} S_x M_y$, чтобы доказать, что любое значение скорости $R < \max_\pi \mathcal{I}_1(\pi, M)$ достижимо. Поскольку декодирование M в $\mathcal{H}$ произвольно, отсюда следует, что любая скорость $R < C_1$ достижима, поэтому $C_1 \leq C_u$.*

Пока что было доказано, что

$$C_1 \leq C \leq C_\chi.$$

Из прямого утверждения теоремы кодирования, которое мы собираемся доказать ниже, будет следовать, что $C = C_\chi$. Следующий результат означает, что $C_1 < C$ для каналов с существенно квантовым выходом. Таким образом, для c-q каналов без памяти возможно увеличение пропускной способности благодаря использованию сцепленных декодирований. Отсюда следует, что тогда как для классического канала без памяти $C_n = nC_1$ (см. предложение 4.3.1), в квантовом случае возможна строгая супераддитивность классической информации: $C_n > nC_1$. Причина этого кроется в том, что на выходе составного квантового канала существуют сцепленные наблюдаемые. Можно сказать, что это есть двойственное проявление корреляций ЭПР. Последние возникают, когда рассматривается сцепленное состояние составной квантовой системы, а измерения локальны, т. е. несцеплены. Строгая супераддитивность информации имеет место для состояний-произведений и обусловлена существованием сцепленных наблюдаемых.

Предложение 5.4.4 *Равенство $C_1 = C$ имеет место тогда и только тогда, когда операторы $\pi_x^0 S_x$ коммутируют для распределения π^0, максимизирующего $\chi(\pi)$.*

Доказательство. Очевидно, что условие является достаточным. Обратно, пусть $C_1 = C$, тогда используя предложение 5.3.1, компактность множества распределений $\{\pi\}$ и непрерывность шенноновской информации $\mathcal{I}_1(\pi^1, M^1)$, имеем $C_1 = \mathcal{I}_1(\pi^1, M^1)$ для некоторого распределения π^1 и наблюдаемой M^1, поэтому $\mathcal{I}_1(\pi^1, M^1) = \chi(\pi^0)$. Отсюда следует, что $\chi(\pi^1) = \chi(\pi^0)$, потому что в противном случае, используя неравенство (5.15), получаем

$$\mathcal{I}_1(\pi^1, M^1) \leq \chi(\pi^1) < \chi(\pi^0).$$

Таким образом, мы можем заменить π^0 на π^1, $\mathcal{I}_1(\pi^1, M^1) = \chi(\pi^1)$ и необходимость следует из второго утверждения теоремы 5.3.2. □

Следующие примеры (подробности см. в разделе 5.8) иллюстрируют неравенство $C_1 < C$.

Пример Для канала с тремя "равноугольными" чистыми состояниями, значение

$$C_1 = 1 - h_2\left(\frac{1+\sqrt{3}/2}{2}\right) \approx 0.645\,\text{бит} \tag{5.34}$$

достигается для не-равномерного распределения $\pi = [1/2, 1/2, 0]$ и наблюдаемой, оптимальной для двух равновероятных состояний S_0, S_1. Таким образом, $C_1 < C = 1$ и последовательность $\{C_n\}$ является строго супераддитивной. Заметим, что равномерное распределение $\pi = [1/3, 1/3, 1/3]$ и соответствующая информационно-оптимальная наблюдаемая дают меньшую величину $\log(3/2) \approx 0.585$ бит.

Пример Шенноновская пропускная способность c-q канала с двумя чистыми состояниями равна

$$C_1 = \max_{\pi, M} \mathcal{I}(\pi, M) = 1 - h_2\left(\frac{1+\sqrt{1-\epsilon^2}}{2}\right), \tag{5.35}$$

где максимум достигается для равномерного распределения ($\pi_0 = \pi_1 = \frac{1}{2}$) и четкой наблюдаемой M, задаваемой ортонормированным базисом, расположенным симметрично по отношению к векторам $|\psi_0\rangle, |\psi_1\rangle$ (см. рис. 2.2). Здесь $\epsilon = |\langle\psi_0|\psi_1\rangle| = \cos\alpha$, где α – угол между направлениями векторов. Заметим, что в этом случае информационно-оптимальная наблюдаемая совпадает с полученной согласно критерию максимального правдоподобия (пример 2.3.2), а C_1 – с пропускной способностью (4.22) соответствующего классического двоичного симметричного канала с вероятностью ошибки

$$p = \sin^2(\pi/4 - \alpha/2) = \frac{1-\sin\alpha}{2}.$$

График зависимости C_1, C от ϵ приведен на рис. 5.1.

5.5 Типичные проекторы

Перед тем как приступить к доказательству прямого утверждения теоремы кодирования для величины C_χ, рассмотрим важное понятие типичного подпространства.

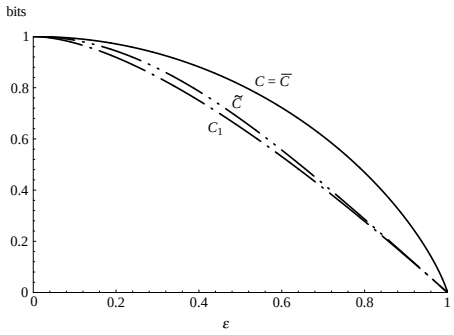


Рис. 5.1. Пропускная способность канала с двумя чистыми состояниями

Рассмотрим тензорное произведение состояний

$$S^{\otimes n} = S \otimes \cdots \otimes S$$

в пространстве $\mathcal{H}^{\otimes n}$. Пусть

$$S = \sum_j \lambda_j |e_j\rangle\langle e_j|,$$

спектральное разложение оператора S, где $\{e_j\}$ ортонормированный базис из собственных векторов, а λ_j – соответствующие собственные значения. Заметим, что λ_j образуют распределение вероятностей с энтропией

$$\sum_j \eta(\lambda_j) = H(S).$$

Тогда спектральное разложение состояния-произведения

$$S^{\otimes n} = S \otimes \cdots \otimes S$$

может быть записано в виде

$$S^{\otimes n} = \sum_J \lambda_J |e_J\rangle\langle e_J|,$$

где $J = (j_1, \ldots, j_n)$, с собственными векторами $|e_J\rangle = |e_{j_1}\rangle \otimes \cdots \otimes |e_{j_n}\rangle$ и собственными значениями $\lambda_J = \lambda_{j_1} \cdot \cdots \cdot \lambda_{j_n}$, образующими распределение вероятностей независимых одинаково распределенных случайных величин. В соответствии с определением 4.1.1, собственный вектор $|e_J\rangle$ будет называться δ-*типичным*, если

$$2^{-n\left(H(S)+\delta\right)} < \lambda_J < 2^{-n\left(H(S)-\delta\right)}.$$

Основываясь на спектральном разложении оператора $S^{\otimes n}$, мы можем определить *δ-типичный проектор* $P^{n,\delta}$ на подпространство $\mathcal{H}^{n,\delta} = P^{n,\delta}\mathcal{H}^{\otimes n}$ соотношением

$$P^{n,\delta} = \sum_{J \in T^{n,\delta}} |e_J\rangle\langle e_J|.$$

Свойства подпространства $\mathcal{H}^{n,\delta}$ типичных векторов аналогичны соответствующим свойствам подмножества $T^{n,\delta}$ типичных слов (см. раздел 4.1) и легко из них получаются:

i. Размерность $\dim \mathcal{H}^{n,\delta} = \operatorname{Tr} P^{n,\delta} = |T^{n,\delta}|$ удовлетворяет неравенству

$$\dim \mathcal{H}^{n,\delta} \leq 2^{n\left(H(S)+\delta\right)};$$

ii. Вклад не-$\delta-$типичных собственных векторов в оператор $S^{\otimes n}$ может быть сделан сколь угодно малым, т. е. для данных $\delta, \varepsilon > 0$ и достаточно больших n

$$\operatorname{Tr} S^{\otimes n}\left(I - P^{n,\delta}\right) < \varepsilon. \tag{5.36}$$

Чтобы это показать, оценим след в базисе из собственных векторов оператора $S^{\otimes n}$ и получим, что он равен

$$\begin{aligned} \sum_{J \in \bar{T}^{n,\delta}} \lambda_J &= \mathsf{P}\left\{\left|-\frac{1}{n}\log\lambda_J - H(S)\right| \geq \delta\right\} \\ &= \mathsf{P}\left\{\left|-\frac{1}{n}\sum_{k=1}^{n}\log\lambda_{j_k} - H(S)\right| \geq \delta\right\}, \end{aligned} \tag{5.37}$$

где P отвечает распределению вероятностей $\{\lambda_J\}$, а последняя вероятность может быть сделана сколь угодно малой для больших n согласно закону больших чисел.

iii. Для достаточно больших n

$$(1-\varepsilon)2^{n\left(H(S)-\delta\right)} \leq \dim \mathcal{H}^{n,\delta}.$$

Теперь можно так сформулировать некоммутативную версию свойства асимптотической равнораспределенности: при больших n состояние $S^{\otimes n}$ становится похожим на хаотическое состояние в подпространстве $\mathcal{H}^{n,\delta}$ размерности $\approx 2^{nH(S)}$.

Непосредственным приложением этого свойства является квантовый аналог сжатия данных. Как мы уже видели, размерность гильбертова пространства отражает информационный ресурс квантовой системы. С другой стороны, в квантовых вычислениях размерность отражает величину памяти, т. е. число q-битов в регистре, которое желательно минимизировать. Рассмотрим задачу кодирования *чистых* квантовых состояний другими состояниями в гильбертовом пространстве по возможности минимальной размерности, однако без существенной потери "квантовой информации", переносимой состояниями. Для этого рассмотрим квантовый

источник, который производит независимые одинаково распределенные *чистые* состояния $S_x = |\psi_x\rangle\langle\psi_x|$ с распределением p_x. Блочные состояния $S_w = |\psi_w\rangle\langle\psi_w|$, где $|\psi_w\rangle = |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle \in \mathcal{H}^{\otimes n}$, кодируются в некоторые состояния S'_w в подпространстве $\mathcal{H}_d \subset \mathcal{H}^{\otimes n}$, размерность которого $d = \dim \mathcal{H}_d$ должна быть по возможности минимальной. При этом требуется, чтобы *средняя точность воспроизведения*

$$F_n = \sum_w \pi_w \langle\psi_w|S'_w|\psi_w\rangle$$

стремилась к 1 при $n \to \infty$.

Пусть

$$\overline{S}_\pi = \sum_{x=1}^{d} \pi_x |\psi_x\rangle\langle\psi_x|$$

среднее состояние источника.

Теорема 5.5.1 (Сжатие квантовой информации) *i. Для достаточно малых $\varepsilon, \delta > 0$ существует подпространство $\mathcal{H}_d \subset \mathcal{H}^{\otimes n}$ размерности $d = 2^{n(H(\overline{S}_\pi)+\delta)}$ и операторы плотности S'_w в $\mathcal{H}_d$, такие что $F_n > 1 - \varepsilon$ для достаточно больших n;*

ii. При любом выборе подпространства $\mathcal{H}_d \subset \mathcal{H}^{\otimes n}$ с $d = 2^{n(H(\overline{S}_\pi)-\delta)}$ и операторов плотности S'_w в $\mathcal{H}_d$ имеет место $F_n < \varepsilon$ для достаточно больших n.

Доказательство. i. Для состояния источника $|\psi_w\rangle = |\psi_{x_1}\rangle \otimes \cdots \otimes |\psi_{x_n}\rangle$ определим "сжатое" состояние как

$$S'_w = \frac{P^{n,\delta}|\psi_w\rangle\langle\psi_w|P^{n,\delta}}{\langle\psi_w|P^{n,\delta}\psi_w\rangle}, \tag{5.38}$$

где $P^{n,\delta}$ – типичный проектор состояния $\overline{S}_\pi^{\otimes n}$. Оператор S'_w действует в подпространстве $\mathcal{H}^{n,\delta}$, которое имеет размерность не более $d = 2^{n\left(H(\overline{S}_\pi)+\delta\right)}$. Для точности воспроизведения получаем

$$\begin{aligned} F_n &= \sum_w \pi_w \langle\psi_w|S'_w\psi_w\rangle \\ &= \sum \pi_w \langle\psi_w|P^{n,\delta}\psi_w\rangle = \operatorname{Tr} \overline{S}_\pi^{\otimes n} P^{n,\delta}, \end{aligned}$$

так как $\overline{S}_\pi^{\otimes n} = \sum \pi_w |\psi_w\rangle\langle\psi_w|$. Благодаря второму свойству проектора $P^{n,\delta}$, мы можем ограничить это снизу величиной $1 - \varepsilon$.

ii. Учитывая, что $S'_w \le P_d$ для всех S'_w в $\mathcal{H}_d = P_d \mathcal{H}^{\otimes n}$, имеем

$$\begin{aligned} F_n &= \sum \pi_w \langle\psi_w|S'_w\psi_w\rangle \le \operatorname{Tr} \overline{S}_\pi^{\otimes n} P_d \\ &= \operatorname{Tr} \overline{S}_\pi^{\otimes n} P_{\frac{\delta}{2},n} P_d + \operatorname{Tr} \overline{S}_\pi^{\otimes n} \left(I - P_{\frac{\delta}{2},n}\right) P_d \\ &\le \|\overline{S}_\pi^{\otimes n} P_{\frac{\delta}{2},n}\| \cdot \operatorname{Tr} P_d + \operatorname{Tr} \overline{S}_\pi^{\otimes n} \left(I - P_{\frac{\delta}{2},n}\right) \end{aligned}$$

Здесь было использовано следующее свойство следа: если $T, X \geq 0$, то $\mathrm{Tr}\, TX \leq \mathrm{Tr}\, T\|X\|$, что следует из (1.18). Согласно определению типичных собственных векторов, соответствующие собственные значения ограничены сверху величиной

$$\|S^{\otimes n} P^{n,\delta}\| < 2^{-n\left(H(S)-\delta\right)}. \tag{5.39}$$

Используя это неравенство и свойство ii. типичного проектора, получаем, что для достаточно больших n величина (5.39) не превосходит

$$2^{-n\left(H(\overline{S}_\pi)-\delta/2\right)} d + \varepsilon.$$

Таким образом, если $d = 2^{n\left(H(\overline{S}_\pi)-\delta\right)}$, то $F_n \leq 2^{-n\delta/2} + \varepsilon \leq 2\varepsilon$ для достаточно больших n. □

Для доказательства прямого утверждения теоремы кодирования в следующем разделе нам также понадобится соответствующее понятие условно типичного проектора. Рассмотрим блочные состояния $S_w = S_{x_1} \otimes ... \otimes S_{x_n}$, где S_x теперь произвольное (не обязательно чистое) состояние, а $w = (x_1, ..., x_n)$ – слово входного алфавита. Обозначим

$$\bar{H}_\pi(S_{(\cdot)}) = \sum_x \pi_x H(S_x)$$

– квантовый аналог условной энтропии выхода относительно входа. Пусть $P_w = P_w^{n,\delta}$ – спектральный проектор оператора S_w, отвечающий собственным значениям в интервале $(2^{-n[\bar{H}_\pi(S_{(\cdot)})+\delta]}, 2^{-n[\bar{H}_\pi(S_{(\cdot)})-\delta]})$. Более подробно, рассмотрим спектральное разложение

$$S_x = \sum_j \lambda_j^x |e_j^x\rangle\langle e_j^x|.$$

Тогда спектральное разложение оператора S_w имеет вид

$$S_w = \sum_J \lambda_J^w |e_J^w\rangle\langle e_J^w|,$$

где $\lambda_J^w = \lambda_{j_1}^{x_1} \cdot ... \cdot \lambda_{j_n}^{x_n}$ – собственные значения, а $|e_J^w> = |e_{j_1}^{x_1}> \otimes ... \otimes |e_{j_n}^{x_n}>$ – собственные векторы. *Условно типичный проектор* определяется как

$$P_w = \sum_{J \in T_w^{n,\delta}} |e_J^w\rangle\langle e_J^w|,$$

где

$$T_w^{n,\delta} = \{J : 2^{-n[\bar{H}_\pi(S_{(\cdot)})+\delta]} < \lambda_J^w < 2^{-n[\bar{H}_\pi(S_{(\cdot)})-\delta]}\}.$$

Существенные свойства проектора P_w:

i. Согласно определению,

$$P_w \leq S_w 2^{n[\bar{H}_\pi(S_{(\cdot)})+\delta]}; \tag{5.40}$$

ii. Для $\varepsilon > 0$ и достаточно больших n

$$\mathsf{E}\,\mathrm{Tr}\, S_w(I - P_w) \leq \varepsilon, \tag{5.41}$$

где E – математическое ожидание, отвечающее распределению вероятностей

$$\mathsf{P}\{w = (x_1, \ldots, x_n)\} = \pi_{x_1} \cdot \ldots \cdot \pi_{x_n}. \tag{5.42}$$

Чтобы доказать второе свойство, рассмотрим последовательность независимых испытаний с исходами (x_l, j_l); $l = 1, \ldots, n$, где вероятность исхода (x, j) в каждом испытании равна $\pi_x \lambda_j^x$. Тогда

$$\mathsf{E}\,\mathrm{Tr}\, S_w(I - P_w) = \mathsf{P}\{J \in \overline{T_w^{n,\delta}}\}, \tag{5.43}$$

что стремится к 0 при $n \to \infty$, согласно закону больших чисел, см. задачу 4.4.1.

5.6 Доказательство прямого утверждения теоремы кодирования

Согласно следствию 5.1.2, достаточно показать, что минимальная средняя вероятность ошибки $\bar{p}_e(n, 2^{nR})$ стремится к нулю при $n \to \infty$, если $R < C_\chi$.

Рассмотрим среднее состояние

$$\bar{S}_\pi = \sum_x \pi_x S_x$$

и типичный проектор $P = P^{n,\delta}$ состояния $\bar{S}_\pi^{\otimes n}$, определенный в предыдущем разделе. Для заданного набора кодовых слов $W = \{w^{(1)}, \ldots, w^{(N)}\}$ мы можем также построить условно типичные проекторы $P_{w^{(j)}}; j = 1, \ldots, N$. Теперь мы введем специальную субоптимальную наблюдаемую M. Классическая теорема кодирования Шеннона (теорема 4.4.2) отвечает случаю диагональных операторов плотности S_x с условными вероятностями $p(y|x)$ на диагонали. Проектор P является квантовым аналогом индикатора подмножества всех δ-типичных слов на выходе, а $P_{w^{(j)}}$ – тем же для подмножества условно типичных слов на выходе при данном слове $w^{(j)}$ на входе. Однако эти проекторы не обязаны коммутировать, что делает невозможным непосредственное обобщение областей принятия решений (4.41) и соответствующее рассуждение в духе теории множеств. Поэтому мы вводим следующую наблюдаемую M в $\mathcal{H}^{\otimes n}$

$$M_j = \left(\sum_{l=1}^{N} PP_{w^{(l)}}P\right)^{-1/2} PP_{w^{(j)}}P \left(\sum_{l=1}^{N} PP_{w^{(l)}}P\right)^{-1/2}; \quad j = 1, \ldots, N. \tag{5.44}$$

Нормировка посредством квадратного корня необходима, поскольку сумма операторов $PP_{w^{(j)}}P$ может не быть равна единичному оператору, а введение $PP_{w^{(j)}}P$ вместо $P_{w^{(j)}}$ играет роль пересечения с подмножеством всех типичных слов в формуле (4.42). Оператор $(\sum_{l=1}^{N} PP_{w^{(l)}}P)^{-1/2}$ следует понимать как *обобщенный обратный* к оператору $(\sum_{l=1}^{N} PP_{w^{(l)}}P)^{1/2}$, который равен 0 на нулевом подпространстве этого оператора, содержащем в себе область значений проектора $I - P$. Обозначая $\hat{P}$ проектор на область значений оператора $\sum_{l=1}^{N} PP_{w^{(l)}}P$, имеем

$$PP_{w^{(l)}}P \le \hat{P} \le P, \quad l = 1, \ldots, N. \tag{5.45}$$

Для упрощения обозначений мы в дальнейшем будем нумеровать слова индексом w, опуская j, l. Обозначая

$$A_w = P_w P(\sum_{w'=1}^{N} PP_{w'}P)^{-1/2}$$

и используя некоммутативный аналог неравенства Коши-Буняковского (2.25) при $X = A_w, Y = I$, получаем

$$\bar{P}_e(W, M) = \frac{1}{N}\sum_{w=1}^{N}[1 - \operatorname{Tr} S_w A_w^* A_w] \le \frac{1}{N}\sum_{w=1}^{N}[1 - |\operatorname{Tr} S_w A_w|^2] \le \frac{2}{N}\sum_{w=1}^{N}[1 - \operatorname{Tr} S_w A_w],$$

где $\operatorname{Tr} S_w A_w = \operatorname{Tr} S_w P_w P(\sum_{w'=1}^{N} PP_{w'}P)^{-1/2}$ является вещественным числом, заключенным между 0 и 1. Применяя неравенство

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

получаем, учитывая (5.45)

$$-2(\sum_{w'=1}^{N} PP_{w'}P)^{-1/2} \le -3\hat{P} + \sum_{w'=1}^{N} PP_{w'}P \le -3PP_wP + \sum_{w'=1}^{N} PP_{w'}P.$$

Поэтому

$$\bar{P}_e(W, M) \le \frac{1}{N}\sum_{w=1}^{N}[2\operatorname{Tr} S_w - 3\operatorname{Tr} S_w P_w PP_w P + \sum_{w'=1}^{N} \operatorname{Tr} S_w P_w PP_{w'}P]$$

$$= \frac{1}{N}\sum_{w=1}^{N}[2\operatorname{Tr} S_w(I - P_w PP_w P) + \sum_{w':w'\neq w} \operatorname{Tr} S_w P_w PP_{w'}P].$$

Принимая во внимание оценку

$$\begin{aligned}\operatorname{Tr} S_w(I-P_wPP_wP) &= \operatorname{Tr} S_w(I-P_w)PP_wP+\operatorname{Tr} S_w(I-P)P_w\\ &\quad-\operatorname{Tr} S_w(I-P)P_w(I-P)+\operatorname{Tr} S_w(I-P_w)P+\operatorname{Tr} S_w(I-P)\\ &\le 2[\operatorname{Tr} S_w(I-P_w)+\operatorname{Tr} S_w(I-P)],\end{aligned}$$

где было использовано неравенство (1.13), имеем

$$\bar{P}_e(W,M)\le\frac{1}{N}\sum_{w=1}^{N}\{4\operatorname{Tr} S_w(I-P)+4\operatorname{Tr} S_w(I-P_w)+\sum_{w':w'\neq w}\operatorname{Tr} PS_wPP_{w'}\},\tag{5.46}$$

что является основной оценкой, аналогичной (4.43) в классическом случае.

Теперь применим процедуру случайного кодирования, вновь предполагая, что слова $w^{(1)},...,w^{(N)}$ выбираются случайным образом, независимо друг от друга, с распределением вероятностей (5.42) для каждого слова, где π – оптимальное распределение, для которого

$$H(\bar{S}_\pi)-\bar{H}_\pi(S_{(\cdot)})=C_\chi.$$

Тогда

$$\mathsf{E}S_w=\sum_{x_1\dots x_n}\pi_{x_1}\dots\pi_{x_n}S_{x_1}\otimes\dots\otimes S_{x_n}=\bar{S}_\pi^{\otimes n}.$$

Усредняя неравенство (5.46) и учитывая независимость операторов $S_w, P_{w'}$, получаем

$$\mathsf{E}\bar{P}_e(W,M)\le 4\operatorname{Tr}\bar{S}_\pi^{\otimes n}(I-P)+4\mathsf{E}\operatorname{Tr} S_w(I-P_w)+(N-1)\operatorname{Tr}\bar{S}_\pi^{\otimes n}P\mathsf{E}P_{w'}.\tag{5.47}$$

Используя неравенства (5.36), (5.41), выражающие типичность проекторов P, P_w, и неравенство (1.18), имеем

$$\mathsf{E}\bar{P}_e(W,M)\le 8\epsilon+(N-1)\|\bar{S}_\pi^{\otimes n}P\|\operatorname{Tr}\mathsf{E}P_{w'},$$

для $n\ge n(\pi,\epsilon,\delta)$. Согласно свойству (5.39) проектора P,

$$\|\bar{S}^{\otimes n}P\|\le 2^{-n[H(\bar{S}_\pi)-\delta]},$$

а по свойству $i.$ проектора P_w,

$$\operatorname{Tr}\mathsf{E}P_{w'}=\mathsf{E}\operatorname{Tr}P_{w'}\le\mathsf{E}\operatorname{Tr}S_{w'}\cdot 2^{n[\bar{H}_\pi(S_{(\cdot)})+\delta]}=2^{n[\bar{H}_\pi(S_{(\cdot)})+\delta]}.$$

Вспоминая, что $H(\bar{S}_\pi)-\bar{H}_\pi(S_{(\cdot)})=C_\chi$, получаем

$$\bar{p}_e(n,N)\le\mathsf{E}\bar{P}_e(W,M)\le 8\epsilon+N2^{-n[C_\chi-2\delta]}.$$

Таким образом, если $R\le C_\chi-3\delta$,

$$\bar{p}_e(n, 2^{nR}) \leq 8\epsilon + 2^{-n\delta},$$

откуда и следует прямое утверждение теоремы кодирования. □

Небольшая модификация этого рассуждения позволяет доказать экспоненциальное убывание средней вероятности ошибки, а именно: найдется $\beta > 0$, такое что

$$\mathsf{E}\bar{P}_e(W, M) \leq 2^{-n\beta}. \tag{5.48}$$

В самом деле, мы уже получили подобную оценку для последнего слагаемого в (5.47), при условии $N \leq 2^{n[C_\chi - 3\delta]}$, так что остается доказать это для первых двух слагаемых. Но это получается применением неравенства (4.11) к вероятностям больших уклонений (5.37), (5.43).

5.7 Функция надежности для канала с чистыми состояниями

В классическом случае чистые сигнальные состояния соответствуют вырожденным распределениям вероятностей и теорема кодирования очевидным образом дает максимальное значение пропускной способности $\log |\mathcal{X}|$. Однако для чистых квантовых неортогональных состояний утверждение теоремы кодирования остается нетривиальным, хотя ситуация все же упрощается, позволяя получить более сильный результат. Здесь мы дадим другое доказательство прямого утверждения теоремы кодирования в случае чистых состояний $S_x = |\psi_x\rangle\langle\psi_x|$, которое не использует понятия типичного проектора, однако благодаря более тонкому анализу дает оценку, позволяющую количественно оценить показатель экспоненциальной скорости сходимости вероятности ошибки.

Теорема 5.7.1 *При $R < C_\chi$ имеет место оценка*

$$\bar{p}_e(n, 2^{nR}) \leq 2 \cdot 2^{-nE(R)},$$

где

$$E(R) = \max_{0 \leq r \leq 1} [-Rr + \max_{\pi}(-\log \operatorname{Tr} \overline{S}_\pi^{1+r})] > 0. \tag{5.49}$$

Функция $E(R)$ дает нижнюю границу для так называемой *функции надежности* канала, характеризующей экспоненциальную скорость сходимости к нулю вероятности ошибки $\bar{p}_e(n, 2^{nR})$ при $n \to \infty$.

Доказательство. Рассмотрим подпространство пространства $\mathcal{H}^{\otimes n}$, порожденное кодовыми векторами $\psi_{w^{(1)}}, \ldots, \psi_{w^{(N)}}$. Как было показано в разделе 2.2.4, мы можем использовать оператор Грама $G = \sum_j |\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}|$, чтобы построить переполненную систему

$$|\hat{\psi}_j\rangle = G^{-\frac{1}{2}}|\psi_{w^{(j)}}\rangle, \quad j = 1, \ldots, N,$$

а значит, и наблюдаемую M с компонентами

$$M_j = |\hat{\psi}_j\rangle\langle\hat{\psi}_j|, \quad j = 1, \ldots, N, \tag{5.50}$$

которая может быть продолжена на все пространство $\mathcal{H}^{\otimes n}$ путем присоединения проектора M_0 на ортогональное дополнение к $\psi_{w^{(1)}}, \ldots, \psi_{w^{(N)}}$. Это будет нашим субоптимальным декодированием для составного канала. Заметим, что оно отличается от (5.44) отсутствием типичного проектора P (тогда как условно типичные проекторы в случае чистых состояний суть $P_w = |\psi_w\rangle\langle\psi_w|$). Так мы получаем оценку сверху для средней вероятности ошибки через оператор Грама кодовых векторов:

$$\begin{aligned}
\bar{P}_e(W, M) &= \frac{1}{N}\sum_{j=1}^{N}\Big(1 - \langle\hat{\psi}_j|G^{\frac{1}{2}}|\hat{\psi}_j\rangle^2\Big) \\
&\leq \frac{2}{N}\sum_{j=1}^{N}\Big(1 - \langle\hat{\psi}_j|G^{\frac{1}{2}}|\hat{\psi}_j\rangle\Big) \\
&= \frac{2}{N}\big(N - \operatorname{Tr} G^{\frac{1}{2}}\big)
\end{aligned} \tag{5.51}$$

Предполагая, что кодовые слова $\{w^{(j)}\}$ выбираются случайно, независимо и с распределением (5.42), получаем

$$\mathsf{E}\bar{P}_e(W, M) \leq \frac{2}{N}\Big(N - \mathsf{E}\operatorname{Tr} G^{\frac{1}{2}}\Big). \tag{5.52}$$

Используя неравенство

$$-2x^{\frac{1}{2}} \leq x^2 - 3x\ , \quad x \geq 0, \tag{5.53}$$

в сочетании с очевидным $-2x^{\frac{1}{2}} \leq 0$, имеем

$$-2G^{\frac{1}{2}} \leq \begin{cases} G^2 - 3G \\ 0 \end{cases}. \tag{5.54}$$

Математическое ожидание оператора Грама равно

$$\begin{aligned}
\mathsf{E}G &= \mathsf{E}\sum_{j=1}^{N}|\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}| = N\mathsf{E}|\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}| \\
&= N\sum_{x_1\ldots x_n}\pi_{x_1}\ldots\pi_{x_n}\big(|\psi_{x_1}\rangle\otimes\ldots\otimes|\psi_{x_n}\rangle\big)\big(\langle\psi_{x_1}|\otimes\ldots\otimes\langle\psi_{x_n}|\big) \\
&= N\Big[\sum_x \pi_x|\psi_x\rangle\langle\psi_x|\Big]^{\otimes n} \ = \ N\overline{S}_\pi^{\otimes n}.
\end{aligned} \tag{5.55}$$

Аналогично

$$\begin{aligned}\mathsf{E}\left(G^2-G\right) &= \mathsf{E}\Big(\sum_{j,k=1}^{N}|\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}|\psi_{w^{(k)}}\rangle\langle\psi_{w^{(k)}}| - \\ &\qquad -\sum_{j=1}^{N}|\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}|\Big) \\ &= \mathsf{E}\left(\sum_{j\neq k}|\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}|\psi_{w^{(k)}}\rangle\langle\psi_{w^{(k)}}|\right) \\ &= N(N-1)\Big(\overline{S}_{\pi}^{\otimes n}\Big)^2, \end{aligned} \tag{5.56}$$

так что в сочетании с (5.54) получаем

$$-2\mathsf{E}G^{\frac{1}{2}} \leq \begin{cases} N(N-1)\Big(\overline{S}_{\pi}^{\otimes n}\Big)^2 - 2N\overline{S}_{\pi}^{\otimes n} \\ 0 \end{cases}. \tag{5.57}$$

Пусть $\{|e_J\rangle\}$ – ортонормированный базис из собственных векторов и $\{\lambda_J\}$ – соответствующие собственные значения оператора $\overline{S}_{\pi}^{\otimes n}$. Тогда

$$-2\langle e_J|\mathsf{E}G^{-\frac{1}{2}}|e_J\rangle \leq \begin{cases} N(N-1)\lambda_J^2 - 2N\lambda_J \\ 0 \end{cases}, \tag{5.58}$$

и используя неравенства $\min(a,b) \leq a^r b^{1-r}$ и $2^{1-r} \leq 2$ при $0 \leq r \leq 1$, получаем

$$N\lambda_J \min\left[(N-1)\lambda_J, 2\right] - 2N\lambda_J \leq 2N(N-1)^r\lambda_J^{1+r} - 2N\lambda_J, \tag{5.59}$$

откуда

$$-2\mathsf{E}\operatorname{Tr}G^{\frac{1}{2}} \leq -2N + 2N(N-1)^r \operatorname{Tr}\left(\bar{S}_{\pi}^{\otimes n}\right)^{1+r}.$$

Возвращаясь к (5.52), мы можем оценить среднюю вероятность ошибки, минимизированную по всем кодам, как

$$\bar{p}_e(n,N) \leq 2N^r\left(\operatorname{Tr}\overline{S}_{\pi}^{1+r}\right)^n$$

для любого $0 \leq r \leq 1$. Полагая $N = 2^{nR}$, получаем

$$\bar{p}_e(n,2^{nR}) \leq 2\cdot 2^{n\left(Rr+\log\operatorname{Tr}\overline{S}_{\pi}^{1+r}\right)}. \tag{5.60}$$

Теперь можно использовать свободу в выборе параметров r и π. Заметим, что $-\log\operatorname{Tr}\overline{S}_{\pi}^{1+r}$ является вогнутой функцией от r, что доказывается вычислением второй производной (**Задача**). Также

$$\frac{d}{dr}\Big|_{r=0}\left(-\log\operatorname{Tr}\overline{S}_{\pi}^{1+r}\right) = -\operatorname{Tr}\overline{S}_{\pi}\log\overline{S}_{\pi} = H(\overline{S}_{\pi}). \tag{5.61}$$

Возьмем в качестве π распределение, для которого $H(\overline{S}_{\pi}) = C_{\chi}$. Тогда при $R < C_{\chi}$ имеем $\bar{p}_e(n,2^{nR}) \leq 2\cdot 2^{-nE(R)}$, где $E(R)$ дается выражением (5.49). □

5.8 Комментарии

1. Проблема определения пропускной способности квантового канала связи сформировалась в 1960-е годы и восходит к более ранним классическим трудам Габора и Бриллюэна, поставившим вопрос о квантово-механических пределах точности и скорости передачи информации (см. книги Курикши [14], Митюгова [19] и обзор Кэйвса и Драммонда [70]). Эти работы заложили физические основы и подняли проблему адекватного математического рассмотрения круга вопросов, связанного с классической пропускной способностью квантовых каналов.

2. Доказательство теоремы 5.3.2, включающее критерий достижимости верхней границы и опирающееся на сравнение свойств выпуклости величин в (5.15), было дано Холево [34].

3. Детальное сравнение величин C_1, C_χ для различных каналов проведено в работе Хирота и др. [111].

Сильное обращение

$$\lim_{n\to\infty} \bar{p}_e(n, 2^{nR}) = 1 \quad \text{при} \quad R \geq C_\chi$$

получено в работе Огава и Нагаока, см. [92].

Формула (5.35) вытекает из работы Левитина [121], который нашел максимум величины $\mathcal{I}_1(\pi, M)$ по четким наблюдаемым с двумя значениями, и из работы Шора [144]. В последней работе было показано, что в данном случае достаточно ограничиться такими наблюдаемыми, и что $\max_M \mathcal{I}(\pi, M)$ является вогнутой симметричной функцией от π, откуда вытекает, что оптимальное распределение является равномерным.

Канал с тремя чистыми состояниями был введен в работе Холево [33], где было показано, что для равномерного распределения π максимум $\max_M \mathcal{I}(\pi, M) = \log(3/2) \approx 0.585$ бит по всем наблюдаемым строго больше максимума по четким наблюдаемым, равного $\log\left(\sqrt{3}/\sqrt[3]{2}\right) \approx 0.459$ бит. Полное исследование проблемы максимизации величины $\mathcal{I}_1(\pi, M)$ для произвольного симметричного семейства чистых состояний в двумерном пространстве дано в работе Сасаки, Барнета, Джоза, Осаки и Хирота [135], где также была предложена экспериментальная реализация оптимальной процедуры измерения. Детальный анализ канала с тремя линейно независимыми чистыми состояниями в вещественном трехмерном пространстве проведен в работе Шора [144], где в частности показано, что информационно-оптимальная наблюдаемая в этой задаче имеет в общем случае 6 значений (т. е. максимальное количество, допускаемое предложением 5.3.1 в случае $d = 3$).

4. Идея сжатия данных в квантовой теории информации принадлежит Шумахеру; доказательство теоремы 5.5.1 дано в совместной работе Шумахера и Джоза [109]. Обзор результатов различных авторов, относящихся к значительно более сложной проблеме сжатия для квантового источника со смешанными состояниями имеется в книге Хайаши [92].

5. Доказательство прямой теоремы кодирования для произвольного c-q канала дано Холево [100], а также Шумахером и Вестморлендом [136]. Другое доказательство в духе подхода Чисара и Кернера [45] к классическим теоремам кодирования предложено впоследствии Винтером [156]. Этому предшествовало рассмотрение канала с чистыми состояниями в работе Джоза, Шумахера, Вестморленда, Вуттерса и Хаусладена [91]. Следует отметить, что еще до этого шенноновский метод случайного кодирования для квантового канала с чистыми почти ортогональными состояниями был применен Стратоновичем и Ванцяном [28], которые на основании первого приближения для минимальной вероятности ошибки пришли к выводу, что этот метод позволяет получить лишь значение

$$\tilde{C} = -\log \min_{\pi} \mathrm{Tr} \bar{\rho}_{\pi}^2 < C. \tag{5.62}$$

На самом деле надлежащим образом примененный метод случайного кодирования в случае канала с произвольными чистыми состояниями позволил получить даже оценку функции надежности в духе классических результатов Галлагера [5], см. работу Бурнашева и Холево [1], причем величина 5.62 оказалась равной так называемой урезанной пропускной способности (cut-off rate). В работе [1] также высказана до сих пор не доказанная гипотеза о функции надежности для общего канала со смешанными состояниями. В этом случае, как и в ряде других, рассмотрение смешанных состояний переводит проблему на качественно более высокий уровень сложности.

Часть III

6. Квантовые эволюции и каналы

6.1 Эволюции квантовой системы

В этом параграфе понятие квантового канала будет рассмотрено с общей точки зрения. Как мы видели в главе 4, классический канал полностью определяется переходной матрицей вероятностей:

$$\mathcal{X} \overset{p(y|x)}{\longrightarrow} \mathcal{Y}.$$

Такой канал описывает аффинное преобразование $p_x \to p'_y = \sum_x p(y|x)p_x$ классических состояний (распределений вероятностей) $p = \{p_x\}$ на некотором входном алфавите $\mathcal{X}$ в классические состояния $p' = \{p'_y\}$ на выходном алфавите $\mathcal{Y}$. По аналогии, в квантовом случае мы сначала рассмотрим аффинные отображения, которые переводят операторы плотности в операторы плотности $\Phi : \mathfrak{S}(\mathcal{H}) \to \mathfrak{S}(\mathcal{H})$,

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

Задача 6.1.1 *Аффинное отображение Φ множества состояний $\mathfrak{S}(\mathcal{H})$ в себя единственным образом продолжается до отображения линейного пространства $\mathfrak{T}(\mathcal{H})$ со следующими свойствами:*

i. Φ линейно: $\Phi[\sum_j c_j T_j] = \sum_j c_j \Phi[T_j]$, $c_j \in \mathbb{C}$;
ii. Φ положительно: $T \in \mathfrak{T}(\mathcal{H})$, $T \geq 0 \Rightarrow \Phi[T] \geq 0$;
iii. Φ сохраняет след: $\operatorname{Tr} \Phi[T] = \operatorname{Tr} T$, $T \in \mathfrak{T}(\mathcal{H})$.

Указание: Следуя доказательству теоремы 2.2.1, постройте линейное продолжение отображения Φ на $\mathfrak{T}(\mathcal{H})$; свойства ii., iii. доказываются непосредственной проверкой.

Определение 6.1.1 *Для любого отображения $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H})$, удовлетворяющему свойствам i.-iii., сопряженное отображение $\Phi^*: \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{H})$ определяется формулой*

$$\operatorname{Tr} \Phi[T]X = \operatorname{Tr} T\Phi^*[X], \quad T \in \mathfrak{T}(\mathcal{H}),\ X \in \mathfrak{B}(\mathcal{H}). \tag{6.1}$$

Задача 6.1.2 *Свойства отображения Φ из задачи 6.1.1 равносильны следующим:*

i. Φ^* *линейно;*
ii. Φ^* *положительно:* $X \geq 0 \Rightarrow \Phi^*[X] \geq 0$;
iii. Φ^* *унитально:* $\Phi^*[I] = I$.

Отображение Φ описывает эволюцию квантовой системы в терминах состояний (картина Шредингера), в то время как отображение Φ^* – в терминах наблюдаемых (картина Гейзенберга).

Пример 6.1.1

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

Следующий результат характеризует все обратимые эволюции.

Теорема 6.1.1 *Пусть Φ – аффинное взаимно-однозначное отображение выпуклого множества квантовых состояний $\mathfrak{S}(\mathcal{H})$ на себя, тогда*

$$\Phi[S] = USU^*, \quad S \in \mathfrak{S}(\mathcal{H}), \tag{6.2}$$

где U - унитарный или антиунитарный оператор.

В случае унитарного оператора продолжение Φ до линейного положительного отображения имеет тот же вид (6.2) для всех $S \in \mathfrak{T}(\mathcal{H})$. *Антиунитарный* оператор U характеризуется свойствами:

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

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

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

Учитывая, что в общем случае $T^* = \Lambda T^\top \Lambda^*$, получаем, что в случае антиунитарного U линейное продолжение Φ дается формулой

$$\Phi[T] = UT^*U^*, \quad T \in \mathfrak{T}(\mathcal{H}). \tag{6.3}$$

* В физике операция комплексного сопряжения связана с обращением времени.

Доказательство. Доказательство в случае q-бита ($\dim \mathcal{H} = 2$) дано в разделе 6.8; из него вытекает, что в общем случае сужение Φ на подмножество состояний, сосредоточенных на любом двумерном подпространстве $\mathcal{H}_2 \subset \mathcal{H}$, дается формулой (6.2), где U, однако, может зависеть от выбранного подпространства. Доказательство в общем случае состоит в том, чтобы, производя подходящие деформации этого подпространства, показать, что эта зависимость сводится к несущественному фазовому множителю.

Пусть $\psi_1, \psi_2 \in \mathcal{H}$. Поскольку аффинное взаимно-однозначное отображение множества квантовых состояний на себя взаимно-однозначно переводит крайние точки в крайние точки, то

$$\Phi[|\psi_j\rangle\langle\psi_j|] = |\phi_j\rangle\langle\phi_j|; \quad j = 1, 2.$$

Рассматривая двумерное подпространство $\mathcal{H}_2 \subset \mathcal{H}$, содержащее ψ_1, ψ_2 и применяя упомянутый выше результат для q-бита, получаем из формул (6.2), (6.3)

$$\Phi[|\psi_1\rangle\langle\psi_2|] = z|\phi_1\rangle\langle\phi_2| \text{ либо } \quad z|\phi_2\rangle\langle\phi_1|, \tag{6.4}$$

где $|z| = 1$. Отсюда

$$|\langle\phi_1|\phi_2\rangle| = |\langle\psi_1|\psi_2\rangle|.$$

Пусть теперь $\{e_j; j = 1, \ldots, d\}$ – ортонормированный базис в $\mathcal{H}$, тогда

$$\Phi[|e_j\rangle\langle e_j|] = |h_j\rangle\langle h_j|,$$

где $\{h_j\}$ – ортонормированный базис в $\mathcal{H}$. Из (6.4) следует, что

$$\Phi[|e_j\rangle\langle e_k|] = z_{jk}|h_j\rangle\langle h_k| \text{ либо } \quad z_{kj}|h_k\rangle\langle h_j|. \tag{6.5}$$

Если $\Phi[|e_1\rangle\langle e_2|] = z_{12}|h_1\rangle\langle h_2|$, но $\Phi[|e_1\rangle\langle e_3|] = z_{31}|h_3\rangle\langle h_1|$, то

$$\Phi[|e_1\rangle\langle e_2 + e_3|] = z_{12}|h_1\rangle\langle h_2| + z_{31}|h_3\rangle\langle h_1|$$

является оператором ранга 2, что противоречит (6.4). Повторяя это рассуждение, получаем, что в (6.5) для всех j, k имеет место одна и та же из двух альтернатив.

Рассмотрим вторую из них, когда

$$\Phi[|e_j\rangle\langle e_k|] = z_{kj}|h_k\rangle\langle h_j|,$$

тогда, полагая $\psi = e_1 + \cdots + e_d$, получаем, что оператор

$$\Phi[|\psi\rangle\langle\psi|] = \sum_{j,k=1}^{d} z_{kj}|h_k\rangle\langle h_j|$$

является проектором ранга 1, поэтому $z_{kj} = a_k\bar{a}_j$, где $|a_j| = 1$ для всех j. Обозначая U антиунитарный оператор, для которого $U|e_j\rangle = a_j h_j$, получаем (6.3). Аналогично, в случае первой альтернативы получаем $\Phi[T] = UTU^*$, где U – унитарный оператор. □

6.2 Вполне положительные отображения

Обобщая предыдущие рассмотрения, рассмотрим линейное отображение Φ, действующее из $\mathfrak{T}(\mathcal{H}_1)$ в $\mathfrak{T}(\mathcal{H}_2)$, тогда сопряженное отображение Φ^* действует из $\mathfrak{B}(\mathcal{H}_2)$ в $\mathfrak{B}(\mathcal{H}_1)$.

Определение 6.2.1 *Отображение Φ^* (Φ) называется вполне положительным, если выполнено одно из следующих эквивалентных условий:*

i. отображение $\Phi^* \otimes \mathrm{Id}_n$ положительно для всех $n = 1, 2, \ldots$, где Id_n обозначает тождественное отображение в алгебре всех $n \times n$-матриц $\mathfrak{B}_n$;

ii. для любых конечных наборов векторов $\{\varphi_j\} \subset \mathcal{H}_2$, $\{\psi_j\} \subset \mathcal{H}_1$ выполнено следующее неравенство

$$\sum_{j,k} \langle \varphi_j | \Phi[|\psi_j\rangle\langle\psi_k|]\varphi_k\rangle \geq 0.$$

Чтобы доказать эквивалентность этих условий, рассмотрим пространства $\mathcal{H}_i^{(n)}$, которые являются прямыми ортогональными суммами n копий $\mathcal{H}_i, i = 1, 2$, где n – количество векторов в наборах $\{\varphi_j\}, \{\psi_j\}$. Как мы уже видели в разделе 3.1.1, $\mathcal{H}_i^{(n)}$ может быть отождествлено с пространством $\mathcal{H}_i \otimes \ell_n^2$. Таким образом, обозначая

$$\varphi^{(n)} = \sum_{j=1}^{n} \oplus \varphi_j, \quad \psi^{(n)} = \sum_{j=1}^{n} \oplus \psi_j,$$

имеем

$$\sum_{j,k} \langle \varphi_j | \Phi[|\psi_j\rangle\langle\psi_k|]\varphi_k\rangle = \langle \psi^{(n)} | (\Phi^* \otimes \mathrm{Id}_n) \left[|\varphi^{(n)}\rangle\langle\varphi^{(n)}| \right] |\psi^{(n)}\rangle,$$

откуда вытекает $i. \Rightarrow ii.$ Обратно, любой положительный оператор $A^{(n)}$ в $\mathcal{H}_2^{(n)} \simeq \mathcal{H}_2 \otimes \ell_n^2$ может быть представлен в виде суммы положительных операторов $|\varphi^{(n)}\rangle\langle\varphi^{(n)}|$ ранга 1. Следовательно, из свойства $ii.$, означающего положительность отображения $\Phi^* \otimes \mathrm{Id}_n$ на таких операторах, вытекает $i.$

Задача 6.2.1 *Если отображение Φ удовлетворяет условию $i.$ при $n = 2$, то для него выполняется неравенство Кэдисона*

$$\Phi^*[A]^* \Phi^*[A] \leq \Phi^*[A^* A], \tag{6.6}$$

где A – произвольный оператор.

Задача 6.2.2 *Покажите, что для операции транспонирования в некотором базисе условие $i.$ нарушается уже при $n = 2$. Указание: Рассмотрите условие $ii.$, где $\{\psi_j\}$ – базис, в котором происходит транспонирование.*

Понятие вполне положительного отображения было введено Стайнспрингом, который доказал важный результат, обобщающий теорему Наймарка. Мы получим уточнение этого результата в конечномерном случае.

Теорема 6.2.1 *Для любого вполне положительного отображения* $\Phi^*\colon \mathfrak{B}(\mathcal{H}_2) \to \mathfrak{B}(\mathcal{H}_1)$ *найдутся гильбертово пространство* $\mathcal{H}_0$ *и оператор* $V\colon \mathcal{H}_1 \to \mathcal{H}_2 \otimes \mathcal{H}_0$, *такие что*

$$\Phi^*[X] = V^*(X \otimes I_0)V, \qquad X \in \mathfrak{B}(\mathcal{H}_2). \tag{6.7}$$

Двойственным образом,

$$\Phi[S] = \operatorname{Tr}_{\mathcal{H}_0} VSV^*, \qquad S \in \mathfrak{T}(\mathcal{H}_1). \tag{6.8}$$

Отображение Φ^* *унитально, т. е.* $\Phi^*[I_2] = I_1$ *(отображение* Φ *сохраненяет след) тогда и только тогда, когда оператор* V *изометричен.*

Доказательство. Рассмотрим алгебраическое тензорное произведение $\mathcal{L} = \mathcal{H}_1 \otimes \mathfrak{B}(\mathcal{H}_2)$, порожденное элементами $\Psi = \psi \otimes X,\ \psi \in \mathcal{H}_1,\ X \in \mathfrak{B}(\mathcal{H}_2)$. Введем в $\mathcal{L}$ псевдоскалярное произведение с квадратом нормы

$$\| \sum_j \psi_j \otimes X_j \|^2 = \sum_{j,k} \langle \psi_j | \Phi^*[X_j^* X_k] | \psi_k \rangle.$$

Эта величина неотрицательна, так как блочный оператор $A^{(n)} = \left[X_j^* X_k\right]_{j,k=\overline{1,n}}$ в $\mathcal{H}_2^{(n)}$ положителен. После факторизации по подпространству $\mathcal{L}_0$ элементов с нулевой нормой мы получим гильбертово пространство $\mathcal{K} = \mathcal{L}/\mathcal{L}_0$. Определим V и π соотношениями $V\psi = \psi \otimes I$ и $\pi[Y]\Psi = \pi[Y](\psi \otimes X) = \psi \otimes YX$. Нетрудно проверить, что эти определения выдерживают факторизацию. Тогда π – *-гомоморфизм алгебры $\mathfrak{B}(\mathcal{H}_2)$ в $\mathfrak{B}(\mathcal{K})$, т. е. линейное отображение, сохраняющее алгебраические операции и инволюцию: $\pi[XY] = \pi[X]\pi[Y], \pi[X^*] = \pi[X]^*$. Отображение π унитально, более того,

$$\langle \varphi | \Phi^*[X] | \psi \rangle = \langle \varphi \otimes I | \psi \otimes X \rangle = \langle \varphi | V^* \pi[X] V | \psi \rangle, \qquad X \in \mathfrak{B}(\mathcal{H}_2),$$

т. е.

$$\Phi^*[X] = V^* \pi[X] V. \tag{6.9}$$

Однако согласно лемме 6.2.2 (см. ниже) любой унитальный *-гомоморфизм алгебры $\mathfrak{B}(\mathcal{H})$ унитарно эквивалентен отображению $\pi[X] = X \otimes I_0$, где I_0 – единичный оператор в некотором гильбертовом пространстве $\mathcal{H}_0$, поэтому мы можем положить $\mathcal{K} = \mathcal{H}_2 \otimes \mathcal{H}_0$, и представление Стайнспринга (6.9) принимает вид (6.7).

Двойственное представление (6.8) вытекает из (6.7) и определения сопряженного отображения (6.1). □

Лемма 6.2.2 *Пусть π унитальный *-гомоморфизм алгебры $\mathfrak{B}(\mathcal{H})$ в алгебру $\mathfrak{B}(\mathcal{K})$, тогда найдутся гильбертово пространство $\mathcal{H}_0$ и изометричное отображение U пространства $\mathcal{K}$ на $\mathcal{H}\otimes\mathcal{H}_0$, такие что*

$$\pi[X] = U^*\left(X\otimes I_0\right)U, \quad X\in\mathfrak{B}(\mathcal{H}). \tag{6.10}$$

Доказательство. Выберем ортономированный базис $\{e_j; j=1,\dots,d\}$ в $\mathcal{H}$, и рассмотрим матричные единицы $|e_j\rangle\langle e_k|$ и их образы

$$V_{jk} = \pi[|e_j\rangle\langle e_k|]\in\mathfrak{B}(\mathcal{K}).$$

Поскольку π *-гомоморфизм, операторы V_{jk} удовлетворяют тем же алгебраическим соотношениям, что и матричные единицы

$$V_{jk}V_{lm} = \delta_{kl}V_{jm}, \quad V_{jk}^* = V_{kj}. \tag{6.11}$$

Рассмотрим подпространство $\mathcal{H}_0 = V_{11}\mathcal{K}\subseteq\mathcal{K}$ и определим $U:\mathcal{K}\to\mathcal{H}\otimes\mathcal{H}_0$ соотношением

$$U\psi = \sum_j |e_j\rangle\otimes V_{1j}\psi.$$

(Заметим, что $V_{1j}\psi = V_{11}V_{1j}\psi\in\mathcal{H}_0$.) Тогда

$$\langle U\psi_1|\left(X\otimes I_0\right)U\psi_2\rangle = \sum_{j,k}\langle e_j|Xe_k\rangle\langle V_{1j}\psi_1|V_{1k}\psi_2\rangle. \tag{6.12}$$

Поскольку в силу (6.11)

$$\langle V_{1j}\psi_1|V_{1k}\psi_2\rangle = \langle\psi_1|V_{jk}\psi_2\rangle = \langle\psi_1|\pi[|e_j\rangle\langle e_k|]\psi_2\rangle,$$

правая часть (6.12) равна $\sum_{j,k}\langle e_j|Xe_k\rangle\langle\psi_1|\pi[|e_j\rangle\langle e_k|]\psi_2\rangle = \langle\psi_1|\pi[X]\psi_2\rangle$. Отсюда следует (6.10). Полагая $X=I$ и используя унитальность π, получаем, что U – изометричное отображение $\mathcal{K}$ в $\mathcal{H}\otimes\mathcal{H}_0$.

Задача 6.2.3 *Покажите, что образ $\mathcal{K}$ при отображении U совпадает с $\mathcal{H}\otimes\mathcal{H}_0$.*

□

Представление Стайнспринга неединственно. Представление, для которого размерность $\dim\mathcal{H}_0$ минимальна, называется *минимальным*.

Теорема 6.2.3 *Пусть*

$$\Phi^*[X] = \tilde{V}^*(X\otimes\tilde{I}_0)\tilde{V}, \qquad X\in\mathfrak{B}(\mathcal{H}_2) \tag{6.13}$$

– другое представление ($\tilde{I}_0$ – единичный оператор в пространстве $\tilde{\mathcal{H}}_0$). Тогда существует частичная изометрия W_0 из $\mathcal{H}_0$ в $\tilde{\mathcal{H}}_0$, такая что

$$(I_2 \otimes W_0)V = \tilde{V}; \tag{6.14}$$

если оба представления минимальны, то W_0 изометрично отображает $\mathcal{H}_0$ на $\tilde{\mathcal{H}}_0$.

Доказательство. Для представления (6.7) рассмотрим подпространство

$$\mathcal{M} = \{(X \otimes I_0)V\psi : \psi \in \mathcal{H}_1, X \in \mathfrak{B}(\mathcal{H}_2)\} \subset \mathcal{K} = \mathcal{H}_2 \otimes \mathcal{H}_0.$$

Пространство $\mathcal{M}$ инвариантно относительно умножения на операторы вида $Y \otimes I_0$, следовательно, оно имеет вид $\mathcal{M} = \mathcal{H}_2 \otimes \mathcal{M}_0$, $\mathcal{M}_0 \subseteq \mathcal{H}_0$. Для минимального представления имеем $\mathcal{M}_0 = \mathcal{H}_0$, поскольку в противном случае найдется собственное подпространство.

Рассмотрим аналогичное подпространство $\tilde{\mathcal{M}} = \mathcal{H}_2 \otimes \tilde{\mathcal{M}}_0$ пространства $\tilde{\mathcal{K}} = \mathcal{H}_2 \otimes \tilde{\mathcal{H}}_0$ для другого представления (6.13). Определим оператор W, действующий из $\mathcal{M}$ в $\tilde{\mathcal{M}}$, равенством

$$W(X \otimes I_0)V\psi = (X \otimes \tilde{I}_0)\tilde{V}\psi. \tag{6.15}$$

Согласно (6.7), (6.13), нормы векторов и их образов при отображении W равны $\langle\psi|\Phi[X^*X]|\psi\rangle$, поэтому W изометричен.

Из (6.15) получаем, что для всех $Y \in \mathfrak{B}(\mathcal{H}_2)$ справедливо равенство

$$W(YX \otimes I_0)V\psi = (Y \otimes \tilde{I}_0)W(X \otimes \tilde{I}_0)\tilde{V}\psi$$

и, следовательно,

$$W(Y \otimes I_0) = (Y \otimes \tilde{I}_0)W \tag{6.16}$$

на $\mathcal{M}$. Продолжая оператор W на $\mathcal{K}$, полагая его равным нулю на ортогональном дополнении к $\mathcal{M}$, получаем, что (6.16) выполнено на $\mathcal{K}$. Отсюда следует, что $W = I_2 \otimes W_0$, где W_0 изометрично отображает $\mathcal{M}_0$ на $\tilde{\mathcal{M}}_0$. Из соотношения (6.15) следует (6.14). □

Следствие 6.2.1 (Представление Крауса) *Отображение Φ^* вполне положительно тогда и только тогда, когда оно может быть представлено в виде*

$$\Phi^*[X] = \sum_k V_k^* X V_k, \qquad X \in \mathfrak{B}(\mathcal{H}_2), \tag{6.17}$$

где $V_k : \mathcal{H}_1 \to \mathcal{H}_2$, или, двойственно,

$$\Phi[T] = \sum_k V_k T V_k^*, \qquad T \in \mathfrak{T}(\mathcal{H}_1). \tag{6.18}$$

Отображение Φ^* унитально (Φ сохраняет след) тогда и только тогда, когда

$$\sum V_k^* V_k = I. \tag{6.19}$$

Доказательство. Полагая $I_0 = \sum |e_j^0\rangle\langle e_j^0|$, где e_j^0 – ортонормированный базис в $\mathcal{H}_0$, рассмотрим операторы V_j, действующие из $\mathcal{H}_1$ в $\mathcal{H}_2$, которые определяются формулой

$$\langle\phi|\, V_j\psi\rangle = \langle\phi \otimes e_j^0|\, V\psi\rangle, \quad \phi \in \mathcal{H}_2, \psi \in \mathcal{H}_1. \tag{6.20}$$

Для операторов, определяемых формулой (6.20), мы будем иногда использовать обозначение $V_j = \langle e_j^0|V$. Представление (6.17) следует тогда из формулы (6.7). □

Как и представление Стайнспринга, представление Крауса неединственно. Представление Крауса с минимальным количеством ненулевых компонент называется *минимальным*.

Задача 6.2.4 *Опираясь на теорему 6.2.3, покажите, что для любых двух представлений Крауса вполне положительного отображения Φ с операторами $V_j, \tilde{V}_k$, найдется прямоугольная частично изометрическая матрица $[u_{kj}]$, такая, что $\tilde{V}_k = \sum_j u_{kj} V_j$. Если оба представления минимальны, то эта матрица унитарна.*

Задача 6.2.5 *Пусть $M = \{M_x\}$ – наблюдаемая и*

$$M_x = V^* E_x V = \tilde{V}^* \tilde{E}_x \tilde{V} \tag{6.21}$$

– два расширения Наймарка для M в пространствах $\mathcal{K}, \tilde{\mathcal{K}}$, тогда найдется частичная изометрия $W : \mathcal{K} \to \tilde{\mathcal{K}}$, такая что

$$WV = \tilde{V}, \quad WE_x = \tilde{E}_x W.$$

Если расширения минимальны в смысле размерностей $\mathcal{K}$ и $\tilde{\mathcal{K}}$, то W – изометрическое отображение $\mathcal{K}$ на $\tilde{\mathcal{K}}$.

6.3 Определение канала

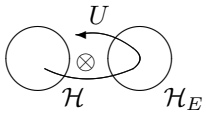


Рис. 6.1. *Открытая квантовая система.*

Дадим физическую интерпретацию свойства полной положительности. Рассмотрим (вообще говоря, необратимую) эволюцию открытой системы, взаимодействующей с окружением (см. рис. 6.1). Обозначим $\mathcal{H}$ гильбертово пространство системы, $\mathcal{H}_E$ – пространство окружения, и пусть S_E – начальное состояние окружения. Предположим, что обратимая эволюция, описывающая взаимодействие системы с окружением, задается унитарным оператором U. Тогда эволюция системы дается формулой

$$\Phi[S] = \mathrm{Tr}_{\mathcal{H}_E}\, U\,(S \otimes S_E)\, U^*. \tag{6.22}$$

Теорема 6.3.1 *Всякое вполне положительное сохраняющее след отображение Φ с $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$ может быть расширено до эволюции открытой системы, взаимодействующей с окружением, так что выполнено соотношение (6.22).*

Доказательство. Рассмотрим пространство $\mathcal{H}_E$ N-мерных векторов, где N – количество компонент в представлении Крауса для Φ. Пространство $\mathcal{H} \otimes \mathcal{H}_E$ может быть представлено как прямая сумма N копий пространства $\mathcal{H}$, т. е. как пространство векторов-столбцов $\left[\, \psi_1, \ldots, \psi_N \,\right]^\top$, $\psi_j \in \mathcal{H}$; операторы в этом пространстве суть блочные матрицы $\left[X_{jk}\right]_{j,k=1,\ldots,N}$, $X_{jk} \in \mathfrak{B}(\mathcal{H})$. Рассмотрим вектор $|\psi_E\rangle = [1, 0, \ldots, 0]^\top$, тогда

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

Введем оператор

$$U = \begin{bmatrix} V_1 & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ V_N & \ldots & \ldots \end{bmatrix}, \tag{6.23}$$

где первый столбец состоит из операторов $V_1, \ldots, V_N$, а остальные столбцы могут быть выбраны таким образом, что оператор будет унитарным; возможность такого выбора вытекает из (6.19). Имеем

$$U\left(S \otimes |\psi_E\rangle\langle\psi_E|\right) U^* = \left[V_j S V_k^*\right]_{j,k=1,\ldots,N}.$$

Частичный след в $\mathcal{H} \otimes \mathcal{H}_E$ по пространству $\mathcal{H}_E$ есть сумма диагональных элементов матрицы, что соответствует представлению Крауса для отображения Φ. □

Предыдущие рассуждения мотивируют следующее определение. Будем обозначать A систему на входе канала, B – систему на выходе.

Определение 6.3.1 *Каналом в картине Шредингера называется линейное вполне положительное сохраняющее след отображение $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$. Двойственным образом, каналом в картине Гейзенберга называется линейное вполне положительное унитальное отображение $\Phi^* : \mathfrak{B}(\mathcal{H}_B) \to \mathfrak{B}(\mathcal{H}_A)$.*

Канал Φ называется *бистохастическим*, если он отображает хаотическое состояние в $\mathcal{H}_A$ в хаотическое состояние в $\mathcal{H}_B$, $\Phi[I_A/d_A] = I_B/d_B$. Если размерности входной и выходной систем совпадают, $d_A = d_B$, то бистохастический канал Φ является унитальным. В этом случае Φ^* обладает также свойством сохранения следа, и как Φ, так и Φ^* являются каналами в картинах Шредингера и Гейзенберга.

Пусть $S = S_A$ – входное состояние канала Φ. Обозначим $\mathcal{L} = \mathrm{supp}S$ носитель оператора плотности S. Рассмотрим очищение $S_{AR} = |\psi_{AR}\rangle\langle\psi_{AR}|$ состояния S_A, где R обозначает очищающую систему и

$$|\psi_{AR}\rangle = \sum_j \sqrt{\lambda_j}|e_j\rangle \otimes |h_j\rangle; \quad \lambda_j > 0,\, e_j \in \mathcal{L},$$

см. (3.9). Обозначим через Id_R тождественное отображение в $\mathfrak{T}(\mathcal{H}_R)$, тогда тензорное произведение $\Phi \otimes \mathrm{Id}_R$ описывает канал, действующий из AR в BR, причем R не изменяется. В связи с этим R называется также *эталонной* системой. Выходным состоянием канала является

$$(\Phi \otimes \mathrm{Id}_R)[S_{AR}] = S_{BR}. \tag{6.24}$$

Ввиду того, что состояние S_{AR} в общем случае является сцепленным, такую операцию иногда называют "передачей сцепленности".

Предложение 6.3.2 *Состояние S_{BR} единственным образом определяет ограничение $\Phi_{\mathcal{L}}$ канала Φ на состояния $\tilde{S}$ с $\mathrm{supp}\tilde{S} \subset \mathcal{L}$. В частности, если S_A невырожденное, то S_{BR} единственным образом определяет канал Φ.*

Доказательство. Имеем

$$(\Phi \otimes \mathrm{Id}_R)[S_{AR}] = \sum_{j,k} \sqrt{\lambda_j}\sqrt{\lambda_k}\Phi[|e_j\rangle\langle e_k|] \otimes |h_j\rangle\langle h_k|. \tag{6.25}$$

Набор операторов $\Phi[|e_j\rangle\langle e_k|]$ единственным образом определяет $\Phi_{\mathcal{L}}$, откуда и следует утверждение. □

В частности, взяв максимально сцепленное состояние S_{AR}, для которого $\lambda_j = 1/d; j = 1, \ldots, d$, получаем из (6.25)

$$\Phi[S] = d\,\mathrm{Tr}_R\, S_{BR}(I_B \otimes S_R^\top), \tag{6.26}$$

где $S_R^\top = \sum_{j,k}\langle e_k|S|e_j\rangle|h_j\rangle\langle h_k|$. Это взаимно-однозначное соответствие между каналами и состояниями S_{BR} иногда называют *соответствием Чоя-Ямиловского.*

6.4 Каналы, разрушающие сцепленность

В этом разделе вводится важный класс каналов, содержащих классический этап переработки информации. Мы начнем рассмотрение с частных случаев.

i. Классически-квантовый (c-q) канал:

$$\Phi[S] = \sum_x \langle e_x|S|e_x\rangle S_x,$$

где $\{e_x\}$ – ортонормированный базис в $\mathcal{H}_A$, а S_x – операторы плотности в $\mathcal{H}_B$, соответствующие значениям символа x. Каналы этого типа, которые определяются отображением $x \to S_x$, описывающим кодирование классического входного сигнала x в квантовое состояние S_x, фактически уже рассматривались в главе 5.

ii. Квантово-классический (q-c) канал:

$$\Phi[S] = \sum_y |e_y\rangle\langle e_y| \operatorname{Tr} SM_y,$$

где $\{e_y\}$ - ортонормированный базис в $\mathcal{H}_B$, а $M = \{M_y\}$– наблюдаемая в $\mathcal{H}_A$. Такой канал соответствует измерению наблюдаемой M (см. далее раздел 6.5); при этом возникает распределение вероятностей на множестве исходов измерения. Сопряженный канал действует по формуле

$$\Phi^*[X] = \sum_y \langle e_y|X|e_y\rangle M_y.$$

Представление Стайнспринга (6.7) для q-c канала Φ связано с расширением Наймарка для наблюдаемой M: полагая $X = |e_y\rangle\langle e_y|$, получаем

$$M_y = V^* \left(|e_y\rangle\langle e_y| \otimes I_0\right) V,$$

где $E = \{|e_y\rangle\langle e_y| \otimes I_0\}$ – четкая наблюдаемая в $\mathcal{K}$.

iii. Композиция q-c и c-q каналов с одним и тем же ортонормированным базисом $\{e_j\}$ дает q-c-q канал

$$\Phi[S] = \sum_j S_j \operatorname{Tr} SM_j, \tag{6.27}$$

для которого *i*. и *ii*. являются частными случаями. Такой канал содержит промежуточное классическое звено преобразования информации (в котором она представляется символом j).

Определение 6.4.1 *Канал Φ из A в B обладает свойством разрушения сцепленности, если для произвольного входного состояния S_{AR} канала $\Phi \otimes \mathrm{Id}_R$ его выходное состояние S_{BR} – несцепленное, т. е. является смесью состояний-произведений*

$$S_{BR} = \sum_j p_j S_B^j \otimes S_R^j, \tag{6.28}$$

где $\{p_j\}$– распределение вероятностей.

Понятие сцепленности состояний будет подробнее рассмотрено в разделе 7.4.

Предложение 6.4.1 *Для того, чтобы канал обладал свойством разрушения сцепленности, необходимо и достаточно, чтобы он был q-c-q каналом.*

Доказательство. Действительно, для канала (6.27) и любого входного состояния S_{AR} канала $\Phi \otimes \mathrm{Id}_R$ выполняется соотношение (6.28) с $S_B^j = S_j, p_j S_R^j = \mathrm{Tr}_A\, S_{AR} M_j$. Обратно, пусть канал Φ разрушает сцепленность, тогда, беря в качестве S_{AR} максимально сцепленное состояние и используя соотношения (6.28), (6.26), получаем представление (6.27), в котором $S_j = S_B^j, M_j = dp_j (S_R^j)^T$. □

Пример 6.4.1

Важным частным случаем c-q канала является *полностью деполяризующий канал*, который описывает необратимую эволюцию к финальному состоянию S_f:

$$\Phi[S] = S_f \cdot \mathrm{Tr}\, S, \quad S \in \mathfrak{T}(\mathcal{H}).$$

Пример 6.4.2

Пусть $p \in [0,1]$, тогда канал $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H} \oplus \mathbb{C})$ определяемый соотношением

$$\Phi(S) = \begin{bmatrix} (1-p)S & 0 \\ 0 & p\mathrm{Tr}S \end{bmatrix}$$

сохраняет входное состояние S с вероятностью $1-p$ и "стирает" его с вероятностью p, посылая сигнал о стирании. Этот канал называется *квантовым стирающим каналом.*

Задача 6.4.1 *Докажите полную положительность и найдите разложения Крауса для всех рассмотренных выше каналов.*

6.5 Процессы квантовых измерений

Важнейший пример необратимой эволюции — изменение состояния квантовой системы в результате производимого над ней измерения. *Полное идеальное* квантовое измерение связывается с ортонормированным базисом $|e_x\rangle$, векторы которого индексированы возможными исходами измерения x. Постулируется, что система, находящаяся перед измерением в состоянии S, в результате такого измерения переходит в состояния

$|e_x\rangle\langle e_x|$ с вероятностями $\langle e_x|\, Se_x\rangle$. Таким образом, статистический ансамбль после измерения разбивается на подансамбли, соответствующие различным исходам x, и в целом описывается состоянием

$$S' = \sum_x |e_x\rangle\langle e_x|\, Se_x\rangle\langle e_x|. \tag{6.29}$$

Отметим, что отображение $S \to S'$ является частным случаем q-c канала, отвечающего наблюдаемой $M = \{|e_x\rangle\langle e_x|\}$.

При *неполном* идеальном измерении некоторые исходы группируются, порождая 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*, идеальное измерение четкой наблюдаемой E дает значение x с вероятностью

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

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

$$S_x = \frac{E_x S E_x}{\operatorname{Tr} E_x S E_x}, \quad \text{если} \quad p_x > 0.$$

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

$$S' = \sum_x p_x S_x = \sum_x E_x S E_x = \Phi[S]. \tag{6.30}$$

Заметим, что в отличие от классической формулы Байеса, квантовое состояние (6.30) всего ансамбля после идеального измерения может отличаться от начального состояния S. Таким образом, даже идеальное квантовое измерение не сводится к простому наблюдению выходного символа и включает в себя воздействие, которое изменяет состояние системы, даже если исходы "не считываются". В этом принципиальное отличие квантовых наблюдаемых от классических случайных величин, наблюдение которых не изменяет статистический ансамбль, а сводится к простому отбору его представителей в соответствии со значениями случайных величин.

В картине Шредингера изменение системы в результате измерения описывается вполне положительным отображением

$$\Phi^*[X] = \sum_x E_x X E_x = \Phi[X], \tag{6.31}$$

которое может быть интерпретировано как условное математическое ожидание относительно алгебры операторов, коммутирующих со всеми проекторами E_x; при этом выполнены характеристические свойства:

i. $\Phi^2 = \Phi$;
ii. $\Phi[X\Phi[Y]] = \Phi[X]\Phi[Y]$, $X, Y \in \mathfrak{B}(\mathcal{H})$.

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 $\mathcal{H}_0$, над которой затем производится идеальное измерение четкой наблюдаемой E^0. Тогда, в соответствии с постулатом фон Неймана-Людерса, вероятность исхода x равна

$$p_x = \operatorname{Tr} U(S \otimes S_0)U^*(I \otimes E^0_x),$$

а апостериорное состояние описывается оператором плотности S_x, удовлетворяющим соотношению

$$p_x S_x = \operatorname{Tr}_{\mathcal{H}_0} U(S \otimes S_0)U^*(I \otimes E^0_x) \equiv \Phi_x[S]. \tag{6.32}$$

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 описывают 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 (не-единственный) инструмент и косвенное измерение. Рассмотрим произвольное представление $M_x = V^*_x V_x$; построим по нему унитарный оператор (6.23) в пространстве $\mathcal{H} \otimes \mathcal{H}_E$, тогда семейство вполне положительных отображений

$$\Phi_x[S] = \operatorname{Tr}_{\mathcal{H}_E} U(S \otimes |\psi_E\rangle\langle\psi_E|)U^*(I \otimes |e_x\rangle\langle e_x|), \tag{6.33}$$

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

$$M_x = \mathrm{Tr}_{\mathcal{H}_E} \left(I \otimes |\psi_E\rangle\langle\psi_E|\right) U^* \left(I \otimes |e_x\rangle\langle e_x|\right) U = \Phi_x^*[I]. \tag{6.34}$$

Таким образом, процесс косвенного измерения наблюдаемой M реализуется пробной системой $\mathcal{H}_E$ в начальном состоянии $|\psi_E\rangle\langle\psi_E|$, унитарным оператором U в пространстве $\mathcal{H} \otimes \mathcal{H}_E$, и четкой наблюдаемой $\{|e_x\rangle\langle e_x|\}$ в $\mathcal{H}_E$.

Задача 6.5.1 *Для полного идеального измерения с $M_x = |e_x\rangle\langle e_x|$ положим $V_x = M_x$ и построим унитарный оператор U по формуле (6.23) (в этом случае $\mathcal{H}_E = \mathcal{H}$). Покажите, что*

$$U(|e_x\rangle \otimes |e_1\rangle) = |e_x\rangle \otimes |e_x\rangle.$$

Для произвольного вектора $\psi \in \mathcal{H}$

$$U(|\psi\rangle \otimes |e_1\rangle) = \sum_x \langle e_x|\psi\rangle |e_x\rangle \otimes |e_x\rangle,$$

таким образом, взаимодействие U порождает сцепленность между исходной и пробной системами, которая позволяет свести идеальное измерение в исходной системе к измерению четкой наблюдаемой $\{|e_x\rangle\langle e_x|\}$ в пробной системе.

6.6 Комплементарные каналы

Рассмотрим три квантовые системы A, B, C с пространствами $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ и линейный оператор $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_C$; соотношения

$$\Phi_B[S] = \mathrm{Tr}_{\mathcal{H}_C} VSV^*, \quad \Phi_C[S] = \mathrm{Tr}_{\mathcal{H}_B} VSV^*; \quad S \in \mathfrak{T}(\mathcal{H}_A) \tag{6.35}$$

определяют два вполне положительных отображения $\Phi_B : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$, $\Phi_C : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_C)$, которые мы называем взаимно *комплементарными* . Если V – изометрия, то оба отображения сохраняют след, то есть являются каналами.

Из теоремы Стайнспринга вытекает, что для заданного вполне положительного отображения всегда существует комплементарное. Кроме того, комплементарное отображение единственно в следующем смысле: для данного вполне положительного отображения Φ_B, любые два отображения $\Phi_C, \Phi_{C'}$, комплементарные к Φ_B эквивалентны в том смысле, что найдется частичная изометрия $W : \mathcal{H}_C \to \mathcal{H}_{C'}$, такая что

$$\Phi_{C'}[S] = W\Phi_C[S]W^*, \quad \Phi_C[S] = W^*\Phi_{C'}[S]W, \tag{6.36}$$

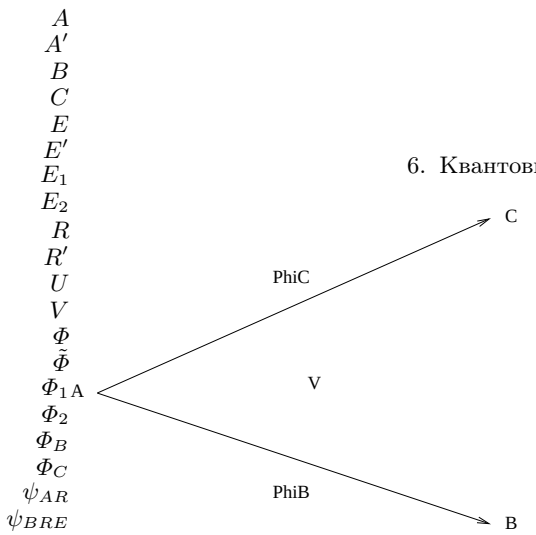


Рис. 6.2. *Комплементарный канал.*

для всех S. Если размерности пространств $\mathcal{H}_C, \mathcal{H}_{C'}$ минимальны, то W является изометрией $\mathcal{H}_C$ на $\mathcal{H}_{C'}$. Таким образом, комплементарность есть отношение на классах эквивалентности вполне положительных отображений.

Если $A = B$ – открытая квантовая система, взаимодействующая с окружением $C = E$ в произвольном начальном состоянии S_E, а Φ_B – соответствующий канал

$$\Phi_B[S] = \mathrm{Tr}_E U(S \otimes S_E)U^*, \tag{6.37}$$

описывающий изменение состояния A, то конечное состояние окружения является выходным для канала

$$\Phi_E[S] = \mathrm{Tr}_B U(S \otimes S_E)U^*, \tag{6.38}$$

Мы будем называть каналы Φ_B, Φ_E *слабо комплементарными*. Если же состояние окружения – чистое, $S_E = |\psi_E\rangle\langle\psi_E|$, то, вводя изометрию $V = U|\psi_E\rangle : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$, мы получаем, что Φ_E является комплементарным каналом к Φ_B. Если состояние S_E не является чистым, то, рассматривая его очищение $S_{E'}$ в пространстве $\mathcal{H}_{E'} = \mathcal{H}_E \otimes \mathcal{H}_R$ и полагая $U' = U \otimes I_R$, получаем представление вида (6.38) с заменой E на E', где $S_{E'}$ – уже чистое состояние; следовательно, $\Phi_{E'}$ комплементарен к Φ_B и $\Phi_E[S] = \mathrm{Tr}_R \Phi_{E'}[S]$. В отличие от комплементарного, слабо комплементарный канал не единствен и зависит от представления (6.37).

Для упрощения формул будем использовать обозначение $\tilde{\Phi}$ для отображения, комплементарного к Φ.

Предположим, что вполне положительное отображение $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H}')$ задано представлением Крауса

$$\Phi[S] = \sum_{\alpha=1}^{\tilde{d}} V_\alpha S V_\alpha^*, \tag{6.39}$$

тогда комплементарное отображение $\tilde{\Phi} : \mathfrak{T}(\mathcal{H}) \to \mathfrak{M}_{\tilde{d}}$ задается формулой:

$$\tilde{\Phi}[S] = \left[\mathrm{Tr} V_\alpha S V_\beta^*\right]_{\alpha,\beta=\overline{1,\tilde{d}}} = \sum_{\alpha,\beta=1}^{\tilde{d}} \left(\mathrm{Tr} S V_\beta^* V_\alpha\right) |e_\alpha\rangle\langle e_\beta|, \tag{6.40}$$

где $\{e_\alpha\}$ – канонический базис в координатном пространстве $\ell^2_{\tilde{d}}$, которое играет роль $\mathcal{H}_C$, так что $\tilde{d} = d_C$. Это утверждение вытекает из того, что $V = \sum_{\alpha=1}^{\tilde{d}} \oplus V_\alpha$ является отображением $\mathcal{H}_A$ в $\sum_{\alpha=1}^{\tilde{d}} \oplus \mathcal{H}_B \simeq \mathcal{H}_B \otimes \mathcal{H}_C$, для которого $\Phi, \tilde{\Phi}$ задаются частичными следами (6.35), см. задачу 3.1.4. Записывая след в пространстве $\mathcal{H}_B$ в ортонормированном базисе $\{e_j^B\}$, получаем представление Крауса для комплементарного отображения

$$\tilde{\Phi}[S] = \sum_{j=1}^{d_B} \tilde{V}_j S \tilde{V}_j^*, \tag{6.41}$$

где $\langle e_\alpha | \tilde{V}_j = \langle e_j^B | V_\alpha$ (см. определение (6.20)).

Задача 6.6.1 *Покажите непосредственным вычислением, что применение аналогичной процедуры к $\tilde{\Phi}$ дает отображение $\tilde{\tilde{\Phi}}$, которое унитарно эквивалентно Φ.*

Пример Рассмотрим канал

$$\Phi[S] = S \otimes S_C, \tag{6.42}$$

где S_C – фиксированное состояние в пространстве $\mathcal{H}_C$. Если S_C – чистое состояние, то этот канал унитарно эквивалентен идеальному (тождественному) каналу Id. Записывая спектральное разложение

$$S_C = \sum_{\alpha=1}^{\tilde{d}} \lambda_\alpha |e_\alpha\rangle\langle e_\alpha|,$$

получаем представление Крауса с операторами $V_\alpha = \sqrt{\lambda_\alpha}(I \otimes |e_\alpha\rangle)$, следовательно, из (6.40) вытекает

$$\tilde{\Phi}[S] = \left[\lambda_\alpha \delta_{\alpha\beta} \mathrm{Tr} S\right]_{\alpha,\beta=\overline{1,\tilde{d}}} \simeq S_C \mathrm{Tr} S,$$

т. е. $\tilde{\Phi}$ – полностью деполяризующий канал. Для произвольного состояния S_C идеальный канал Id слабо комплементарен к этому каналу.

Этот пример иллюстрирует тот факт, что идеальная передача квантовой информации из A в B равносильна отсутствию утечки информации из A в окружение C, и наоборот. Далее мы увидим, что приближенная версия этого свойства – чем канал Φ ближе к идеальному, тем комплементарный канал $\tilde{\Phi}$ ближе к полностью деполяризующему – играет решающую роль в теореме кодирования для квантовой информации.

Полностью деполяризующий канал является частным случаем канала, разрушающего сцепленность (q-c-q) (см. раздел 6.4); вычисления могут быть обобщены на весь этот класс. Пусть $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$ канал вида (6.27), тогда он имеет представление Крауса с операторами V_α ранга 1:

$$\Phi[S] = \sum_{\alpha=1}^{\tilde{d}} |\varphi_\alpha\rangle\langle\psi_\alpha|S|\psi_\alpha\rangle\langle\varphi_\alpha|, \tag{6.43}$$

где

$$\sum_{\alpha=1}^{\tilde{d}} |\psi_\alpha\rangle\langle\varphi_\alpha|\varphi_\alpha\rangle\langle\psi_\alpha| = I.$$

Задача 6.6.2 *Используя спектральное разложение операторов S_j, M_j в (6.27), покажите, что существование представления Крауса вида (6.43) является необходимым и достаточным для того, чтобы канал Φ разрушал сцепленность.*

Комплементарный канал $\tilde{\Phi} : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{M}_{\tilde{d}}$ имеет вид

$$\tilde{\Phi}[S] = [c_{\alpha\beta}\langle\psi_\alpha|S|\psi_\beta\rangle]_{\alpha,\beta=\overline{1,\tilde{d}}} = \sum_{\alpha,\beta=1}^{\tilde{d}} c_{\alpha\beta}|e_\alpha\rangle\langle\psi_\alpha|S|\psi_\beta\rangle\langle e_\beta|, \tag{6.44}$$

где $c_{\alpha\beta} = \langle\varphi_\beta|\varphi_\alpha\rangle$. Заметим, что произвольная положительно определенная матрица $[c_{\alpha\beta}]$ может быть представлена как матрица Грама некоторой системы векторов в гильбертовом пространстве (разложение Колмогорова). В частном случае, когда система $\{\psi_\alpha\}_{\alpha=\overline{1,\tilde{d}}}$ является ортонормированным базисом в $\mathcal{H}$, соотношение (6.44) задает вполне положительное отображение, называемое *диагональным*. Из (6.43) мы видим, что диагональные каналы являются комплементарными к c-q каналам. Для q-c каналов, $\{\varphi_\alpha\}_{\alpha=\overline{1,\tilde{d}}}$ является ортонормированным базисом в $\mathcal{H}$, так что $c_{\alpha\beta} = \delta_{\alpha\beta}$, и комплементарный канал также является q-c каналом.

Используя разложение Колмогорова $c_{\alpha\beta} = \sum_j \bar{v}_{\beta j} v_{\alpha j}$ и обозначая

$$\tilde{V}_j = \sum_{\alpha=1}^{\tilde{d}} v_{\alpha j}|e_\alpha\rangle\langle\psi_\alpha|, \tag{6.45}$$

получаем представление Крауса (6.41) для комплементарного канала. Для диагональных отображений $|\psi_\alpha\rangle = |e_\alpha\rangle$, следовательно, из (6.44) можно видеть, что диагональные каналы характеризуются представлением Крауса с одновременно диагонализуемыми операторами $\tilde{V}_j$.

Определение 6.6.1 *Пусть $\{p_j\}$ – конечное распределение вероятностей и $\Phi_j : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\mathcal{H}'_j)$ – некоторое семейство каналов. Канал $\Phi : \mathfrak{T}(\mathcal{H}) \to \mathfrak{T}(\sum_j \oplus \mathcal{H}'_j)$ называется ортогональной выпуклой комбинацией каналов Φ_j,*

$$\Phi = \sum_j \oplus p_j \Phi_j,$$

если $\Phi[S] = \sum_j \oplus p_j \Phi_j[S]$ *для всех* $S \in \mathfrak{S}(\mathcal{H})$.

Задача 6.6.3 *Покажите, что канал, комплементарный к ортогональной выпуклой комбинации каналов является ортогональной выпуклой комбинацией комплементарных каналов:*

$$\widetilde{\sum_j \oplus p_j \Phi_j} = \sum_j \oplus p_j \tilde{\Phi}_j.$$

В частности, используя пример (6.42) (с $\tilde{d} = \dim \mathcal{H}_C = 1$*), получаем, что комплементарный к стирающему каналу* Φ_p *унитарно эквивалентен стирающему каналу* Φ_{1-p}.

6.7 Ковариантные каналы

Прежде чем мы перейдем к дальнейшим примерам, рассмотрим важную конструкцию – дискретную версию операторов Вейля и канонических коммутационных соотношений (ККС) в конечномерном гильбертовом пространстве. Фиксируем ортонормированный базис $\{e_k; k = 1, \ldots, d\}$ в $\mathcal{H}$. *Дискретными операторами Вейля* называются унитарные операторы в $\mathcal{H}$, определяемые соотношением

$$W_{\alpha\beta} = U^\alpha V^\beta; \quad \alpha, \beta = 0, \ldots, d-1, \tag{6.46}$$

где

$$V|e_k\rangle = \exp\left(\frac{2\pi i k}{d}\right)|e_k\rangle; \quad U|e_k\rangle = |e_{k+1(\mathrm{mod} d)}\rangle; \quad k = 0, \ldots, d-1. \tag{6.47}$$

Заметим, что $W_{00} = I$. Операторы Вейля удовлетворяют дискретному аналогу канонических коммутационных соотношений Вейля - Сигала для бозонных систем, которые будут подробно рассмотрены в главе 11:

$$\begin{aligned} W_{\alpha\beta} W_{\alpha'\beta'} &= \exp \frac{2\pi i \beta \alpha'}{d} W_{\alpha+\alpha', \beta+\beta'} \\ &= \exp \frac{2\pi i \left(\beta' \alpha - \beta \alpha'\right)}{d} W_{\alpha'\beta'} W_{\alpha\beta}. \end{aligned} \tag{6.48}$$

Первое равенство выражает тот факт, что отображение $(\alpha, \beta) \to W_{\alpha\beta}$ является проективным представлением аддитивной циклической группы $\mathbb{Z}_d \times \mathbb{Z}_d$.

Заметим, что оператор A, коммутирующий со всеми операторами Вейля $W_{\alpha,\beta}$, кратен единичному оператору. В самом деле, из $[A, V] = 0$,

согласно (6.47), вытекает, что $A|e_k\rangle = a_k|e_k\rangle$, а из $[A,U]=0$ следует $a_k \equiv a$. Отсюда вытекает важное тождество

$$\sum_{\alpha,\beta=0}^{d-1} W_{\alpha\beta} S W_{\alpha\beta}^* = d(\operatorname{Tr} S) I, \tag{6.49}$$

так как из (6.48) следует, что левая часть коммутирует со всеми операторами Вейля. Константа в правой части находится из равенства следов. В действительности, равенство (6.49) является следствием соотношений ортогональности для неприводимого унитарного представления произвольной компактной группы.

Заметим, что в случае $d=2$ дискретные операторы Вейля сводятся к матрицам Паули; а именно, полагая $|e_0\rangle = |\uparrow\rangle, |e_1\rangle = |\downarrow\rangle$, имеем

$$W_{01} = V = \sigma_z, \quad W_{10} = U = \sigma_x, \quad W_{11} = UV = -i\sigma_y, \tag{6.50}$$

при этом дискретные ККС переходят в правила умножения (2.17).

Определение 6.7.1 *Деполяризующий канал в* $\mathcal{H}, \dim \mathcal{H} = d$, *определяется соотношением*

$$\Phi[\mathcal{S}] = (1-p)S + p\frac{I}{d}\operatorname{Tr} S, \quad 0 \le p \le \frac{d^2}{d^2-1}. \tag{6.51}$$

Если $p \le 1$, то это соотношение описывает смесь идеального канала Id и полностью деполяризующего канала, который переводит любое состояние в хаотическое $\bar{S} = \frac{I}{d}$. Для любого значения параметра $0 \le p \le \frac{d^2}{d^2-1}$ полная положительность следует из представления Крауса

$$\Phi[\mathcal{S}] = \left(1 - p\frac{d^2-1}{d^2}\right) S + \frac{p}{d^2} \sum_{\alpha+\beta>0} W_{\alpha\beta} S W_{\alpha\beta}^*, \tag{6.52}$$

которое получается при подстановке выражения для $\operatorname{Tr} S$ из (6.49) в (6.51). В этом выражении все коэффициенты положительны, что следует из условия $0 \le p \le \frac{d^2}{d^2-1}$.

Деполяризующий канал удовлетворяет соотношению $\Phi^* = \Phi$ и является унитальным, а значит, и бистохастическим.

Пусть G – группа и $g \to V_g^j; g \in G;$ – два (проективных) унитарных представления группы G в $\mathcal{H}_j; j = A, B$. Канал $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$ называется *ковариантным*, если

$$\Phi[V_g^A S (V_g^A)^*] = V_g^B \Phi[S] (V_g^B)^* \tag{6.53}$$

для всех $g \in G$ и S. Если представление V_g^B – неприводимое, то канал Φ является бистохастическим. В самом деле, из свойства $V_g^B \Phi[I_A] = \Phi[I_A] V_g^B$ (для всех g) и из неприводимости V_g^B следует, что оператор $\Phi[I_1]$ пропорционален I_B, а коэффициент пропорциональности получается из равенства следов.

Задача 6.7.1 *Покажите, что деполяризующий канал характеризуется свойством унитарной ковариантности: $\Phi[U\mathcal{S}U^*] = U\Phi[\mathcal{S}]U^*$ для произвольного унитарного оператора U в $\mathcal{H}$.*

Задача 6.7.2 *Покажите, что канал, комплементарный к ковариантному каналу, сам является ковариантным:*

$$\Phi[V_g^A S (V_g^A)^*] = V_g^E \Phi[S](V_g^E)^*,$$

где $g \to V_g^E$ – (проективное) унитарное представления группы G в $\mathcal{H}_E$. Указание: используйте тот факт, что для произвольного ковариантного канала имеется представление Крауса (6.18), в котором V_j удовлетворяют соотношениям

$$V_g^B V_j (V_g^A)^* = \sum_k d_{jk}(g) V_k,$$

где $g \to D(g) = [d_{jk}(g)]$ некоторое матричное унитарное представление группы G. Отсюда следует, что комплементарное отображение (6.40) ковариантно, причем роль второго представления V_g^E играет $D(g)$.

6.8 q-битные каналы

Рассмотрим каналы в пространстве состояний q-бита $\mathcal{H}_2$, которые соответствуют аффинным преобразованиям шара Блоха (единичного шара в $\mathbb{R}^3$), удовлетворяющим дополнительному ограничению, налагаемому полной положительностью.

Предложение 6.8.1 *Произвольное положительное сохраняющее след отображение Φ пространства $\mathfrak{T}(\mathcal{H}_2)$ может быть представлено в виде*

$$\Phi[S] = U_2 \Lambda[U_1 S U_1^*] U_2^*, \tag{6.54}$$

где U_1, U_2 – унитарные операторы, а Λ имеет следующую каноническую форму в базисе матриц Паули

$$\Lambda[I] = I + \sum_{\gamma=x,y,z} t_\gamma \sigma_\gamma, \quad \Lambda[\sigma_\gamma] = \lambda_\gamma \sigma_\gamma, \quad \gamma = x, y, z, \tag{6.55}$$

а λ_γ, t_γ – действительные числа.

Доказательство. Ограничение Φ на множество состояний q-бита является аффинным преобразованием, поэтому оно отображает состояние $S(\boldsymbol{a})$, задаваемое соотношением (2.8), в состояние $S(T\boldsymbol{a}+\boldsymbol{b})$, где T – вещественная 3×3-матрица, $\boldsymbol{b}$– вектор в $\mathbb{R}^3$. Используя полярное разложение для T и спектральное разложение для $|T|$, получаем

$$T = O|T| = O_2 L O_1,$$

где O, O_1, O_2 – ортогональные матрицы, $\det O_1 = \det O_2 = 1$, а $L = \mathrm{diag}[\lambda_x, \lambda_y, \lambda_z]$ – диагональная матрица с диагональным элементами λ – вещественными, но необязательно неотрицательными. Таким образом,

$$\Phi[S(\boldsymbol{a})] = S(O_2[L(O_1\boldsymbol{a}) + \boldsymbol{t}]), \tag{6.56}$$

где $\boldsymbol{t} = O_2^{-1}\boldsymbol{b}$.

Матрицы O_1, O_2 описывают повороты шара Блоха в $\mathbb{R}^3$. Покажем, что они отвечают преобразованиям квантовых состояний, которые порождаются унитарными операторами U в $\mathcal{H}$ согласно формуле (6.2). Достаточно рассмотреть поворот O вокруг оси z на угол φ. При этом единичный вектор $\boldsymbol{a}$ с углами Эйлера θ, ϕ преобразуется в вектор $O\boldsymbol{a}$ с углами $\theta, \phi+\varphi$, поэтому вектор состояния (2.9) в $\mathcal{H}$ переходит в вектор

$$|\psi(O\boldsymbol{a})\rangle = \begin{bmatrix} \cos\frac{\theta}{2} e^{-i(\phi+\varphi)/2} \\ \sin\frac{\theta}{2} e^{i(\phi+\varphi)/2} \end{bmatrix} = U|\psi(\boldsymbol{a})\rangle, \tag{6.57}$$

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

$$S(O\boldsymbol{a}) = US(\boldsymbol{a})U^* \tag{6.58}$$

для всех $\boldsymbol{a}$.

Учитывая (6.56), получаем доказываемое соотношение (6.54). □

Задача 6.8.1 *Покажите, что повороту шара Блоха на угол φ вокруг оси $\boldsymbol{a}$ отвечает унитарный оператор*

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

Рассмотрим подробнее отображения вида (6.55). Разумеется, условие полной положительности накладывает нетривиальные ограничения на параметры λ_γ, t_γ. Наиболее прозрачным является случай $t_\gamma \equiv 0$, когда отображение Φ, и, следовательно Λ, унитальны. В этом случае Λ является сжатием шара Блоха вдоль осей x, y, z с коэффициентами $|\lambda_x|, |\lambda_y|, |\lambda_z|$, которое комбинируется с отражениями, если некоторые из чисел λ_γ отрицательны (например, $\lambda_y < 0$ влечет отражение относительно плоскости xz). Транспонированию матрицы плотности $S(\boldsymbol{a})$, которое заведомо не является вполне положительным отображением, отвечает Λ с параметрами $\lambda_x = 1, \lambda_y = -1, \lambda_z = 1$. Отсюда, в частности, получаем, что соотношение типа (6.58) имеет место и для ортогональных матриц O с $\det O = -1$, однако им соответствуют антиунитарные преобразования U. Поскольку всякое взаимно-однозначное аффинное преобразование шара Блоха, очевидно, задается ортогональной матрицей O, отсюда вытекает теорема Вигнера-Кэдисона 6.1.1 для q-бита.

Задача 6.8.2 *Используя правила умножения (2.17), покажите, что в случае* $t_\gamma \equiv 0$

$$\Lambda[S] = \sum_{\gamma=0,x,y,z} \mu_\gamma \sigma_\gamma S \sigma_\gamma, \tag{6.59}$$

где

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

Неотрицательность этих чисел необходима и достаточна для полной положительности отображения Λ*, а значит, и* Φ*.*

Задача 6.8.3 *Покажите, что q-битные унитальные каналы (6.59) ковариантны по отношению к проективному унитарному представлению группы* $\mathbb{Z}_2 \times \mathbb{Z}_2$*, определяемому соотношениями (6.50).*

6.9 Комментарии

1. Доказательство теоремы 6.1.1, восходящей к утверждению Вигнера [4], следует книге Дэвиса [75] (см. раздел 2.3 и комментарии к нему).

2. Понятие вполне положительного отображения введено Стайнспрингом [149] в более общем контексте C^*-алгебр. Математический обзор свойств вполне положительных отображений, включая доказательство неравенства в задаче 6.2.1, имеется в статье Штермера (с. 85-106 сборника [84]). В конечномерном случае, имеющем свою специфику, вполне положительные отображения подробно рассматривались в работе Чоя [71].

В основе приведенного доказательства теоремы 6.2.1 лежит конструкция Гельфанда-Наймарка-Сигала: построение гильбертова пространства и представления в этом пространстве, опирающееся на объект с подходящими свойствами положительности. Такое доказательство с небольшими изменениями проходит для вполне положительных отображений C^*-алгебр. В случае алгебры $\mathfrak{B}(\mathcal{H})$ лемма 6.2.2 позволяет получить более подробное представление (6.7). Представление Крауса подробно обсуждается в книге [117].

3. На свойство полной положительности динамики открытой квантовой системы было обращено внимание в работах Холево [31], Линдблада [125]. Большое число работ посвящено изучению марковской динамики открытой квантовой системы, которая описывается полугруппой вполне положительных отображений. Обзор теории квантовых динамических полугрупп и квантовых случайных процессов см. в книге [36].

Соответствие Чоя-Ямилковского вытекает из результатов работы [71]. В квантовой теории информации оно используется весьма плодотворно;

например, представление Крауса (6.17) в конечномерном случае легко получается из (6.26), см. [22], теорема 8.1.

4. Каналы типа q-c-q были введены в работе Холево [35]. Подробное рассмотрение каналов, разрушающих сцепленность, включая доказательство предложения 6.4.1, дано в работе Городецкого, Шора и Рускаи [108].

5. Детальное обсуждение математических моделей квантового измерения см., например, в книгах Людвига [128], Дэвиса [75], Крауса [117], Холево [36], работах Линдблада [123], Озава [131].

6. Понятие комплементарного канала введено в работе Деветака и Шора [79]. Результаты настоящего раздела принадлежат Холево [40], а также Кингу, Мацумото, Натансону и Рускаи [115].

7. Дискретные операторы Вейля использовались в знаменитой статье Беннета, Брассара, Крепо, Джоза, Переса и Вуттерса [58] о телепортации. Относительно ковариантности комплементарного канала (задача 6.7.2) см. работу Холево [40].

8. Реальные примеры q-битных каналов подробно рассматриваются в книге Нильсена и Чанга [22]. Общая структура q-битных каналов подробно описана в работе Рускаи, Шарека и Вернер [134].

7. Квантовая энтропия и информация

7.1 Квантовая относительная энтропия

Определение 7.1.1 *Пусть S, T – операторы плотности. Если оператор T невырожден, то квантовая относительная энтропия определяется формулой*

$$H(S;T) = -H(S) - \operatorname{Tr} S \log T = \operatorname{Tr} S(\log S - \log T). \tag{7.1}$$

Случай, когда $\operatorname{supp} S \subseteq \operatorname{supp} T$ *(*$\operatorname{supp} S$ *обозначает носитель S, см. раздел 1.3), сводится к предыдущему путем рассмотрения сужения операторов S, T на* $\operatorname{supp} T$*, где T невырожден. Если же* $\operatorname{supp} S \not\subseteq \operatorname{supp} T$*, то полагаем $H(S;T) = +\infty$.*

Задача 7.1.1 *Пусть λ_j, μ_k – собственные значения, а $|e_j\rangle, |h_k\rangle$ – собственные векторы, соответственно, операторов плотности S, T, тогда*

$$H(S;T) = \sum_{j,k} |\langle e_j | h_k \rangle|^2 \, (\lambda_j \log \lambda_j - \lambda_j \log \mu_k). \tag{7.2}$$

Предложение 7.1.1

$$H(S;T) \geq \frac{\log e}{2} \operatorname{Tr}(S-T)^2 \geq 0, \tag{7.3}$$

причем равенство имеет место тогда и только тогда, когда $S = T$.

Доказательство. Используя для функции (4.2) формулу

$$\eta(\lambda) - \eta(\mu) = (\lambda - \mu)\eta'(\mu) + \frac{1}{2}(\lambda - \mu)^2 \eta''(\xi),$$

где $0 \leq \lambda < \xi < \mu$, получаем неравенство

$$\lambda \log \lambda - \lambda \log \mu \geq (\lambda - \mu) \log \mathrm{e} + \frac{\log \mathrm{e}}{2}(\mu - \lambda)^2;$$

подстановка этого неравенства в (7.2) приводит к требуемому результату. □

Следствие 7.1.1 (Субаддитивность квантовой энтропии) *Пусть S_{12} – оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2$ с частичными следами $S_1 = \mathrm{Tr}_{\mathcal{H}_2} S_{12}$ и $S_2 = \mathrm{Tr}_{\mathcal{H}_1} S_{12}$. Тогда*

$$H(S_1) + H(S_2) \leq H(S_{12}), \tag{7.4}$$

причем равенство имеет место тогда и только тогда, когда $S_{12} = S_1 \otimes S_2$.

Доказательство. Доказательство вытекает из равенства $H(S_1)+H(S_2)-H(S_{12}) = H(S_{12}; S_1 \otimes S_2)$ и неравенства (7.3).

Подобно энтропии, относительная энтропия обладает следующими свойствами:

Задача 7.1.2 *i. Изометрическая инвариантность: $H(VSV^*; VS'V^*) = H(S; S')$ для произвольного оператора V, изометричного на носителях операторов S, S';*
ii. Аддитивность: $H(S_1 \otimes S_2; S_1' \otimes S_2') = H(S_1; S_1') + H(S_2; S_2')$.

7.2 Монотонность относительной энтропии

Теорема 7.2.1 *Для произвольных операторов плотности S, T в $\mathcal{H}_1$ и произвольного канала Φ из $\mathcal{H}_1$ в $\mathcal{H}_2$*

$$H(S; T) \geq H(\Phi[S]; \Phi[T]). \tag{7.5}$$

Этот важный результат вытекает из доказываемой ниже теоремы 7.2.2, которая устанавливает аналогичное свойство монотонности для целого класса функций пары квантовых состояний.

Рассмотрим гильбертово пространство $L^2(\mathcal{H})$ операторов в $\mathcal{H}$ со скалярным произведением

$$(X, Y) = \mathrm{Tr}\, X^* Y.$$

Пусть S, T – невырожденные операторы. В $L^2(\mathcal{H})$ введем оператор L_T левого умножения на T и оператор R_S правого умножения на S, а именно, $L_T X = TX, R_S X = XS$. Эти операторы являются коммутирующими эрмитовыми положительными операторами в $L^2(\mathcal{H})$.

Относительная g-энтропия операторов S, T определяется соотношением

$$H_g(S; T) = (S^{1/2}, g(L_T R_S^{-1}) S^{1/2}), \tag{7.6}$$

для любой функции $g \in \mathcal{G}$, где $\mathcal{G}$ – класс функций вида

$$g(w) = a(w-1) + b(w-1)^2 + \int_0^\infty \frac{(w-1)^2}{w+s} d\nu(s), \tag{7.7}$$

где a – действительное число, $b \geq 0$ и ν – положительная мера на $[0, \infty)$, такая, что $\int_0^\infty \frac{1}{s+1} d\nu(s) < \infty$ (относительно внутренней характеризации этого класса см. раздел 7.7). Для доказательства следующей теоремы важно лишь то обстоятельство, что функция $g(w)$, входящая в определение относительной g-энтропии, имеет такое представление.

Теорема 7.2.2 *Для произвольной $g \in \mathcal{G}$ относительная g-энтропия $H_g(S;T)$ обладает свойством монотонности*

$$H_g(S;T) \leq H_g(\Phi[S];\Phi[T]), \tag{7.8}$$

для любого канала Φ.

Теорема 7.2.1 тогда вытекает из двух наблюдений:

i. функция $g(w) = -\log w$ принадлежит классу $\mathcal{G}$. В самом деле, легко проверяется, что

$$-\ln w = \int_0^\infty \left[\frac{1}{w+s} - \frac{1}{1+s}\right] ds = -(w-1) + \int_0^\infty \frac{(w-1)^2}{(w+s)(s+1)^2} ds.$$

ii.
$$H_{-\log}(S;T) = H(S;T).$$

Действительно,

$$-\log(L_T R_S^{-1})X = \log R_S X - \log L_T X = X(\log S) - (\log T)X$$

в силу коммутирования операторов L_T, R_S, откуда $-\log(L_T R_S^{-1})S^{1/2} = (\log S - \log T)S^{1/2}$ и результат вытекает из формулы (7.6).

Доказательство теоремы 7.2.2.

Лемма 7.2.3 *Для произвольной функции $g \in \mathcal{G}$,*

$$\begin{aligned} H_g(S;T) = {} & b \operatorname{Tr}(T-S)S^{-1}(T-S) \\ & + \int_0^\infty \operatorname{Tr}(T-S)(L_T + sR_S)^{-1}(T-S) d\nu(s) \end{aligned} \tag{7.9}$$

Доказательство. Вводя обозначение $\Delta_{T,S} = L_T R_S^{-1}$, заметим, что

$$(\Delta_{T,S} - I)S^{1/2} = (T-S)S^{-1/2} = R_{S^{-1/2}}(T-S), \tag{7.10}$$

так что

$$H_{w-1}(S;T) = \operatorname{Tr} S^{1/2}(T-S)S^{-1/2} = 0, \tag{7.11}$$

и, следовательно, линейный член в (7.7) не дает вклада в (7.9). Для $g(w) = \frac{(w-1)^2}{(w+s)}$ получаем, используя (7.10),

$$\begin{aligned} H_g(S;T) &= \Big((\Delta_{T,S}-I)S^{1/2}, (\Delta_{T,S}+sI)^{-1}(\Delta_{T,S}-I)S^{1/2}\Big) \\ &= \operatorname{Tr}\left[(T-S)(\Delta_{T,S}+sI)^{-1}R_{S^{-1}}(T-S)\right] \\ &= \operatorname{Tr}(T-S)(L_T+sR_S)^{-1}(T-S). \end{aligned} \tag{7.12}$$

При $s=0$ имеем

$$H_{(w-1)^2/w}(S;T) = \operatorname{Tr}(T-S)T^{-1}(T-S) = H_{(w-1)^2}(T;S). \tag{7.13}$$

Подставляя эти соотношения в (7.7), получаем (7.9). □

Если не учитывать первый член в (7.7) (который не дает вклада в $H_g(S;T)$), то функции (7.7) являются "континуальными выпуклыми комбинациями" функций $(w-1)^2$ и $\frac{(w-1)^2}{w+s}$, $s\geq 0$; таким образом, достаточно доказать монотонность относительной g-энтропии для таких функций g, т. е. для (7.13).

Доказательство теоремы 7.2.2 теперь вытекает из интегрального представления (7.9) и следующей леммы

Лемма 7.2.4 *Для произвольного канала Φ, $s\geq 0$ и оператора A*

$$\operatorname{Tr} A^*(L_T+sR_S)^{-1}A \geq \operatorname{Tr}\Phi[A^*](L_{\Phi[T]}+sR_{\Phi[S]})^{-1}\Phi[A].$$

Доказательство. Так как операторы L_T, R_S являются положительными в $L^2(\mathcal{H})$, то оператор L_T+sR_S также положителен. Положим $X=(L_T+sR_S)^{-1/2}A-(L_T+sR_S)^{1/2}\Phi^*[B]$, где $B=(L_{\Phi[T]}+sR_{\Phi[S]})^{-1}\Phi[A]$. Тогда $\operatorname{Tr}X^*X\geq 0$, так что

$$\begin{aligned} \operatorname{Tr}A^*(L_T+sR_S)^{-1}A - \operatorname{Tr}A^*\Phi^*[B] - \operatorname{Tr}\Phi^*[B^*]A \\ + \operatorname{Tr}\Phi^*[B^*](L_T+sR_S)\Phi^*[B]\geq 0. \end{aligned} \tag{7.14}$$

Имеем

$$\operatorname{Tr}A^*\Phi^*[B]+\operatorname{Tr}\Phi^*[B^*]A = 2\operatorname{Tr}\Phi[A^*](L_{\Phi[T]}+sR_{\Phi[S]})^{-1}\Phi[A],$$

таким образом для доказательства леммы достаточно показать, что последний член в (7.14) меньше или равен, чем правая часть в (7.14). Имеем

$$\begin{aligned} \operatorname{Tr}\Phi^*[B^*](L_T+sR_S)\Phi^*[B] &= \operatorname{Tr}[\Phi^*[B^*]\Phi^*[B]S+\Phi^*[B^*]sT\Phi^*[B]] \\ &= \operatorname{Tr}[\Phi^*[B^*]\Phi^*[B]S+\Phi^*[B]\Phi^*[B^*]sT] \\ &\leq \operatorname{Tr}[\Phi^*[B^*B]S+\Phi^*[BB^*]sT], \end{aligned}$$

где неравенство вытекает из положительности S,T и неравенства Кэдисона (6.6), которое в случае канала Φ принимает вид

$$\Phi^*[B^*]\Phi^*[B]\leq\Phi^*[B^*B].$$

Далее, используя соотношение $\operatorname{Tr}\Phi^*[B^*B]S = \operatorname{Tr} B^*B\Phi[S]$, имеем

$$\begin{aligned}\operatorname{Tr}\Phi^*[B^*](L_T + sR_S)\Phi^*[B] &\le \operatorname{Tr}(B^*B\Phi[S] + BB^*s\Phi[T]) \\ &= \operatorname{Tr} B^*(B\Phi[S] + s\Phi[T]B) \\ &= \operatorname{Tr} B^*(L_{\Phi[T]} + sR_{\Phi[S]})(B) = \operatorname{Tr} B^*\Phi[A] \\ &= \operatorname{Tr}\Phi[A^*](L_{\Phi[T]} + sR_{\Phi[S]})^{-1}\Phi[A].\end{aligned}$$

□

Это доказывает теорему 7.2.2 в случае невырожденных S, T. В общем случае либо $H(S;T) = +\infty$ и неравенство (7.5) тривиально, либо $\operatorname{supp} S \subseteq \operatorname{supp} T$, и тогда можно считать, что $\operatorname{supp} T = \mathcal{H}$, т. е. T невырожден, а для $H(S;T)$ имеет место соотношение (7.1). Аппроксимируя произвольный оператор S невырожденными операторами и используя непрерывность энтропии (см. раздел 7.3), убеждаемся в справедливости теоремы в общем случае. □

Следствие 7.2.1 (обобщенная H-теорема) *Пусть Φ – бистохастический (унитальный) канал в $\mathcal{H} = \mathcal{H}_1 = \mathcal{H}_2$. Тогда*

$$H(\Phi[S]) \ge H(S). \tag{7.15}$$

Доказательство. Из (7.1) следует, что

$$H(S) = \log d - H(S;\overline{S}), \tag{7.16}$$

где $\overline{S} = \frac{1}{d}I$ – хаотическое состояние, а значит, и

$$H(\Phi[S]) = \log d - H(\Phi[S];\overline{S}),$$

в силу того, что $\Phi[\overline{S}] = \overline{S}$. Теперь утверждение непосредственно вытекает из теоремы о монотонности 7.2.1. □

Покажем, как из теоремы 7.2.1 можно получить верхнюю границу для шенноновской информации (5.15). Вводя среднее состояние $\bar{S}_\pi = \sum \pi_x S_x$, имеем важное соотношение

$$\chi(\pi) \equiv H(\bar{S}_\pi) - \sum \pi_x H(S_x) = \sum_x \pi_x H(S_x;\bar{S}_\pi) \tag{7.17}$$

Выберем произвольный базис $\{e_y\}$ в $\mathcal{H}$ и рассмотрим q-с канал $\Psi[S] = \sum_y \operatorname{Tr} SM_y|e_y\rangle\langle e_y|$, отвечающий измерению наблюдаемой $M = \{M_y\}$, тогда

$$\Psi[S_x] = \sum_y p_M(y|\,x)|e_y\rangle\langle e_y|; \quad \Psi[\bar{S}_\pi] = \sum_y \left(\sum_x \pi_x p_M(y|\,x)\right)|e_y\rangle\langle e_y|,$$

где $p_M(y|x) = \operatorname{Tr} S_x M_y$, – операторы плотности, диагональные в базисе $\{e_y\}$. Поэтому квантовая относительная энтропия на выходе канала Ψ превращается в классическую, и соотношение (7.17) влечет

$$\sum_x \pi_x H(\Psi[S_x]; \Psi[\bar{S}_\pi]) = \mathcal{I}_1(\pi, M).$$

Применяя теорему 7.2.1 о монотонности относительной энтропии к каналу Ψ и еще раз используя соотношение (7.17), получаем искомое неравенство.

Теорема 7.2.5 *Следующие свойства эквивалентны:*

i. монотонность относительной энтропии;

ii. сильная субаддитивность квантовой энтропии: пусть S_{123} – оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, тогда, используя очевидные обозначения для частичных следов,

$$H(S_{123}) + H(S_2) \leq H(S_{12}) + H(S_{23}); \tag{7.18}$$

iii. совместная выпуклость относительной квантовой энтропии $H(S;T)$ по аргументам S, T.

Доказательство. i. $\Rightarrow$ ii. Обозначая через $\bar{S}_\alpha = (\dim \mathcal{H}_\alpha)^{-1} I_\alpha$ хаотическое состояние в $\mathcal{H}_\alpha$, имеем

$$H(S_{123}; \bar{S}_{123}) = H(S_{123}; S_{12} \otimes \bar{S}_3) + H(S_{12}; \bar{S}_{12}),$$

$$H(S_{23}; \bar{S}_{23}) = H(S_{23}; S_2 \otimes \bar{S}_3) + H(S_2; \bar{S}_2).$$

Вычитая одно равенство из другого и используя (7.16), получаем

$$H(S_{12}) + H(S_{23}) - H(S_{123}) - H(S_2) = H(S_{123}; S_{12} \otimes \bar{S}_3) - H(S_{23}; S_2 \otimes \bar{S}_3).$$

Рассмотрим канал $\Phi[S_{123}] = S_{23}$. Применяя к нему теорему 7.2.1, получаем, что правая часть неотрицательна, т. е. убеждаемся в справедливости свойства сильной субаддитивности.

ii. $\Rightarrow$ iii. Рассмотрим оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$ вида

$$S_{123} = \sum_{kl} |e_k\rangle\langle e_k| \otimes A_{kl} \otimes |h_l\rangle\langle h_l|,$$

где $\{e_k\}$ – ортонормированный базис в $\mathcal{H}_1$, $\{h_l\}$ ортонормированный базис в $\mathcal{H}_3$, A_{kl} – положительные операторы в $\mathcal{H}_2$. Тогда

$$S_{12} = \sum_k |e_k\rangle\langle e_k| \otimes \sum_l A_{kl}, \quad S_{23} = \sum_k A_{kl} \otimes |h_l\rangle\langle h_l|, \quad S_2 = \sum_{kl} A_{kl},$$

и свойство сильной субаддитивности влечет следующее неравенство

$$-\sum_{kl} H(A_{kl}) + \sum_k H\left(\sum_l A_{kl}\right) + \sum_l H\left(\sum_k A_{kl}\right) - H\left(\sum_{kl} A_{kl}\right) \geq 0,$$

где введено обозначение $H(A) = -\mathrm{Tr} A \log A$ для любого положительного оператора A. Тогда получаем, в частности, для любых положительных операторов A_1, A_2, B_1, B_2,

$$H(A_1 + A_2 + B_1 + B_2) - H(A_1 + A_2) - H(B_1 + B_2) \leq$$

$$\leq H(A_1 + B_1) - H(A_1) - H(B_1) + H(A_2 + B_2) - H(A_2) - H(B_2),$$

что равносильно совместной выпуклости функции

$$\Delta(A, B) = H(A + B) - H(A) - H(B)$$

по аргументам A, B. Из следующей леммы вытекает, что относительная энтропия $H(A; B)$ также является выпуклой как супремум семейства выпуклых функций.

Лемма 7.2.6 *Для любых двух операторов плотности A, B*

$$H(A; B) = \sup_{\lambda > 0} \lambda^{-1} \chi_\lambda(A; B), \tag{7.19}$$

где

$$\chi_\lambda(A; B) = H(\lambda A + (1-\lambda)B) - \lambda H(A) - (1-\lambda)H(B) = \Delta(\lambda A, (1-\lambda)B) + h(\lambda).$$

Доказательство. Из доказанного выше следует, что функция $\chi_\lambda(A; B)$ выпукла по A, B. Как функция $\lambda \in [0, 1]$ она вогнута и равна нулю при $\lambda = 0, 1$ (это было доказано в разделе 5.3). Значит, при фиксированных A, B

$$\sup_{\lambda > 0} \lambda^{-1} \chi_\lambda(A; B) = \lim_{\lambda \to 0} \lambda^{-1} \chi_\lambda(A; B) = \frac{d}{d\lambda} |_{\lambda=0} \chi_\lambda(A; B) \ .$$

Задача 7.2.1 *Пусть F – непрерывно дифференцируемая функция в интервале $(0, +\infty)$. Покажите, что для любых положительных операторов A, B выполняется*

$$\frac{d}{d\lambda} \mathrm{Tr}\, F(\lambda A + (1-\lambda)B) = \mathrm{Tr}\, F'(\lambda A + (1-\lambda)B)(A - B), \tag{7.20}$$

Указание: докажите (7.20) в случае, когда F – многочлен, а затем используйте аппроксимацию.

Пусть теперь supp $A \subset$ supp B, так что $H(A;B) < +\infty$. Без ограничения общности будем считать, что оператор B невырожден. Вычисляя производную от

$$H(\lambda A + (1-\lambda)B) = \operatorname{Tr} \eta(\lambda A + (1-\lambda)B),$$

по формуле (7.20), получаем, что

$$\frac{d}{d\lambda} |_{\lambda=0} \chi_\lambda(A;B) = \operatorname{Tr} A(\log A - \log B) + \log \mathrm{e} \operatorname{Tr}(B-A) = H(A;B),$$

откуда вытекает (7.19).

Задача 7.2.2 *Проведите доказательство в случае* $H(A;B) = +\infty$.

□

iii. $\Rightarrow i$. Рассмотрим оператор плотности S_{12} в $\mathcal{H}_1 \otimes \mathcal{H}_2$ и положим $S_1 = \operatorname{Tr}_2 S_{12}$. Используя свойство (6.49) дискретных операторов Вейля, имеем

$$S_1 \otimes \bar{S}_2 = \frac{1}{d^2} \sum_{\alpha,\beta=0}^{d-1} (I_1 \times W_{\alpha\beta}) S_{12} (I_1 \times W_{\alpha\beta})^*.$$

Используя выпуклость относительной энтропии и свойства из упражнения 7.1.2, получаем неравенство

$$H(S_1; S_1') \le H(S_{12}; S_{12}'),$$

которое является свойством монотонности для частичных следов. Но по теореме 6.2.1, всякий канал может быть представлен в виде суперпозиции изометрического вложения (сохраняющего относительную энтропию) и частичного следа. Таким образом, свойство i. доказано. □

Используя (7.16) и выпуклость относительной энтропии, мы попутно получаем

Следствие 7.2.2 *Квантовая энтропия вогнута: для произвольных состояний* S_j *и распределения вероятностей* $\{p_j\}$,

$$H\Big(\sum_j p_j S_j\Big) \ge \sum_j p_j H(S_j).$$

Задача 7.2.3 *Докажите вогнутость непосредственно, используя спектральное разложение оператора* $\sum_j p_j S_j$ *и вогнутость функции* $\eta(t)$.

7.3 Свойства непрерывности

В рассматриваемом здесь конечномерном случае квантовая энтропия является непрерывной функцией от состояния. В самом деле, функция $\eta(x) = -x\log x$ (равномерно) непрерывна на отрезке $[0,1]$, значит, из $S_n \to S$ вытекает $\|\eta(S_n) - \eta(S)\| \to 0$, где $\|\cdot\|$ – операторная норма. Следовательно,

$$|H(S_n) - H(S)| = |\operatorname{Tr}(\eta(S_n) - \eta(S))| \leq d\,\|\eta(S_n) - \eta(S)\| \to 0.$$

Следующая лемма дает более точную оценку, которая понадобится в дальнейшем

Лемма 7.3.1 *Пусть S_1, S_2 – два оператора плотности в $d-$мерном гильбертовом пространстве, причем $\|S_1 - S_2\|_1 \leq \frac{1}{e}$. Тогда*

$$|H(S_1) - H(S_2)| \leq \log d \cdot \|S_1 - S_2\|_1 + \eta\left(\|S_1 - S_2\|_1\right). \tag{7.21}$$

Так как функция $\eta(x)$ монотонно возрастает при $x \in \left[0, \frac{1}{e}\right]$, то из леммы следует оценка

$$|H(S_1) - H(S_2)| \leq \log d \cdot \|S_1 - S_2\|_1 + \frac{\log e}{e}. \tag{7.22}$$

Доказательство. Пусть $\lambda_1 \geq \lambda_2 \geq \ldots (\mu_1 \geq \mu_2 \geq \ldots)$ – собственные значения оператора S_1 (соответственно, оператора S_2).

Задача 7.3.1 *Обозначая $\Delta_i = |\lambda_i - \mu_i|$, покажите, что*

$$\Delta = \sum_{i=1}^{d} \Delta_i \leq \|S_1 - S_2\|_1. \tag{7.23}$$

Задача 7.3.2 *Докажите неравенство*

$$|\eta(y) - \eta(x)| \leq \eta(y - x), \quad 0 \leq x \leq y \leq 1.$$

Комбинируя результаты этих задач, получаем

$$\begin{aligned} |H(S_1) - H(S_2)| &\leq \sum_{i=1}^{d} |\eta(\lambda_i) - \eta(\mu_i)| \leq \sum_{i=1}^{d} |\eta(\Delta_i)| \\ &= \Delta \sum_{i=1}^{d} |\eta(\Delta_i/\Delta)| + \eta(\Delta) \leq \Delta \log d + \eta(\Delta). \end{aligned}$$

Принимая во внимание монотонность $\eta(x)$ и используя неравенство (7.23), получаем (7.21). □

Задача 7.3.3 *Покажите, что относительная энтропия полунепрерывна снизу в следующем смысле: если $S_n \to S$ и $S'_n \to S'$, то*

$$\lim \inf_{n\to\infty} H(S_n; S'_n) \geq H(S; S').$$

Указание: используйте непрерывность энтропии, представление (7.19) и тот факт, что супремум семейства непрерывных функций полунепрерывен снизу.

Как мы позже увидим в главе 10, в бесконечномерном гильбертовом пространстве свойство полунепрерывности снизу выполнено как для энтропии, так и для относительной энтропии, тогда как непрерывность энтропии уже не имеет места.

7.4 Информационная корреляция, сцепленность формирования и условная энтропия

В классическом случае мерой количества информации, содержащейся в одной из случайных величин X, Y относительно другой, является взаимная информация $I(X;Y) = H(X)+H(Y)-H(XY)$. В квантовой статистике нет прямого аналога совместной энтропии $H(XY)$, так как совместное распределение наблюдаемых существует лишь в специальных случаях. Рассмотрим один из таких случаев. Пусть S_{12}– состояние составной квантовой системы в гильбертовом пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2$. В этой ситуации очевидным аналогом шенноновской взаимной информации является

$$I(1;2) = H(S_1) + H(S_2) - H(S_{12}) = H(S_{12}; S_1 \otimes S_2). \tag{7.24}$$

Будем называть эту величину *информационной корреляцией*. Из предложения 7.1.1 вытекает, что $I(1;2) \geq 0$ и $I(1;2) = 0$ тогда и только тогда, когда $S_{12} = S_1 \otimes S_2$.

Рассмотрим составную систему и наблюдаемые M_1, M_2 в соответствующих подсистемах. Их совместное распределение существует и дается формулой $p(x,y) = \operatorname{Tr} S_{12} M_{1x} M_{2y}$. Обозначим через $I(M_1; M_2)$ шенноновскую взаимную информацию между исходами измерений наблюдаемых M_1, M_2.

Задача 7.4.1 *Покажите, что для чистого состояния $S_{12} = |\psi\rangle\langle\psi|$*

$$I(1;2) = 2 \max_{M_1, M_2} I(M_1, M_2). \tag{7.25}$$

В случае чистого состояния S_{12} имеем $H(S_{12}) = 0$, и используя теорему 3.1.2, получаем

$$I(1;2) = H(S_1) + H(S_2) = 2H(S_1). \tag{7.26}$$

Энтропия $H(S_1) = H(S_2)$ является естественной мерой сцепленности чистого состояния S_{12}. Она равна нулю для несцепленных состояний и принимает максимальное значение $\log d$ для состояния (3.7), что объясняет название "максимально сцепленное" для этого состояния.

Вопрос определения меры сцепленности для смешанных состояний является значительно более сложным. Наибольший интерес для нас представляет *сцепленность формирования*, которая определяется для произвольного состояния S_{12} в $\mathcal{H}_1 \otimes \mathcal{H}_2$ как

$$E_F(S_{12}) = \inf \sum_j \pi_j H(S_1^j), \tag{7.27}$$

где инфимум берется по всем возможным разложениям

$$S_{12} = \sum_j \pi_j S_{12}^j$$

состояния S_{12} в выпуклую комбинацию состояний S_{12}^j; в терминах выпуклого анализа, $E_F(S_{12})$ является *выпуклой оболочкой* непрерывной функции $S_{12} \to H(S_1)$.

Предложение 7.4.1 *Функция $S_{12} \to E_F(S_{12})$ является непрерывной и выпуклой, а инфимум в (7.27) достигается на выпуклой комбинации не более, чем $(d_1 d_2)^2$ чистых состояний S_{12}^j. Имеет место соотношение двойственности:*

$$E_F(S_{12}) = \max_{A_{12}} \operatorname{Tr} S_{12} A_{12}, \tag{7.28}$$

где максимум берется по всем эрмитовым операторам A_{12} из $\mathcal{H}_1 \otimes \mathcal{H}_2$, удовлетворяющим условию

$$\operatorname{Tr} T_{12} A_{12} \le H(T_1), \tag{7.29}$$

для всех (чистых) состояний T_{12} в $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Доказательство. Обозначим $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, $S = S_{12}$ и $f(S) = H(\operatorname{Tr}_{\mathcal{H}_2} S)$; функция f является непрерывной вогнутой функцией на компактном выпуклом множестве квантовых состояний $\mathfrak{S} = \mathfrak{S}(\mathcal{H})$. Доказательство, приведенное ниже, практически не использует других специальных свойств f и $\mathfrak{S}$, и может быть получено из общих фактов выпуклого анализа, но мы дадим здесь самостоятельный вывод.

Конечный набор состояний с соответствующими вероятностями обычно называют *ансамблем*. Для доказательства теоремы нам понадобится "континуальное" обобщение этого понятия, описываемое вероятностной мерой $\pi(dS)$ на множестве состояний $\mathfrak{S}(\mathcal{H})$ (концепция таких обобщенных ансамблей подробно обсуждается в главе 10). При таком подходе обычные ансамбли описываются мерами с конечным носителем. Используя общие факты теории меры (см. комментарии), можно показать, что

множество $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$ всех вероятностных мер на $\mathfrak{S}(\mathcal{H})$ является компактным в топологии слабой сходимости, которая определяется как сходимость интегралов от всевозможных ограниченных непрерывных функций на $\mathfrak{S}(\mathcal{H})$.

Рассмотрим функционал

$$F(\pi) = \int_{\mathfrak{S}} f(T)\pi(dT), \tag{7.30}$$

на $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$, который для мер π с конечным носителем совпадает с минимизируемым выражением в (7.27). Согласно определению, этот функционал является непрерывным аффинным функционалом на $\mathcal{P}(\mathfrak{S}(\mathcal{H}))$. Нас будет интересовать его минимум на замкнутом выпуклом подмножестве $\mathcal{P}_S$ вероятностных мер π с фиксированным барицентром

$$S = \int_{\mathfrak{S}(\mathcal{H})} T\pi(dT).$$

Непрерывный функционал $F(\pi)$ достигает минимума на выпуклом множестве $\mathcal{P}_S$. Из вогнутости f следует, что можно выбрать минимизирующую меру π с носителем на чистых состояниях, так как всегда можно произвести спектральные разложения для всех операторов плотности $T = T_{12}$ на чистые состояния без изменения барицентра и увеличения значения $F(\pi)$.

Покажем, что всякая крайняя точка $\pi \in \mathcal{P}_S$ имеет носитель, состоящий не более чем из $(d_1 d_2)^2$ состояний. Заметим, что состояния из носителя π линейно независимы в том смысле, что если

$$\int_{\mathfrak{S}(\mathcal{H})} c(T)T\pi(dT) = 0 \tag{7.31}$$

для некоторой вещественной ограниченной измеримой функции $c(T)$, то $c(T) = 0$ почти всюду по мере π. Действительно, из (7.31) следует, что $\pi = \frac{1}{2}\pi_+ + \frac{1}{2}\pi_-$, где

$$\pi_\pm(dT) = (1 \pm \varepsilon c(T))\pi(dT)$$

для достаточно малого ε; из того, что π крайняя точка, следует $\pi_\pm(dT) = \pi(dT)$, откуда $c(T) = 0(\mathrm{mod}\pi)$. Рассматривая матричные элементы (7.31), получаем, что вещественнное гильбертово пространство $L^2_{\mathbb{R}}(\pi)$ натянуто на $(d_1 d_2)^2$ функций

$$T \to \Re\langle e_j|T|e_k\rangle;\ 1 \le j \le k \le d_1 d_2,$$

$$T \to \Im\langle e_j|T|e_k\rangle;\ 1 \le j < k \le d_1 d_2.$$

Действительно, любая функция $c(T)$, ортогональная к этим функциям в $L^2_{\mathbb{R}}(\pi)$, должна равняться нулю $(\mathrm{mod}\pi)$. Оставшаяся часть доказательства предоставляется в качестве задачи:

Задача 7.4.2 *Покажите, что если* $\dim L^2(\pi) = n$, *то носитель* π *состоит из* n *точек.*

Неравенство $\geq$ в соотношении двойственности (7.28) получается путем интегрирования неравенства $f(T) \geq \operatorname{Tr} TA$ по любой вероятностной мере $\pi \in \mathcal{P}_S$. Для доказательства обратного неравенства рассмотрим только меры с фиксированным (но произвольным) конечным носителем, содержащим носитель оптимальной (минимизирующей) меры π^0. Тогда минимизация становится задачей в конечномерном пространстве, с конечным числом линейных ограничений (7.29). Заметим, что эти ограничения включают в себя нормировку вероятностной меры. Применяя метод Лагранжа, получаем, что найдется эрмитов оператор Λ, такой, что π^0 минимизирует функционал

$$F(\pi) - \operatorname{Tr}\left(\int_{\mathfrak{S}} T\pi(dT)\right)\Lambda = \int_{\mathfrak{S}} [f(T) - \operatorname{Tr} T\Lambda]\pi(dT) \tag{7.32}$$

по всем (ненормированным) положительным мерам π с заданным конечным носителем. Здесь эрмитов оператор Λ описывает набор вещественных множителей Лагранжа для ограничений (7.29). Отсюда следует, что

$$f(T) - \operatorname{Tr} T\Lambda \geq 0; \quad f(T) - \operatorname{Tr} T\Lambda = 0(\mathrm{mod}\pi^0). \tag{7.33}$$

Интегрируя второе равенство по π^0 и учитывая (7.29), получаем равенство $\operatorname{Tr} S\Lambda = \int_{\mathfrak{S}} f(T)\pi^0(dS) = F(\pi^0)$, которое вместе с неравенством в (7.33) означает, что Λ является решением двойственной задачи и что соотношение (7.28) имеет место с общим значением $\operatorname{Tr} S\Lambda$.

Из соотношения двойственности вытекает, что функция $E_F(S)$ полунепрерывна снизу, как максимум непрерывных (аффинных) функций $S \to \operatorname{Tr} SA$. Остается показать, что эта функция является и полунепрерывной сверху, т.е.

$$\limsup_{n\to\infty} E_F(S^{(n)}) \leq E_F(S^{(n)}),$$

для любой последовательности состояний $S^{(n)} \to S$. Пусть

$$E_F(S) = \sum_{j=1}^{N} \pi_j^0 H(\operatorname{Tr}_{\mathcal{H}_2} S_j), \tag{7.34}$$

где $S = \sum_{j=1}^{N} \pi_j^0 S_j$. Тогда, если оператор S невырожден, положим

$$S^{(n)} = \sum_{j=1}^{N} \pi_j^{(n)} S_j^{(n)},$$

где

$$\pi_j^{(n)} = t^{(n)}\pi_j^0,$$

$$S_j^{(n)} = (t^{(n)})^{-1}(S^{(n)})^{1/2}S^{-1/2}S_jS^{-1/2}(S^{(n)})^{1/2},$$

и

$$t^{(n)} = \mathrm{Tr}(S^{(n)})^{1/2}S^{-1/2}S_jS^{-1/2}(S^{(n)})^{1/2}.$$

Эти выражения имеют смысл и для вырожденного S, если под $S^{-1/2}$ понимается обобщенный обратный оператор к $S^{1/2}$, т. е. оператор, который равен обратному на носителе и нулю на ортогональном дополнении. Тогда

$$E_F(S^{(n)}) \leq \sum_{j=1}^{N} \pi_j^{(n)} H(\mathrm{Tr}_{\mathcal{H}_2} S_j^{(n)}),$$

и выражение в левой части сходится к (7.34). Поэтому

$$\limsup_{n\to\infty} E_F(S^{(n)}) \leq E_F(S^{(n)}),$$

то есть E_F является полунепрерывной сверху, и, следовательно, непрерывной. □

Вернемся к случаю чистых состояний S_{12}. Заметим, что в классической статистике частичные следы чистого состояния (маргинальные распределения вырожденного распределения) вновь являются чистыми, и операция очищение не имеет аналога. В этой связи отметим следующее необычное свойство квантового аналога условной энтропии. В классическом случае условная энтропия всегда неотрицательна

$$H(X|Y) = H(XY) - H(Y) = \sum_y p_y H(X|Y=y) \geq 0. \tag{7.35}$$

Определяя *квантовую условную энтропию* как

$$H(1|2) = H(S_{12}) - H(S_2),$$

получаем, что эта величина отрицательна, если S_{12} – чистое сцепленное состояние. Другими словами, в отличие от классической энтропии, квантовая энтропия не является монотонной по отношению к расширению системы: $H(S_1) \not\leq H(S_{12})$. Тем не менее, как и в классическом случае, имеют место следующие свойства

i. Монотонность условной энтропии: $H(1|23) \leq H(1|2)$;
ii. $H(1|2) + H(1|3) \geq 0$.

Задача 7.4.3 *Покажите, что эти свойства являются переформулировкой свойства строгой субаддитивности энтропии (7.18).*

Другое полезное свойство квантовой условной энтропии:

Следствие 7.4.1 *Условная энтропия $H(1|2)$ является вогнутой функцией состояния S_{12}.*

Доказательство. Имеем

$$\begin{aligned} H(1|2) &= H(S_{12}) - H(S_2) \\ &= \log d_1 - \operatorname{Tr} S_{12}(\log S_{12} + I_{12} \log d_1 - I_1 \otimes \log S_2) \\ &= \log d_1 - H(S_{12}; \overline{S}_1 \otimes S_2), \end{aligned}$$

и утверждение вытекает из совместной выпуклости относительной энтропии и аффинности отображения $S_{12} \to \overline{S}_1 \otimes S_2$. □

Благодаря этим свойствам, квантовая условная энтропия является полезным аналитическим инструментом в квантовой теории информации.

7.5 Обменная энтропия

Пусть Φ – канал с входным пространством $\mathcal{H}_A$ и выходным пространством $\mathcal{H}_B$, и $S = S_A$ – оператор плотности в $\mathcal{H}_A$, задающий входное состояние канала. По теореме 6.3.1, канал допускает представление Стайнспринга

$$\Phi[S_A] = \operatorname{Tr}_{\mathcal{H}_E} V S_A V^*, \tag{7.36}$$

где V – изометрический оператор из $\mathcal{H}_A$ в $\mathcal{H}_{BE} = \mathcal{H}_B \otimes \mathcal{H}_E$, а $\mathcal{H}_E$ можно рассматривать как пространство "окружения". Передача состояния S_A порождает состояние в пространстве окружения

$$S_E = \operatorname{Tr}_{\mathcal{H}_B} V S_A V^* = \tilde{\Phi}[S_A], \tag{7.37}$$

где $\tilde{\Phi}$ – комплементарный канал, см. раздел 6.6 . В дальнейшем мы упростим обозначения для энтропий, опуская обозначения состояний, т. е. будем писать $H(A)$ вместо $H(S_A)$ и т. п., как это уже делалось в предыдущей главе.

Определение 7.5.1 *Обменной энтропией называется величина $H(S_E) = H(\tilde{\Phi}[S_A])$, равная выходной энтропии окружения. Мы будем обозначать ее $H(S, \Phi)$ или просто $H(E)$.*

Чтобы показать, что величина $H(S, \Phi)$ не зависит от выбора представления для канала, введем эталонную систему $\mathcal{H}_R$. Состояние S_A может быть очищено до чистого состояния $|\psi_{AR}\rangle\langle\psi_{AR}| \in \mathcal{H}_A \otimes \mathcal{H}_R$. Имеем

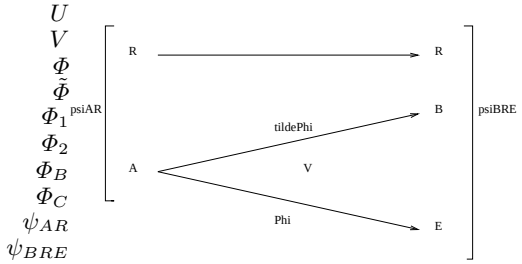


Рис. 7.1. *Очищение с эталонной системой.*

$$H\left((\Phi \otimes \mathrm{Id}_R)|\psi_{AR}\rangle\langle\psi_{AR}|\right) = H(S_{BR}). \tag{7.38}$$

Составная система BRE, описываемая гильбертовым пространством $\mathcal{H}_{BRE} = \mathcal{H}_B \otimes \mathcal{H}_R \otimes \mathcal{H}_E$, описывается чистым состоянием

$$|\psi_{BR}\rangle \otimes |\psi_E\rangle = (V \otimes I_R)|\psi_{AR}\rangle. \tag{7.39}$$

Рассматривая разделение $BR|E$, получаем совокупность двух систем в чистом состоянии, значит, согласно (7.26),

$$H(S_E) = H(S_{BR}), \tag{7.40}$$

где согласно (7.38), правая часть зависит только от канала Φ, и не зависит от выбора представления Стайнспринга. С другой стороны, (7.40) представляет собой выражение для выходной энтропии расширенного канала $\Phi \otimes \mathrm{Id}_R$, примененного к чистому состоянию S_{AR}. Так как левая часть не содержит R, то это равенство также показывает, что $H(S_{BR})$ не зависит от способа очищения состояния $S = S_A$.

Используя конструкцию комплементарного канала из (6.40), мы видим, что S_E задается матрицей плотности $\left[\mathrm{Tr}\, SV_j^* V_k\right]_{j,k=1,\ldots,N}$; следовательно, имеет место

Предложение 7.5.1 *Пусть* $\Phi[S] = \sum_{j=1}^N V_j S V_j^*$, *где* $\sum_{j=1}^N V_j^* V_j = I$ *тогда*

$$H(S, \Phi) = H\left(\left[\mathrm{Tr}\, SV_j^* V_k\right]_{j,k=1,\ldots,N}\right). \tag{7.41}$$

Обменная энтропия удовлетворяет следующему полезному неравенству:

Пусть $S = \sum_j p_j S_j$ – *смесь чистых состояний* S_j, *тогда*

$$H(S, \Phi) \geq \sum_j p_j H(\Phi[S_j]). \tag{7.42}$$

Действительно, благодаря выпуклости квантовой энтропии, $H(S, \Phi) = H(\Phi_E[\sum_j p_j S_j]) \geq \sum_j p_j H(\Phi_E[S_j])$. Но $H(\Phi_E[S_j]) = H(\Phi[S_j])$, так как система BE находится в чистом состоянии, если на входе – чистое состояние S_j. □

Пример Рассмотрим деполяризующий канал (6.51) и хаотическое входное состояние $\bar{S} = I_d/d$. Используя представление Крауса (6.52) и свойства операторов Вейля, находим

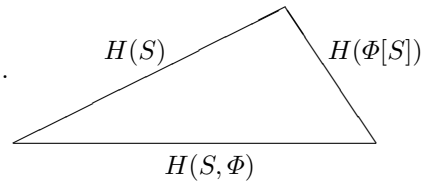
$$\operatorname{Tr} \bar{S} W^*_{\alpha\beta} W_{\alpha'\beta'} = \delta_{(\alpha\beta),(\alpha'\beta')} \begin{cases} \frac{p}{d^2}, & (\alpha\beta) \neq (00), \\ 1 - p\frac{d^2-1}{d^2}, & (\alpha\beta) = (00). \end{cases}$$

Следовательно,

$$H(\bar{S}, \Phi) = -\Big(1 - p\frac{d^2-1}{d^2}\Big) \log \Big(1 - p\frac{d^2-1}{d^2}\Big) - p\frac{d^2-1}{d^2} \log \frac{p}{d^2}. \qquad (7.43)$$

7.6 Квантовая взаимная информация

Выше были рассмотрены три энтропийные величины: входная энтропия $H(S)$, выходная энтропия $H(\Phi[S])$ и обменная энтропия $H(S, \Phi)$. Между ними имеются соотношения, позволяющие рассматривать их как длины сторон треугольника: сумма любых двух из них больше или равна третьей. Например,



$$\begin{aligned} I(S, \Phi) &= H(S) + H(\Phi[S]) - H(S, \Phi) \\ &= H(A) + H(B) - H(E) \\ &= H(R) + H(B) - H(BR) \geq 0 \end{aligned} \qquad (7.44)$$

по свойству субаддитивности квантовой энтропии.

Так как R имеет ту же энтропию, что и A, то величину $I(S, \Phi)$ можно назвать *квантовой взаимной информацией* между входом и выходом. Взаимная информация $I(S, \Phi)$ обладает целым рядом хороших свойств, аналогичных свойствам шенноновской информации.

Предложение 7.6.1 *Квантовая взаимная информация $I(S, \Phi)$*

i. вогнута по S;
ii. выпукла по Φ;
iii. субаддитивна: $I(S_{12}, \Phi_1 \otimes \Phi_2) \leq I(S_1, \Phi_1) + I(S_2, \Phi_2)$;
iv. удовлетворяет неравенствам об обработке информации:

$$I(S, \Phi_2 \circ \Phi_1) \leq \min\{I(S, \Phi_1), I(\Phi_1(S), \Phi_2)\}.$$

Доказательство. Доказательство основано на рассмотрении различных разделений составной системы BRE, находящейся в чистом состоянии (7.39), и на использовании свойств условной энтропии.

i.

$$\begin{aligned} I(S,\Phi) &= H(R) + H(B) - H(E) \\ &= H(BE) + H(B) - H(E) \\ &= H(B|E) + H(B), \end{aligned} \tag{7.45}$$

где первое слагаемое является вогнутой функцией от S_{BE} согласно следствию 7.2.2, а второе – вогнутой функцией от S_B. Остается заметить, что отображения $S \to S_{BE}$ и $S \to S_B$ аффинны.

ii.

$$\begin{aligned} I(S,\Phi) &= H(R) + H(B) - H(E) \\ &= H(BE) + H(B) - H(BR) \\ &= H(R) - H(R|B), \end{aligned}$$

где первый член не зависит от Φ, а второй является вогнутой функцией от состояния $S_{BR} = (\Phi \times \mathrm{Id}_R)[S_{AR}]$, которое аффинно зависит от Φ.

iii. Пусть S_{12} – состояние в A_1A_2 и пусть R – эталонная система, такая что S_{12} – частичный след в A_1A_2 чистого состояния в системе A_1A_2R. Обозначим через E_1, E_2 окружения каналов Φ_1, Φ_2, соответственно. Докажем неравенство

$$\begin{aligned} I(S_{12}, \Phi_1 \otimes \Phi_2) &= H(S_{12}) + H((\Phi_1 \otimes \Phi_2)(S_{12})) - H(S_{12}, \Phi_1 \otimes \Phi_2) \\ &= H(R) + H(B_1B_2) - H(E_1E_2) \\ &\le H(\underbrace{R_1}_{=RA_2}) + H(B_1) - H(E_1) + H(\underbrace{R_2}_{=RA_1}) + H(B_2) - H(E_2). \end{aligned}$$

Так как полная система $B_1B_2E_1E_2R$ описывается чистым состоянием, то это эквивалентно неравенству

$$\begin{aligned} &H(B_1B_2E_1E_2) + H(B_1B_2) - H(E_1E_2) \\ \le\ &H(B_1E_1) + H(B_1) - H(E_1) + H(B_2E_2) + H(B_2) - H(E_2), \end{aligned}$$

откуда

$$\begin{aligned} &H(B_1B_2E_1E_2) + H(E_1) + H(E_2) \\ \le\ &H(E_1E_2) + H(B_1E_1) + H(B_2E_2), \end{aligned}$$

что может быть получено повторным использованием свойства строгой субаддитивности.

Заметим, что последнее неравенство выражает *субаддитивность условной энтропии*:

$$H(B_1B_2|E_1E_2) \leq H(B_1|E_1) + H(B_2|E_2).$$

iv. Обозначим через A входное пространство канала Φ_1, через B выходное пространство для Φ_1, которое является входным для Φ_2, и через C выходное пространство для Φ_2. Через $E_{1,2}$ обозначим окружения каналов $\Phi_{1,2}$. Имеем

$$I(S, \Phi_2 \circ \Phi_1) = H(R) + H(C) - H(E_1E_2). \tag{7.46}$$

Для доказательства первого неравенства об обработке информации заметим, что

$$I(S, \Phi_1) = H(R) + H(B) - H(E_1), \tag{7.47}$$

так что доказываемое неравенство принимает вид $H(C) - H(E_1E_2) \leq H(B) - H(E_1)$. Поскольку системы CRE_1E_2 и BRE_1 описываются чистыми состояниями, это эквивалентно неравенству

$$H(R|E_1E_2) \leq H(R|E_1),$$

которое выполнено в силу монотонности условной энтропии.

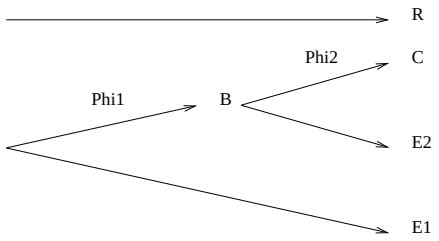


Рис. 7.2. *Композиция каналов.*

Для доказательства второго неравенства заметим, что

$$I(\Phi_1[S], \Phi_2) = H(B) + H(C) - H(E_2) \tag{7.48}$$

так что нужно доказать неравенство $H(R) - H(E_1E_2) \leq H(B) - H(E_2)$. Поскольку системы BRE_1 и CRE_1E_2 описываются чистыми состояниями, то это равносильно неравенству $H(R) - H(CR) \leq H(RE_1) - H(CRE_1)$, или

$$H(C|RE_1) \leq H(C|R).$$

Последнее же неравенство выполнено в силу монотонности условной энтропии. □

7.7 Комментарии

1. Энтропия оператора плотности была введена фон Нейманом [21]. Подробный обзор свойств квантовой энтропии и квантовой относительной энтропии можно найти в статье Верля [152] и книге Ойя и Петца [129].

Можно усилить неравенство (7.3) посредством замены $\mathrm{Tr}(S-T)^2$ на $\|S-T\|_1^2$, см. [139]; доказательство основано на соответствующем классическом неравенстве Пинскера, лемма 12.6.1 [72].

2. Вещественная функция g, определенная на интервале $I \subseteq \mathbb{R}$, называется *операторно-монотонной*, если для любого n и эрмитовых $n \times n$-матриц $A \leq B$ со спектром в I выполняется $g(A) \leq g(B)$. Аналогично определяются операторно-выпуклые функции. Основные результаты теории операторно-монотонных и операторно-выпуклых функций, а также большое количество полезных матричных неравенств можно найти в книге Бхатиа [65]. Используя классические результаты из теории операторно-монотонных и операторно-выпуклых функций, можно показать, что класс $\mathcal{G}$ состоит из операторно-выпуклых функций g на $(0,\infty)$, таких что $g(1)=0$.

Известны и другие доказательства теоремы о монотонности 7.2.1, которые также достаточно сложны. Обычный подход состоит в использовании "свойства выпуклости Либа" [129], [22]. В серии работ [123], [124], [125] Линдблад установил ряд последовательных приближений к свойству монотонности, и в итоге доказал его эквивалентность сильной субаддитивности квантовой энтропии, доказанной ранее Либом и Рускаи [122]. Ульман [151] дал другое доказательство, основанное на интерполяционном методе (см. обзоры [152], [133]). Здесь мы опираемся на работу Лесниевски и Рускаи [120], в которой дается наиболее прямое доказательство. Оно основано на обобщении неравенства Коши- Буняковского и позволяет установить монотонность для целого класса инвариантов пар состояний в "квантовой геометростатистике", инициированной Морозовой и Ченцовым [20]. Эфрос указал на связь этого круга вопросов с операторным обобщением неравенства Йенсена и понятием матричной перспективы операторно-выпуклой функции [82].

Недавно было предложено новое "естественное" доказательство свойства монотонности (Белакович, Зигмунд-Шульце [66] и Хайяши [92]), основанное на операторной интерпретации квантовой относительной энтропии как оптимального показателя экспоненциального убывания ошибки в задаче различения двух квантовых состояний (квантовый аналог леммы Стейна в математической статистике, см. теорему 12.8.1 [72]).

3. Лемма 7.3.1 была доказана Фаннесом [83]. Решение задачи 7.3.1 см. в [129], лемма 1.7. Ауденаерт [54] установил неулучшаемое неравенство

$$|H(S_1) - H(S_2)| \leq \log(d-1) \cdot \|S_1 - S_2\|_1 + h_2(\|S_1 - S_2\|_1).$$

В книге [133] указано на связь этого неравенства с леммой Фано 4.4.3.

4. Информационная корреляция была введена Линдбладом [123], который высказал следующую гипотезу:

$$I(1;2) \leq 2 \max_{M_1, M_2} I(M_1, M_2).$$

Сцепленность формирования введена в основополагающей работе Беннета, ДиВинченцо, Смолина и Вуттерса [60]. Для смешанных состояний имеется целый ряд других важных других мер сцепленности. Количественная теория сцепленности является еще одним разделом квантовой информатики, в котором математические методы получили существенное приложение и развитие, см., например, обзоры Кейля [112], Альбера и др. [52], книгу Хайаши [92].

Определение и свойства выпуклой оболочки, а также общая теорема двойственности выпуклого программирования имеются, например, в книгах Рокафеллара [25], Магарил-Ильяева и Тихомирова [17]. Доказательство полунепрерывности сверху сцепленности формирования опирается на конструкцию из работы Широкова [140].

Квантовая условная энтропия введена Адами и Серфом [50]. Операциональное истолкование отрицательных значений условной энтропии как своего рода кредита информации предложено Городецким, Оппенхаймом и Винтером [107].

5. Обменная энтропия была введена Линдбладом [126] (который однако использовал другую терминологию), и впоследствии, в контексте квантовой теории информации, независимо Барнумом, Нильсеном и Шумахером [56].

6. Квантовая взаимная информация также появилась в работе Линдблада [126], а впоследствии – в работе Адами и Серфа [50], где были подробно рассмотрены ее свойства.

8. Классическая пропускная способность квантового канала связи

8.1 Теорема кодирования

В главе 5 была установлена фундаментальная теорема кодирования для классически-квантового канала связи. В настоящей главе, используя этот результат, мы получим теорему кодирования для произвольного квантового канала.

Рассмотрим канал $\Phi : \mathfrak{T}(\mathcal{H}_1) \to \mathfrak{T}(\mathcal{H}_2)$ и соответствующий составной канал $\Phi^{\otimes n} = \Phi \otimes \cdots \otimes \Phi : \mathfrak{T}(\mathcal{H}_1^{\otimes n}) \to \mathfrak{T}(\mathcal{H}_2^{\otimes n})$. Блочный код для такого составного канала включает c-q канал, кодирующий классические сообщения i состояниями во входном пространстве $\mathcal{H}_1^{\otimes n}$, и q-c канал (наблюдаемую) $M^{(n)}$ в пространстве $\mathcal{H}_2^{\otimes n}$, декодирующий выходные состояния в классические сообщения j:

$$i \to \underbrace{S_i^{(n)}}_{\substack{\text{входное}\\ \text{состояние}}} \to \underbrace{\Phi^{\otimes n}[S_i^{(n)}]}_{\substack{\text{выходное}\\ \text{состояние}}} \stackrel{M^{(n)}}{\to} j$$

Это приводит к следующему определению

Определение 8.1.1 *Код $(\Sigma^{(n)}, M^{(n)})$ размера N для составного канала $\Phi^{\otimes n}$ состоит из кодирования, задаваемого набором $\Sigma^{(n)} = \{S_i^{(n)}; i = 1, \ldots, N\}$ состояний в $\mathcal{H}_1^{\otimes n}$ и декодирования, описываемого наблюдаемой $M^{(n)} = \{M_j^{(n)}; j = 0, 1, \ldots, N\}$ в $\mathcal{H}_2^{\otimes n}$.*

Максимальная вероятность ошибки такого кода равна

$$P_e(\Sigma^{(n)}, M^{(n)}) = \max_{i=1,\ldots,N} \{1 - p_{\Sigma M}(i|i)\}, \tag{8.1}$$

где $p_{\Sigma M}(j|i) = \operatorname{Tr} \Phi^{\otimes n}[S_i^{(n)}] M_j^{(n)}$ – вероятность принятия решение j при условии, что было послано сообщение i. Минимум ошибки $P_e(\Sigma^{(n)}, M^{(n)})$ по всевозможным кодам длины n и размера N по-прежнему обозначается $p_e(n, N)$. *Классическая пропускная способность* $C(\Phi)$ квантового канала Φ определяется как точная верхняя грань достижимых скоростей R, для которых

$$\lim_{n\to\infty} p_e(n, 2^{nR}) = 0.$$

Напомним, что конечное распределение вероятностей π на множестве квантовых состояний $\mathfrak{S}(\mathcal{H})$, которое приписывает вероятности π_i состояниям S_i, называется *ансамблем*. Если задан ансамбль $\pi^{(n)}$ с вероятностями $\{\pi_i^{(n)}\}$ входных состояний $S_i^{(n)}$, то, используя условную вероятность $p_{\Sigma M}(j|i)$, можно найти совместное распределение входа i и выхода j и вычислить шенноновскую информацию $\mathcal{I}_n(\pi^{(n)}, M^{(n)})$. Аналогично предложению 5.4.3 получаем

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} \sup_{\{\pi^{(n)}, M^{(n)}\}} \mathcal{I}_n(\pi^{(n)}, M^{(n)}). \tag{8.2}$$

Отличие от случая c-q канала состоит в том, что здесь нет фиксированного входного алфавита и приходится оптимизировать не только входное распределение $\{\pi^{(n)}\}$ и выходную наблюдаемую $M^{(n)}$, но и всевозможные состояния $S_i^{(n)}$ на входе канала $\Phi^{\otimes n}$.

Предложение 8.1.1 *Классическая пропускная способность канала Φ равна*

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} C_\chi(\Phi^{\otimes n}), \tag{8.3}$$

где

$$C_\chi(\Phi) = \sup_{\{\pi_i, S_i\}} \chi\left(\{\pi_i\}; \{\Phi[S_i]\}\right), \tag{8.4}$$

величина χ определяется соотношением (5.8) и супремум берется по всевозможным входным ансамблям $\pi = \{\pi_i, S_i\}$ в $\mathcal{H}$.

Заметим, что аналогично предложению 5.4.3, в соотношениях (8.2), (8.3) предел при $n \to \infty$ равен супремуму по n благодаря супераддитивности соответствующих последовательностей.

Доказательство. Неравенство $\leq$ в (8.3) вытекает из (8.2) и из верхней границы в теореме 5.3.2. Покажем, что

$$C(\Phi) \geq \lim_{n\to\infty} \frac{1}{n} C_\chi(\Phi^{\otimes n}) \equiv \bar{C}(\Phi). \tag{8.5}$$

Возьмем $R < \bar{C}(\Phi)$, тогда можно выбрать n_0 и ансамбль $\pi^{(n_0)} = \{\pi_i^{(n_0)}, S_i^{(n_0)}\}$ в $\mathcal{H}_1^{\otimes n_0}$, такие что

$$n_0 R < \chi\left(\{\pi_i^{(n_0)}\}; \{\Phi^{\otimes n_0}\left[S_i^{(n_0)}\right]\}\right). \tag{8.6}$$

Рассмотрим c-q канал $\tilde{\Phi}$ в $\mathcal{H}_2^{\otimes n_0}$, определенный соотношением

$$i \to \Phi^{\otimes n_0}[S_i^{(n_0)}]. \tag{8.7}$$

Согласно теореме кодирования 5.2.1 для c-q каналов, пропускная способность канала $\tilde{\Phi}$ равна

$$C(\tilde{\Phi}) = \max_{\pi} \chi\left(\{\pi_i\}; \{\Phi^{\otimes n_0}\left[S_i^{(n_0)}\right]\}\right), \tag{8.8}$$

где состояния фиксированы и максимум берется по распределению вероятностей π. В силу (8.6) это больше, чем $n_0 R$. Обозначая $\tilde{p}_e(n,N)$ минимальную вероятность ошибки для канала $\tilde{\Phi}$, имеем

$$p_e(nn_0, 2^{(nn_0)R}) \leq \tilde{p}_e(n, 2^{n(n_0 R)}), \tag{8.9}$$

поскольку любой код размера N для канала $\tilde{\Phi}^{\otimes n}$ определяет код того же размера для $\Phi^{\otimes nn_0}$. Таким образом, при $R < \bar{C}(\Phi)$, правая, а значит и левая части соотношения (8.9) стремятся к нулю при $n \to \infty$. Рассуждая как в предложении 5.4.3, получаем $p_e(n', 2^{n'R}) \to 0$ при $n' \to \infty$, и следовательно (8.5). □

8.2 χ-пропускная способность

Мы будем называть величину $C_\chi(\Phi)$, определенную в (8.4), *χ-пропускной способностью* канала Φ. Итак,

$$C_\chi(\Phi) = \sup_{\{\pi_i, S_i\}} \left[H\left(\sum_i \pi_i \Phi[S_i]\right) - \sum_i \pi_i H\left(\Phi[S_i]\right) \right], \tag{8.10}$$

что можно также переписать в виде

$$C_\chi(\Phi) = \sup_{S \in \mathfrak{S}(\mathcal{H})} \left[H\left(\Phi\left(S\right)\right) - \hat{H}_\Phi\left(S\right) \right], \tag{8.11}$$

где введена новая характеристика канала

$$\hat{H}_\Phi\left(S\right) = \inf_{\pi: \sum_i \pi_i S_i = S} \sum_i \pi_i H(\Phi(S_i)) \tag{8.12}$$

– *выпуклая оболочка* выходной энтропии $H\left(\Phi[S]\right)$. Здесь инфимум берется по всевозможным ансамблям π с фиксированным средним состоянием $\sum_i \pi_i S_i = S$.

Задача 8.2.1 *Докажите, что для идеального канала* $\hat{H}_{\mathrm{Id}}(S) \equiv 0$.

Величина (8.12) тесно связана со сцепленностью формирования (7.27), а именно,

$$\hat{H}_\Phi\left(S\right) = E_F(VSV^*),$$

где V – изометрия в представлении Стайнспринга (6.8) канала Φ, так что VSV^* является состоянием составной системы "выход-окружение". Доказательство следующей леммы аналогично доказательству предложения 7.4.1.

Лемма 8.2.1 *Выпуклая оболочка выходной энтропии $\hat{H}_\Phi(S)$ является непрерывной выпуклой функцией на $\mathfrak{S}(\mathcal{H})$. Инфимум в (8.12) достигается на ансамбле, состоящем не более чем из d_1^2 чистых состояний, $d_1 = \dim \mathcal{H}_1$.*

Следствие 8.2.1 *Супремум в выражении (8.10) для χ-пропускной способности достигается на ансамбле, состоящем не более чем из d_1^2 чистых состояний.*

Задача 8.2.2 *Используя задачу 4.3.2, докажите необходимость и достаточность для оптимальности ансамбля $\pi = \{\pi_i, S_i\}$, со средним состоянием $\bar{S}_\pi = \sum_i \pi_i S_i$, следующего условия максимальной удаленности: найдется положительное μ, такое что*

$$H(\Phi[S]; \Phi[\bar{S}_\pi]) \leq \mu, \quad \text{для всех (чистых) входных состояний } S, \quad (8.13)$$

причем равенство достигается для членов ансамбля $S = S_j$ с $\pi_j > 0$; при этом с необходимостью $\mu = C_\chi(\Phi)$. Указание:

$$\frac{\partial}{\partial \pi_j} \chi\left(\{\pi_i\}; \{\Phi\left[S_i\right]\}\right) = H(\Phi\left[S_j\right]; \Phi\left[\bar{S}_\pi\right]) - \log \mathrm{e}.$$

χ-пропускная способность может быть вычислена для ряда интересных каналов.

Задача 8.2.3 *Покажите, что $C_\chi(\mathrm{Id}) = \log d$ для идеального канала в d-мерном гильбертовом пространстве. Для квантового стирающего канала $C_\chi(\Phi_p) = (1-p)\log d$, где p – вероятность стирания. В обоих случаях оптимальный ансамбль состоит из равновероятных чистых состояний, отвечающих ортонормированному базису в $\mathcal{H}$.*

Предложение 8.2.2 *Пусть Φ – ковариантный канал (см. раздел 6.7),*

$$\Phi[U_g^{(1)} S (U_g^{(1)})^*] = U_g^{(2)} \Phi[S] (U_g^{(2)})^*.$$

Если представление $U_g^{(1)}$ неприводимо, то

$$C_\chi(\Phi) = H\left(\Phi\left[\frac{I_1}{d_1}\right]\right) - \min_S H(\Phi[S]), \quad (8.14)$$

где минимум берется по множеству чистых состояний.

Доказательство. Предположим сначала, что группа симметрий G конечна. Неприводимость представления означает отсутствие нетривиальных инвариантных подпространств, что эквивалентно следующему свойству: всякий оператор, коммутирующий со всеми $U_g^{(1)}$, кратен единичному оператору. Отсюда, аналогично доказательству соотношения (6.49), получаем

$$|G|^{-1}\sum_{g\in G}U_g^{(1)}S(U_g^{(1)})^*=\frac{I_1}{d_1}. \tag{8.15}$$

Покажем, что

$$\max_{S\in\mathfrak{S}(\mathcal{H})}H\left(\Phi[S]\right)=H\left(\Phi\left[\frac{I_1}{d_1}\right]\right). \tag{8.16}$$

Это вытекает из того, что функция $S\to H\left(\Phi[S]\right)$ вогнута и инвариантна относительно преобразований $S\to U_g^{(1)}SU_g^{A*}$, поэтому для произвольного S

$$H\left(\Phi[S]\right)=|G|^{-1}\sum_{g\in G}H\left(\Phi\left[U_g^{(1)}S(U_g^{(1)})^*\right]\right)\le H\left(\Phi\left[|G|^{-1}\sum_{g\in G}U_g^{(1)}S(U_g^{(1)})^*\right]\right). \tag{8.17}$$

Неравенство $\le$ вытекает тогда из (8.15), так что достаточно показать

$$C_\chi(\Phi)\ge H\left(\Phi\left[\frac{I_1}{d_1}\right]\right)-\min_S H(\Phi[S]). \tag{8.18}$$

Выберем состояние S_0, минимизирующее выходную энтропию. Поскольку выходная энтропия вогнута, она достигает минимума на множестве чистых состояний. Поэтому значение в правой части (8.14) достигается для ансамбля состояний $S_g=U_g^{(1)}S_0U_g^{A*};g\in G$, взятых с равными вероятностями $\pi_g=|G|^{-1}$.

Если группа $|G|$ непрерывна, то применимо аналогичное рассуждение, но оптимизирующее распределение будет непрерывным, а именно, равномерным распределением на G. Затем можно использовать конечную аппроксимацию и свойства непрерывности. □

Пример Для деполяризующего канала (6.51)

$$C_\chi(\Phi)=\log d+\left(1-p\frac{d-1}{d}\right)\log\left(1-p\frac{d-1}{d}\right)+p\frac{d-1}{d}\log\frac{p}{d}, \tag{8.19}$$

что достигается для ансамбля равновероятных чистых состояний, отвечающих ортонормированному базису в $\mathcal{H}$.

В самом деле, канал является неприводимо ковариантным, поэтому применимо предложение 8.2.2. Для произвольного чистого входного состояния, выход $\Phi[|\psi\rangle\langle\psi|]$ имеет простое собственное значение $\left(1-p\frac{d-1}{d}\right)$ и $d-1$ собственных значений $\frac{p}{d}$. Поэтому выходная энтропия равна

$$-\left(1-p\frac{d-1}{d}\right)\log\left(1-p\frac{d-1}{d}\right)-p\frac{d-1}{d}\log\frac{p}{d}$$

для всех чистых входных состояний, откуда следует соотношение (8.19).

Пример Вычислим χ-пропускную способность произвольного q-битного унитального канала Φ. Такой канал неприводимо ковариантен по отношению к представлению (6.50) (см. задачу 6.8.3), поэтому мы вновь можем применить предложение 8.2.2.

При вычислении $\min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S))$ можно считать, что $\Phi = \Lambda$, принимая во внимание унитарную инвариантность квантовой энтропии. Заметим также, что согласно задаче 2.1.2, энтропия q-битного состояния (2.8) равна

$$H(S(\boldsymbol{a})) = h_2\left(\frac{1-|\boldsymbol{a}|}{2}\right). \tag{8.20}$$

Теперь учтем, что шар Блоха сжимается каналом Λ в эллипсоид с полуосями $|\lambda_\gamma|, \gamma = x, y, z$, и минимум выходной энтропии достигается на конце самой длинной полуоси, отвечающем состоянию с собственными значениями $\frac{1\pm\max_\gamma|\lambda_\gamma|}{2}$. Это дает значение χ-пропускной способности

$$C_\chi(\Phi) = 1 - h_2\left(\frac{1-\max_\gamma|\lambda_\gamma|}{2}\right). \tag{8.21}$$

8.3 Проблема аддитивности

8.3.1 Эффект сцепленности в кодировании и декодировании

Одной из наиболее интересных и трудных математических проблем в квантовой теории информации является *гипотеза аддитивности*:

$$C_\chi(\Phi_1\otimes\Phi_2) = C_\chi(\Phi_1) + C_\chi(\Phi_2), \tag{8.22}$$

для произвольных каналов Φ_1, Φ_2. Если свойство (8.22) выполнено для некоторого канала $\Phi = \Phi_1$ и произвольного канала Φ_2, то имеет место равенство

$$C_\chi(\Phi^{\otimes n}) = nC_\chi(\Phi) \tag{8.23}$$

и в силу соотношения (8.3) классическая пропускная способность канала Φ равна его χ-пропускной способности:

$$C(\Phi) = C_\chi(\Phi). \tag{8.24}$$

До сих это было установлено лишь для некоторых важных классов, таких как каналы, разрушающие сцепленность, q-битные унитальные каналы и деполяризующий канал. Более того, недавно появились указания на то, что по крайней мере в высоких размерностях могут существовать каналы, для которых свойства (8.22), (8.23), (8.24) не выполняются.

Аддитивность (4.3.1) шенноновской пропускной способности обусловлена в конечном счете отсутствием сцепленности в классических системах. Чтобы увидеть это, попытаемся обобщить на квантовый случай классическое доказательство аддитивности, использующее условия Куна-Таккера (см. предложение 4.3.1). Пусть Φ_1, Φ_2 – два канала, оптимальные ансамбли которых имеют средние состояния $\bar{S}_{\pi^1}, \bar{S}_{\pi^2}$, и мы хотим доказать, что

$$C_\chi(\Phi_1 \otimes \Phi_2) \leq C_\chi(\Phi_1) + C_\chi(\Phi_2). \tag{8.25}$$

Тогда условие (8.13), примененное к произведению каналов $\Phi_1 \otimes \Phi_2$ требует выполнения неравенства

$$H((\Phi_1 \otimes \Phi_2)(S_{12}); (\Phi_1 \otimes \Phi_2)(\bar{S}_{\pi^1} \otimes \bar{S}_{\pi^2}) \leq C_\chi(\Phi_1) + C_\chi(\Phi_2) \tag{8.26}$$

для *всех* чистых входных состояний S_{12} произведения каналов. Равенство должно выполняться для тензорных произведений членов оптимальных ансамблей с положительными вероятностями, что легко получается из соответствующих равенств для каналов Φ_1, Φ_2, как и неравенство (8.26) для состояний-произведений $S_{12} = S_1 \otimes S_2$. Однако доказательство неравенства (8.26) для *сцепленных* состояний не проще, чем доказательство самой гипотезы аддитивности.

Аддитивность величины $C_\chi(\Phi)$ имела бы удивительное физическое следствие – это означало бы, что использование сцепленных кодовых состояний не увеличивает пропускную способность квантового канала. С другой стороны, мы видели в разделе 5.4, что использование сцепленных декодирований приводит к такому эффекту. Рассмотрим это подробнее.

Кодирование $\Sigma^{(n)} = \{S_i^{(n)}\}$ в определении 8.1.1 назовем *несцепленным*, если

$$S_i^{(n)} = \sum_{x^n} p(x^n|i) S_{x_1} \otimes \cdots \otimes S_{x_n}, \tag{8.27}$$

где S_{x_k} – состояние в k-й копии пространства $\mathcal{H}$, а $p(x^n|i)$ – условная вероятность, описывающая предварительную обработку классического входного сообщения. Напомним, что декодирование $M^{(n)}$ называется несцепленным, если оно имеет вид (5.33).

Для данного квантового канала Φ можно определить четыре классические пропускные способности $C_{1,1}, C_{1,\infty}, C_{\infty,1}, C_{\infty,\infty}$, где первый (второй) индекс относится к кодированию (декодированию), ∞ означает использование произвольных сцепленных блочных кодирований (декодирований), тогда как 1 означает ограничение несцепленными кодированиями (декодированиями) в определении 8.1.1 и, соответственно, в формуле (8.2). Очевидно, что $C_{\infty,\infty} = C(\Phi)$.

Задача 8.3.1 *Используя классическую теорему кодирования, покажите, что* $C_{1,1} = \max_{\pi,M} \mathcal{I}_1(\pi, M)$.

Величину $C_{1,1}$ будем называть *шенноновской пропускной способностью* квантового канала Φ.

Задача 8.3.2 *Используя квантовую теорему кодирования 5.2.1, покажите, что* $C_{1,\infty} = C_\chi(\Phi)$.

Соотношения между четырьмя пропускными способностями показаны на следующей диаграмме:

$$\begin{array}{ccc} C_{\infty,1} & \leq & C_{\infty,\infty}(=C) \\ || & & VI \\ C_{1,1} & \leq & C_{1,\infty}(=C_\chi) \end{array} \tag{8.28}$$

где $\leq$ в данном случае следует понимать как "меньше или равно для всех каналов и строго меньше для некоторых".

Равенство

$$C_{1,1} = C_{\infty,1} \tag{8.29}$$

означает, что использование сцепленных входных состояний при несцепленных выходных наблюдаемых не увеличивает количество информации; оно будет доказано ниже. Верхнее неравенство $C_{\infty,1} \leq C_{\infty,\infty}$ следует из этого равенства и нижнего неравенства, которое выражает возможность строгой супераддитивности по отношению к декодированиям, продемонстрированную в разделе 5.4.

Доказательство равенства (8.29). Пусть Φ некоторый канал и π – распределение вероятностей, приписывающее вероятность π_k входному состоянию S_k, и пусть $M = \{M_j\}$ – наблюдаемая, измеряемая на выходе канала. Обозначим $\mathcal{I}_\Phi(\pi, M)$ классическую шенноновскую информацию, отвечающую входному распределению π и переходной вероятности $p(j|k) = \operatorname{Tr}\Phi[S_k]M_j = \operatorname{Tr} S_k\Phi^*[M_j]$. Чтобы доказать (8.29), достаточно установить неравенство

$$\max_{\{\pi,M^1,M^2\}} \mathcal{I}_{\Phi_1\otimes\Phi_2}(\pi, M^1\otimes M^2)$$

$$\leq \max_{\{\pi^1,M^1\}} \mathcal{I}_{\Phi_1}(\pi^1, M^1) + \max_{\{\pi^2,M^2\}} \mathcal{I}_{\Phi_2}(\pi^2, M^2), \tag{8.30}$$

В терминах классической взаимной информации имеем

$$\mathcal{I}_\Phi(\pi, M) = I(X;Y),$$

где X – входная случайная величина, принимающая значения k, а Y – выходная величина, принимающая значения j. Чтобы доказать (8.30), рассмотрим состояния S_k в гильбертовом пространстве $\mathcal{H}_1\otimes\mathcal{H}_2$ канала $\Phi_1\otimes\Phi_2$, с наблюдаемой $M^1\otimes M^2$. Тогда условная вероятность равна

$$p(j_1, j_2|k) = \operatorname{Tr} S_k(\Phi_1^*[M_{j_1}]\otimes\Phi_2^*[M_{j_2}]) = p^1(j_1|j_2,k)p^2(j_2|k), \tag{8.31}$$

где

$$p^1(j_1|j_2,k) = \mathrm{Tr}_1\, S^1_{j_2,k}\Phi_1^*[M_{j_1}], \quad p^2(j_2|k) = \mathrm{Tr}_2\, S^2_k\Phi_2^*[M_{j_2}],$$

и

$$S^2_k = \mathrm{Tr}_1\, S_k, \qquad S^1_{j_2,k} = \frac{\mathrm{Tr}_2\, S_k(I\otimes\Phi_2^*[M_{j_2}])}{\mathrm{Tr}\, S_k(I\otimes\Phi_2^*[M_{j_2}])}.$$

ЗдесьTr$_s$ обозначает частичный след по s-й подсистеме ($s=1,2$) в $\mathcal{H}_1\otimes\mathcal{H}_2$.

Имеем

$$I(X;Y_1Y_2) = H(Y_1Y_2) - H(Y_1Y_2|X),$$

где $H(\cdot), H(\cdot|\cdot)$, соответственно, энтропия и условная энтропия случайных величин. В силу субаддитивности классической энтропии,

$$H(Y_1Y_2) \le H(Y_1) + H(Y_2).$$

С другой стороны,

$$H(Y_1Y_2|X) = H(Y_1|Y_2X) + H(Y_2|X).$$

Сопоставляя, получаем

$$I(X;Y_1Y_2) \le I(XY_2;Y_1) + I(X;Y_2),$$

что в силу (8.31) сводится к

$$\mathcal{I}_{\Phi_1\otimes\Phi_2}(\pi, M^1\otimes M^2) \le \mathcal{I}_{\Phi_1}(\pi^1, M^1) + \mathcal{I}_{\Phi_2}(\pi^2, M^2),$$

где π^1 – распределение, приписывающее вероятности $\pi_k p^2(j_2|k)$ состояниям $S^1_{j_2,k}$, а π^2 – распределение, приписывающее вероятности π_k состояниям S^2_k. Беря максимум по π и S, получаем (8.30). □

8.3.2 Иерархия свойств аддитивности

Важной характеристикой квантового канала Φ, которая уже появлялась в разделе 8.2, является минимальная выходная энтропия

$$\check{H}(\Phi) = \min_{S\in\mathfrak{S}(\mathcal{H})} H(\Phi(S)). \tag{8.32}$$

Соответствующая ей гипотеза аддитивности имеет вид

$$\check{H}\left(\Phi_1\otimes\Phi_2\right) = \check{H}\left(\Phi_1\right) + \check{H}\left(\Phi_2\right). \tag{8.33}$$

В силу вогнутости энтропии, минимум в (8.32) достигается на чистом состоянии $S\in\mathrm{extr}\mathfrak{S}(\mathcal{H})$. В силу этого, классический аналог величины $\check{H}(\Phi)$ аддитивен, поскольку всякое чистое состояние составной классической системы является произведением чистых состояний подсистем. В

квантовом случае для тензорного произведения $\mathcal{H}_1 \otimes \mathcal{H}_2$, описывающего составную систему,

$$\mathrm{extr}\mathfrak{S}(\mathcal{H}_1 \otimes \mathcal{H}_2) \supsetneqq \mathrm{extr}\mathfrak{S}(\mathcal{H}_1) \times \mathrm{extr}\mathfrak{S}(\mathcal{H}_2), \tag{8.34}$$

благодаря наличию сцепленных состояний, так что для аддитивности (8.33) нет очевидной причины.

Ясно, что неравенство $\leq$ всегда имеет место в (8.33). Отсюда следует, что (8.33) очевидно выполняется для каналов с нулевой минимальной выходной энтропией, т.е. таких, для которых существует чистое выходное состояние, например, для идеального и стирающего каналов.

Задача 8.3.3 *Свойства (8.22) и (8.33) равносильны для двух ковариантных каналов, удовлетворяющих условиям предложения 8.2.2. Указание: используйте тот факт, что тензорное произведение неприводимых представлений двух групп симметрий G_1, G_2 является неприводимым представлением группы $G_1 \times G_2$.*

Рассмотрим теперь выпуклую оболочку выходной энтропии $\hat{H}_\Phi(S)$, определенную соотношением (8.12). Гипотетическое свойство супераддитивности для этой величины формулируется следующим образом: *для произвольного состояния $S_{12} \in \mathfrak{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ и каналов Φ_1, Φ_2*

$$\hat{H}_{\Phi_1 \otimes \Phi_2}(S_{12}) \geq \hat{H}_{\Phi_1}(S_1) + \hat{H}_{\Phi_2}(S_2), \tag{8.35}$$

где S_1, S_2 – частичные следы состояния S_{12} в $\mathcal{H}_1, \mathcal{H}_2$.

Свойство супераддитивности (8.35) связано со следующим свойством сцепленности формирования: пусть $\mathcal{H}_1 = \mathcal{H}_1^A \otimes \mathcal{H}_1^B$ и $\mathcal{H}_2 = \mathcal{H}_2^A \otimes \mathcal{H}_2^B$, и пусть S_{12}^{AB} – произвольное состояние в $\mathcal{H}_1 \otimes \mathcal{H}_2$, тогда

$$E_F(S_{12}^{AB}) \geq E_F(S_1^{AB}) + E_F(S_2^{AB}), \tag{8.36}$$

где сцепленность формирования определятся по отношению к разбиению $A|B$. А именно, если бы гипотеза (8.36) выполнялась для всех состояний S_{12}^{AB}, то гипотеза (8.35) выполнялась бы для всех каналов Φ_1, Φ_2 и всех состояний S_{12}, тогда как если бы (8.35) выполняласьдля всех S_{12} и каналов частичного следа, то (8.36) выполнялась бы для всех состояний S_{12}^{AB}.

Предложение 8.3.1 *Для данных каналов Φ_1, Φ_2 свойство супераддитивности (8.35) влечет аддитивность как минимальной выходной энтропии (8.33), так и χ-пропускной способности (8.22).*

Доказательство. В самом деле, пусть S_{12}^0 доставляет минимум функции $H(\Phi_1 \otimes \Phi_2)(S_{12})$, тогда

$$\begin{aligned}\check{H}(\Phi_1 \otimes \Phi_2) &= H((\Phi_1 \otimes \Phi_2)(S_{12}^0)) \geq \hat{H}_{\Phi_1 \otimes \Phi_2}(S_{12}^0) \\ &\geq \hat{H}_{\Phi_1}(S_1^0) + \hat{H}_{\Phi_2}(S_2^0) \geq \check{H}(\Phi_1) + \check{H}(\Phi_2),\end{aligned}$$

откуда следует (8.33). С другой стороны, (8.35) и субаддитивность квантовой энтропии влечет

$$\begin{aligned} & H((\Phi_1 \otimes \Phi_2)(S_{12})) - \hat{H}_{\Phi_1 \otimes \Phi_2}(S_{12}) \\ \leq\ & H((\Phi_1 \otimes \Phi_2)(S_{12})) - \hat{H}_{\Phi_1}(S_1) - \hat{H}_{\Phi_2}(S_2) \\ \leq\ & \left[H(\Phi_1(S_1)) - \hat{H}_{\Phi_1}(S_1)\right] + \left[H(\Phi_2(S_2)) - \hat{H}_{\Phi_2}(S_2)\right]. \end{aligned} \tag{8.37}$$

Используя (8.11), получаем

$$C_\chi(\Phi_1 \otimes \Phi_2) \leq C_\chi(\Phi_1) + C_\chi(\Phi_2),$$

т. е. (8.22). □

Далее, свойство (8.35) тесно связано с аддитивностью χ-пропускной способности с ограничениями на входе. Пусть F – положительный оператор во входном гильбертовом пространстве $\mathcal{H}$ канала, E – положительная постоянная. Рассмотрим χ-пропускную способность при линейном ограничении $\operatorname{Tr} SF \leq E$ на входное состояние S:

$$C_\chi(\Phi, F, E) = \max_{S:\operatorname{Tr} SF \leq E} \left[H\left(\Phi\left(S\right)\right) - \hat{H}_\Phi\left(S\right)\right]. \tag{8.38}$$

Задача 8.3.4 *Пусть имеется два канала Φ_1, Φ_2 с соответствующими ограничениями F_1, F_2. Тогда неравенство (8.35) вместе с (8.11) влечет*

$$C_\chi(\Phi_1 \otimes \Phi_2, F_1 \otimes I_2 + I_1 \otimes F_2, E) = \max_{E_1+E_2=E} \left[C_\chi(\Phi_1, F_1, E_1) + C_\chi(\Phi_2, F_2, E_2)\right]. \tag{8.39}$$

Замечательно, однако, что все сформулированные гипотезы аддитивности оказываются равносильными в следующем смысле.

Теорема 8.3.2 *Гипотезы (8.33), (8.22), (8.23) и (8.35) глобально эквивалентны в том смысле, что если одна из них выполняется для всех каналов, то и другие выполняются для всех каналов. Более того, они глобально эквивалентны супераддитивности сцепленности формирования (8.36) и аддитивности χ-пропускной способности с произвольными входными ограничениями.*

Из этой теоремы также вытекает, что контрпример к одной из этих гипотез влечет существование контрпримеров ко всем остальным.

Мы не будем приводить доказательство этой теоремы и ограничимся рассмотрением некоторых случаев, в которых удается доказать наиболее сильное свойство (8.35).

Предложение 8.3.3 *Пусть Φ_2 произвольный канал. Свойство (8.35) выполняется, если Φ_1 ортогональная выпуклая сумма (см. определение 6.6.1) идеального канала и канала $\Phi_{(0)}$, такого, что это свойство выполняется для пары $\Phi_{(0)}$, Φ_2.*

Отсюда следует, что для такого Φ_1 выполняется также аддитивность минимальной выходной энтропии (8.33) и $\chi-$ пропускной способности (8.22). Например, это имеет место для стирающего канала, поскольку он является ортогональной прямой суммой идеального канала и канала, разрушающего сцепленность, для которого свойство (8.35) будет установлено в разделе 8.3.3.

Доказательство. Обозначим $\Phi^{(q)} = q\mathrm{Id} \oplus (1-q)\Phi^{(0)}$. Тогда

$$H(\Phi^{(q)}(S)) = qH(S) + (1-q)H(\Phi^{(0)}(S)) + h_2(q),$$

где $h_2(q) = -q\log q - (1-q)\log(1-q)$ – двоичная энтропия,

$$\hat{H}_{\Phi^{(q)}}(S) = (1-q)\hat{H}_{\Phi^{(0)}}(S) + h_2(q), \tag{8.40}$$

поскольку минимум в выражении для $\hat{H}_{\Phi^{(0)}}(S)$ достигается для ансамбля чистых состояний S_j, для которого $H(S_j) = 0$. Для произвольного канала Φ_2 имеем

$$\Phi^{(q)} \otimes \Phi_2 = q(\mathrm{Id} \otimes \Phi_2) \oplus (1-q)(\Phi^{(0)} \otimes \Phi_2).$$

Тогда

$$\begin{aligned}\hat{H}_{\Phi^{(q)}\otimes\Phi_2}(S_{12}) &\geq q\hat{H}_{\mathrm{Id}\otimes\Phi_2}(S_{12}) + (1-q)\hat{H}_{\Phi^{(0)}\otimes\Phi_2}(S_{12}) + h_2(q)\\ &\geq q\hat{H}_{\Phi_2}(S_2) + (1-q)\left[\hat{H}_{\Phi^{(0)}}(S_1) + \hat{H}_{\Phi_2}(S_2)\right] + h_2(q),\end{aligned}$$

где была использована супераддитивность величины $\hat{H}_{\mathrm{Id}\otimes\Phi_2}(S_{12}), \hat{H}_{\Phi^{(0)}\otimes\Phi_2}(S_{12})$ и тот факт, что $\hat{H}_{\mathrm{Id}}(S_1) \equiv 0$. Используя (8.40), получаем, что правая часть равна $\hat{H}_{\Phi^{(q)}}(S_1) + \hat{H}_{\Phi_2}(S_2)$.

Задача 8.3.5 *Используя разложение Шмидта (теорема 3.1.2), покажите, что $H(\Phi[S]) = H(\tilde{\Phi}[S])$ для любого чистого состояния S, где $\tilde{\Phi}$ – канал, комплементарный к Φ. Поэтому если какое-либо из свойств (8.35), (8.33) выполняется для каналов Φ_1, Φ_2, то аналогичное свойство выполняется для комплементарных каналов $\tilde{\Phi}_1, \tilde{\Phi}_2$.*

Отсюда в частности следует, что предложение 8.3.3 имеет место с заменой идеального канала на комплементарный ему полностью деполяризующий.

8.3.3 Некоторые энтропийные неравенства

Получим неравенства, которые позволяют в некоторых случаях доказать супераддитивность выпуклой оболочки (8.35), и, следовательно, аддитивность минимальной выходной энтропии (8.33) и χ-пропускной способности (8.22).

Рассмотрим состояние S и измерение, описываемое семейством операторов A_k, удовлетворяющих условию нормировки $\sum_k A_k^* A_k = I$, так

что исход k появляется с вероятностью $\pi_k = \operatorname{Tr} SA_k^* A_k$ и приводит к апостериорному состоянию системы $S_k = A_k S A_k^* / \pi_k$. Энтропии начального и апостериорных состояний связаны *неравенством Линдблада-Озава*

$$H(S) \geq \sum_k \pi_k H(S_k). \tag{8.41}$$

Для доказательства положим $\mathcal{H} = \mathcal{H}_1$ и рассмотрим очищение $|\psi\rangle$ состояния S в $\mathcal{H}_1 \otimes \mathcal{H}_2$. Тогда

$$\pi_k S_k = \operatorname{Tr}_{\mathcal{H}_2}(A_k \otimes I)|\psi\rangle\langle\psi|(A_k \otimes I)^*,$$

и $S_2 = \operatorname{Tr}_{\mathcal{H}_1} S = \sum_k \pi_k S_2^k$, где

$$\pi_k S_2^k = \operatorname{Tr}_{\mathcal{H}_1}(A_k \otimes I)|\psi\rangle\langle\psi|(A_k \otimes I)^*.$$

Применяя теорему 3.1.2 дважды и используя вогнутость квантовой энтропии, получаем

$$H(S) = H(S_2) = H\left(\sum_k \pi_k S_2^k\right) \geq \sum_k \pi_k H\left(S_2^k\right) = \sum_k \pi_k H(S_k).$$

□

Частным случаем является следующее неравенство для составной системы. Пусть S_{12} – состояние составной системы 12 и пусть $\left\{|e_k^2\rangle\right\}$ – ортонормированный базис в $\mathcal{H}_2$. Полагая $A_k = I_1 \otimes |e_k^2\rangle\langle e_k^2|$, получаем

$$H(S_{12}) \geq \sum_k \pi_k H\left(S_1^k\right), \tag{8.42}$$

где $\pi_k S_1^k = \langle e_k^2|S_{12}|e_k^2\rangle$. Это может быть использовано для доказательства гипотезы аддитивности в следующем частном случае.

Предложение 8.3.4 *Свойство (8.35) выполняется в случае, когда $\Phi_1 = \Phi$ произвольный канал в $\mathcal{H}_1$, а $\Phi_2 = \mathrm{Id}_2$ – идеальный канал в $\mathcal{H}_2$.*

Доказательство. Поскольку $\hat{H}_{\mathrm{Id}}(S) \equiv 0$ согласно задаче 8.2.1, мы должны доказать, что

$$\hat{H}_{(\Phi\otimes\mathrm{Id}_2)}(S_{12}) \geq \hat{H}_\Phi(S_1). \tag{8.43}$$

Пусть $S_{12} = \sum_i \pi_i S_{12}^i$ – оптимальное разложение, т. е.

$$\hat{H}_{(\Phi\otimes\mathrm{Id}_2)}(S_{12}) = \sum_i \pi_i H\left((\Phi \otimes \mathrm{Id}_2)[S_{12}^i]\right).$$

Используя (8.42), получаем

$$\sum_i \pi_i H\left((\Phi \otimes \mathrm{Id}_2)[S_{12}^i]\right) \geq \sum_{ik} \pi_i \pi_{ik} H\left(\Phi\left[S_1^{ik}\right]\right), \tag{8.44}$$

где $\pi_{ik} S_1^{ik} = \langle e_k^2 | S_{12}^i | e_k^2 \rangle$. Теперь мы имеем разложение S_1 в состояния S_1^{ik} с вероятностями $\pi_i \pi_{ik}$ для канала Φ, поэтому правая часть соотношения (8.44) больше или равна $\hat{H}_\Phi(S_{12})$.

Предложение 8.3.5 *Свойство (8.35) выполняется в случае, когда $\Phi_1 = \Phi$ – канал, разрушающий сцепленность (см. раздел 6.4), а $\Phi_2 = \Psi$ – произвольный канал.*

Лемма 8.3.6 *Пусть $\pi_j \geq 0, \sum \pi_j = 1$, и S_1^j, S_2^j – произвольные состояния систем $1, 2$. Тогда*

$$H\left(\sum_j \pi_j S_1^j \otimes S_2^j\right) \geq H\left(\sum_j \pi_j S_1^j\right) + \sum_j \pi_j H\left(S_2^j\right).$$

Доказательство. Введем обозначения

$$(\bar{S}_{12})_\pi = \sum_j \pi_j S_1^j \otimes S_2^j; \quad (\bar{S}_1)_\pi = \sum_j \pi_j S_1^j.$$

Тогда, используя свойство аддитивности энтропии, доказываемое неравенство можно переписать как

$$H((\bar{S}_{12})_\pi) - \sum_j \pi_j H(S_1^j \otimes S_2^j) \geq H((\bar{S}_1)_\pi) - \sum_j \pi_j H(S_1^j),$$

что в силу тождества (7.17) равносильно

$$\sum_j \pi_j H(S_1^j \otimes S_2^j; (\bar{S}_{12})_\pi) \geq \sum_j \pi_j H(S_1^j; (\bar{S}_1)_\pi).$$

Последнее же неравенство вытекает из свойства монотонности относительной энтропии (по отношению к взятию частичного следа по второй системе). □

Доказательство. Возвращаясь к доказательству предложения, предположим, что Φ – канал, разрушающий сцепленность, так что $\Phi[S] = \sum_j S_1^j \mathrm{Tr} S M_1^j$, где $\left\{M_1^j\right\}$ является наблюдаемой в системе 1, тогда

$$(\Phi \otimes \mathrm{Id}_2)[S_{12}] = \sum_j p_j S_1^j \otimes S_2^j, \tag{8.45}$$

где $p_j S_2^j = \mathrm{Tr}_{\mathcal{H}_1} S_{12}\left(M_1^j \otimes I_2\right)$ для произвольного состояния S_{12} составной системы. Беря частичные следы, получаем, в частности,

$$\Phi[S_1] = \sum_j p_j S_1^j, \quad S_2 = \sum_j p_j S_2^j. \tag{8.46}$$

Если Ψ – произвольный канал в системе 2, то

$$(\Phi \otimes \Psi)[S_{12}] = \sum_j p_j S_1^j \otimes \Psi[S_2^j]. \tag{8.47}$$

Имеем

$$\hat{H}_{\Phi\otimes\Psi}(S_{12}) = \inf \sum_i \pi_i H\left((\Phi \otimes \Psi)[S_{12}^i]\right), \tag{8.48}$$

где $S_{12} = \sum_i \pi_i S_{12}^i$ – произвольное разложение. Записывая соотношение (8.45) для S_{12}^i, получаем

$$(\Phi \otimes \mathrm{Id}_2)[S_{12}^i] = \sum_j p_{ij} S_1^{ij} \otimes S_2^{ij},$$

причем имеют место разложения

$$S_1 = \sum_i \pi_i S_1^i, \quad S_2 = \sum_{ij} \pi_i p_{ij} S_2^{ij}. \tag{8.49}$$

Используя (8.47), доказанную лемму и (8.46), получаем

$$\begin{aligned} H\left((\Phi \otimes \Psi)[S_{12}^i]\right) &= H\left(\sum_j p_{ij} S_1^{ij} \otimes \Psi[S_2^{ij}]\right) \\ &\geq H\left(\sum_j p_{ij} S_1^{ij}\right) + \sum_j p_{ij} H(\Psi[S_2^{ij}]) \\ &\geq H\left(\Phi[S_1^i]\right) + \sum_j p_{ij} H(\Psi[S_2^{ij}]), \end{aligned}$$

что в сочетании с (8.48), (8.49) дает

$$\hat{H}_{\Phi\otimes\Psi}(S_{12}) \geq \hat{H}_\Phi[S_1] + \hat{H}_\Psi[S_2].$$

□

8.4 Передача классической информации с помощью сцепленного состояния

8.4.1 Выигрыш благодаря сцепленности

Конструкция протокола сверхплотного кодирования (раздел 3.2.3) допускает обобщение на случай пространства произвольной конечной размерности, а также неидеального квантового канала Φ между A и B. Рассмотрим следующий протокол передачи информации через канал Φ. Системы A (приемник) и B (передатчик) имеют общее сцепленное состояние

S_{AB}, которое распределяется им перед началом связи с помощью некоторой процедуры типа описанной в разделе 3.1.4. Будем предполагать, что $\dim \mathcal{H}_A \leq \dim \mathcal{H}_B$. Система A кодирует классические сообщения i, поступающие с вероятностями π_i в операции (каналы) $\mathcal{E}_A^i$, действующие в $\mathcal{H}_A$. Эти операции применяются системой A к ее части состояния S_{AB}, при этом состояние системы AB преобразуется в $(\mathcal{E}_A^i \otimes \mathrm{Id}_B)[S_{AB}]$. После этого A посылает свою часть состояния по данному каналу Φ, в результате чего на конце B становится доступной измерению система AB в состоянии $(\Phi \circ \mathcal{E}_A^i \otimes \mathrm{Id}_B)[S_{AB}]$. Классическая информация извлекается путем измерения некоторой наблюдаемой в пространстве $\mathcal{H}_{AB}$. Разрешается блочное кодирование, так что на самом деле, все это описание относится к n-й тензорной степени пространства $\mathcal{H}_{AB}$. Нас интересует пропускная способность такого протокола, которая называется *пропускной способностью протокола передачи классической информации с помощью сцепленного состояния (entanglement-assisted classical capacity)*.

Максимум по измерениям B может быть оценен с помощью теоремы кодирования 5.2.1 для классической пропускной способности. Сначала введем величину

$$C_{ea}^{(1)}(\Phi) = \sup_{\pi_i, \mathcal{E}_A^i, S_{AB}^i} \chi\left(\{\pi_i\}; (\Phi \otimes \mathrm{Id}_B)\left[S_{AB}^i\right]\right), \tag{8.50}$$

в которой уже учтены измерения в системе B. Используя канал n раз и произвольные (сцепленные) измерения в системе B, получаем величину

$$C_{ea}^{(n)}(\Phi) = C_{ea}^{(1)}(\Phi^{\otimes n}). \tag{8.51}$$

Тогда полная классическая пропускная способность с использованием сцепленного состояния есть

$$C_{ea}(\Phi) = \lim_{n\to\infty} \frac{1}{n} C_{ea}^{(1)}(\Phi^{\otimes n}). \tag{8.52}$$

Следующая теорема дает простое "однобуквенное" выражение для $C_{ea}(\Phi)$ в терминах квантовой взаимной информации.

Теорема 8.4.1 *Классическая пропускная способность с использованием сцепленного состояния равна*

$$C_{ea}(\Phi) = \max_{S_A} I(S_A, \Phi) = \max_{S_A} \left[H(S_A) + H(\Phi(S_A)) - H(S_A; \Phi)\right]. \tag{8.53}$$

Доказательство этой теоремы будет приведено в последующих разделах. Здесь мы проведем сравнение классических пропускных способностей с использованием и без использования сцепленного состояния.

Пример Рассмотрим квантовый стирающий канал Φ_p. Имеем $H(\Phi_p[S_A]) = (1-p)H(S_A) + h_2(p)$. Используя тот факт, что комплементарный канал $\tilde{\Phi}_p = \Phi_{1-p}$ (см. задачу 6.6.3) и определение 7.5.1 обменной энтропии, получаем $H(S_A; \Phi) = H(\tilde{\Phi}[S_A]) = pH(S_A) + h_2(p)$. Поэтому

$$I(S_A, \Phi_p) = H(S_A) + (1-p)H(S_A) - pH(S_A) = 2(1-p)H(S_A),$$

откуда $C_{ea}(\Phi_p) = 2(1-p)\log d$, что в два раза превышает классическую пропускную способность $C(\Phi_p)$ стирающего канала.

Предложение 8.4.2 *Если канал Φ ковариантен, то максимум в (8.53) достигается на инвариантном состоянии S_A. В частности, если Φ неприводимо ковариантен, то максимум достигается на хаотическом состоянии.*

Доказательство. Пусть канал Φ удовлетворяет условию ковариантности (6.53), тогда

$$I(V_g^{(1)} S_A V_g^{(1)*}, \Phi) = I(S_A, \Phi). \tag{8.54}$$

Это следует из выражения (7.44) для взаимной информации, унитарной инвариантности входной и выходной энтропий и соответствующего свойства обменной энтропии:

$$H(V_g^{(1)} S_A V_g^{(1)*}, \Phi) = H(S_A, \Phi). \tag{8.55}$$

Чтобы доказать это свойство, рассмотрим формулу (7.38) для обменной энтропии и заметим, что очищение состояния $V_g^{(1)} S_A V_g^{(1)*}$ дается вектором $(V_g^{(1)} \otimes I_R)|\psi_{AR}\rangle$. Подставляя в (7.38) и вновь используя ковариантность Φ и унитарную инвариантность энтропии, получаем (8.55).

Пусть теперь S_A – оптимальное состояние для канала Φ. Рассмотрим $V_g^{(1)}$-инвариантное состояние

$$\bar{S} = \frac{1}{|G|} \sum_{g \in G} V_g^{(1)} S_A V_g^{(1)*}.$$

В силу предложения 7.6.1, функция $S \to I(S, \Phi)$ вогнута, поэтому

$$I(\bar{S}, \Phi) \geq \frac{1}{|G|} \sum_{g \in G} I(V_g^{(1)} S_A V_g^{(1)*}, \Phi) = I(S_A, \Phi),$$

где последнее равенство следует из (8.54). □

Пример Рассмотрим деполяризующий канал (6.51). Используя унитарную ковариантность, получаем, что $I(S, \Phi)$ достигает максимума на хаотическом состоянии $\bar{S} = \frac{I}{d}$, для которого $H(\bar{S}) = H(\Phi[\bar{S}]) = \log d$. Обменная энтропия $H(\bar{S}, \Phi)$ дается соотношением (7.43), откуда получаем

$$C_{ea}(\Phi) = \log d^2 + \left(1 - p\frac{d^2-1}{d^2}\right) \log \left(1 - p\frac{d^2-1}{d^2}\right) + p\frac{d^2-1}{d^2} \log \frac{p}{d^2}. \tag{8.56}$$

Сравним это с величиной $C_\chi(\Phi)$, которая дается формулой (8.19).

Мы видим, в частности, что $\frac{C_{ea}(\Phi)}{C_\chi(\Phi)} \to d+1$ в пределе сильного шума $p \to 1$ (когда обе пропускные способности стремятся к нулю!) Для максимального значения $p = \frac{d^2}{d^2-1}$

$$C_{ea}(\Phi) = \log \frac{d^2}{d^2-1},$$

$$C_\chi(\Phi) = \frac{1}{d+1} \log \frac{d}{d+1} + \frac{d}{d+1} \log \frac{d^2}{d^2-1}.$$

Отношение $\frac{C_{ea}(\Phi)}{C_\chi(\Phi)}$ монотонно возрастает от значения 5.08 при $d = 2$, приближаясь к асимптоте $2(d+1)$ с ростом d.

Таким образом, сцепленные состояния вновь выступают как "катализатор" при передаче классической информации – они могут в принципе неограниченно увеличить классическую пропускную способность существующего канала, хотя сами по себе передавать информацию не могут.

Задача 8.4.1 *Покажите, что для q-битного унитального канала вида (6.54), где Λ дается соотношением (6.59),*

$$C_{ea}(\Phi) = I(\bar{S}, \Phi) = 2\log 2 + \sum_{\gamma=0,x,y,z} \mu_\gamma \log \mu_\gamma.$$

Обменная энтропия $H(\bar{S}, \Phi)$ может быть вычислена с помощью предложения 7.5.1 и представления Крауса (6.59), приводя к выражению

$$H(\bar{S}, \Phi) = -\sum_{\gamma=0,x,y,z} \mu_\gamma \log \mu_\gamma.$$

Сравните это с пропускной способностью $C_\chi(\Phi)$, даваемой формулой (8.21).

В общем случае имеются простые неравенства, связывающие две классические пропускные способности:

$$C_\chi(\Phi) \le C_{ea}(\Phi) \le \log d + C_\chi(\Phi). \tag{8.57}$$

Доказательство. Пусть S_A – смесь произвольных чистых состояний S_j с вероятностями p_j. Тогда (7.42) влечет

$$I(S_A, \Phi) \le H(\sum_j p_j S_j) + H(\sum_j p_j \Phi(S_j)) - \sum_j p_j H(\Phi(S_j)).$$

Беря максимум и используя (8.53), получаем второе неравенство в (8.57).

Хотя использование сцепленного состояния и представляет собой дополнительный ресурс, первое неравенство не вполне очевидно в силу специального характера процедуры кодирования, используемой стороной A. Чтобы избежать этой трудности, покажем, что определение величины $C_{ea}^{(1)}(\Phi)$, а значит, и $C_{ea}(\Phi)$, может быть сформулировано без явного использования кодирующих операций $\mathcal{E}_A^i$, а именно

$$C_{ea}^{(1)}(\Phi) = \tag{8.58}$$

$$\sup_{\pi_i, \{S_{AB}^i\} \in \Sigma_B} \left[H\left(\sum_i \pi_i (\Phi \otimes \mathrm{Id}_B) \left[S_{AB}^i \right] \right) - \sum_i \pi_i H\left((\Phi \otimes \mathrm{Id}_B) \left[S_{AB}^i \right] \right) \right],$$

где Σ_B обозначает множество семейств состояний $\{S_{AB}^i\}$, удовлетворяющих условию, что частичные следы S_B^i не зависят от i, $S_B^i = S_B$.

Первое неравенство в (8.57) тогда получается, если положить $S_{AB}^i = S_A^i \otimes S_B$. Чтобы доказать (8.58), достаточно установить следующий факт.

Лемма 8.4.3 *Пусть $\{S_{AB}^i\}$ – семейство состояний, удовлетворяющих условию $S_B^i = S_B$. Тогда найдутся чистое состояние S_{AB} и кодирования $\mathcal{E}_A^i$, такие что*

$$S_{AB}^i = (\mathcal{E}_A^i \otimes \mathrm{Id}_B)[S_{AB}]. \tag{8.59}$$

Доказательство. Для простоты предположим сначала, что оператор S_B невырожден. Тогда

$$S_B = \sum_{k=1}^d \lambda_k |e_k^B\rangle\langle e_k^B|,$$

где $\lambda_k > 0$, а $\{|e_k^B\rangle\}$ – ортонормированный базис в $\mathcal{H}_B$. Пусть $\{|e_k^A\rangle\}$ – ортонормированный базис в $\mathcal{H}_A$. Для вектора $|\psi^A\rangle = \sum_{k=1}^d c_k |e_k^A\rangle$ обозначим $|\bar{\psi}^B\rangle = \sum_{k=1}^d \bar{c}_k |e_k^B\rangle$. Отображение $|\psi^A\rangle \to |\bar{\psi}^B\rangle$ является антиизоморфизмом $\mathcal{H}_A$ и $\mathcal{H}_B$. Положим

$$|\psi_{AB}\rangle = \sum_{k=1}^d \sqrt{\lambda_k} |e_k^A\rangle \otimes |e_k^B\rangle,$$

так что $S_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}|$, и определим кодирования соотношением

$$\mathcal{E}_A^i \left[|\psi^A\rangle\langle\phi^A| \right] = \langle \bar{\psi}^B | S_B^{-1/2} S_{AB}^i S_B^{-1/2} | \bar{\phi}^B \rangle, \quad |\psi^A\rangle, |\phi^A\rangle \in \mathcal{H}_A.$$

(Здесь обозначения Дирака используются для определения "частичного" скалярного произведения в пространстве $\mathcal{H}_B$). Нетрудно проверить, что отображения $\mathcal{E}_A^i$ действительно являются каналами, удовлетворяющими соотношению (8.59).

Если оператор S_B вырожден, приведенную выше конструкцию следует модифицировать, заменив $S_B^{-1/2} S_{AB}^i S_B^{-1/2}$ в предыдущей формуле на $\sqrt{S_B^-} S_{AB}^i \sqrt{S_B^-} + P_B^0$, где S_B^- – обобщенный обратный оператора S_B, а P_B^0 – проектор на нулевое подпространство оператора S_B.

□

8.4.2 Доказательство обращения теоремы кодирования

Чтобы доказать неравенство

$$C_{ea}(\Phi) \leq \max_{S_A} I(S_A, \Phi), \tag{8.60}$$

покажем сначала, что

$$C_{ea}^{(1)}(\Phi) \leq \max_{S_A} I(S_A, \Phi). \tag{8.61}$$

Обозначим $\mathcal{E}_A^i$ кодирования, используемые стороной A. Пусть S_{AB} – исходное чистое состояние системы AB, тогда состояние системы AB (соответственно A) после кодирования есть

$$S_{AB}^i = (\mathcal{E}_A^i \otimes \mathrm{Id}_B)[S_{AB}], \quad S_A^i = \mathcal{E}_A^i[S_A]. \tag{8.62}$$

Заметим, что частичное состояние B не изменяется после кодирования, $S_B^i = S_B$. Докажем, что

$$H\left(\sum_i \pi_i (\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^i]\right) - \sum_i \pi_i H\left((\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^i]\right)$$

$$\leq I\left(\sum_i \pi_i S_A^i, \Phi_A\right). \tag{8.63}$$

Максимум левой части по всевозможным $\pi_i, \mathcal{E}_A^i$ равен $C_{ea}^{(1)}(\Phi)$, откуда будет следовать (8.61).

Используя субаддитивность квантовой энтропии, мы можем оценить первое слагаемое в левой части (8.63) как

$$H\left(\sum_i \pi_i \Phi_A[S_A^i]\right) + H(S_B) = H\left(\Phi_A\left[\sum_i \pi_i S_A^i\right]\right) + \sum_i \pi_i H(S_B).$$

Первый член уже дает выходную энтропию из $I\left(\sum_i \pi_i S_A^i; \Phi_A\right)$. Оценим остальные члены

$$\sum_i \pi_i \left[H(S_B) - H\left((\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^i]\right)\right].$$

Покажем сначала, что слагаемое в квадратных скобках не превосходит $H(S_A^i) - H\left((\Phi_A \otimes \mathrm{Id}_{R^i})[S_{AR^i}^i]\right)$, где R^i эталонная система для очищения S_A^i, а $S_{AR^i}^i$ соответствующее чистое состояние. С этой целью рассмотрим

унитарное расширение кодирования $\mathcal{E}_A^i$ с окружением E_i, которое находится в начальном чистом состоянии, в соответствии с теоремой 6.3.1. Из (8.62) видно, что можно считать $R^i = BE_i$ (после взаимодействия, которое затрагивает только AE_i). Тогда, обозначая штрихом состояния после применения канала Φ_A, получаем

$$H(S_B) - H\left((\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^i]\right) = H(S_B) - H(S_{A'B}^i) = -H_i(A'|B), \quad (8.64)$$

где нижний индекс i условной энтропии указывает на состояние $S_{A'B}^i$. Аналогично

$$H(S_A^i) - H\left((\Phi_A \otimes \mathrm{Id}_{R^i})[S_{AR^i}^i]\right) = H(S_{R^i}^i) - H\left(S_{A'R^i}^i\right)$$
$$= -H_i(A'|R^i) = -H_i(A'|BE_i),$$

что больше или равно (8.64), согласно монотонности условной энтропии.

Используя вогнутость функции $S_A \to H(S_A) - H\left((\Phi_A \otimes \mathrm{Id}_R)[S_{AR}]\right)$ которая будет установлена ниже, получаем

$$\sum_i \pi_i \left[H(S_A^i) - H\left((\Phi_A \otimes \mathrm{Id}_{R^i})[S_{AR^i}^i]\right)\right]$$
$$\leq H\left(\sum_i \pi_i S_A^i\right) - H\left((\Phi_A \otimes \mathrm{Id}_R)[\hat{S}_{AR}]\right),$$

где $\hat{S}_{AR}$ – состояние, очищающее $\sum_i \pi_i S_A^i$ с эталонной системой R. Для доказательства вогнутости обозначим E окружение для канала Φ_A, тогда получим

$$H(S_A) - H\left((\Phi_A \otimes \mathrm{Id}_R)[S_{AR}]\right) = H(S_R) - H\left(S_{A'R}\right)$$
$$= H(S_{A'E'}) - H(S_{E'}) = H(A'|E')$$

Условная энтропия $H(A'|E')$ вогнутая функция состояния $S_{A'E'}$ согласно следствию 7.4.1. Отображение $S_A \to S_{A'E'}$ аффинно, поэтому $H(A'|E')$ – вогнутая функция состояния S_A .

Применяя это рассуждение к каналу $\Phi^{\otimes n}$, получаем

$$C_{ea}^{(n)}(\Phi) \leq \max_{S_A} I(S_A, \Phi^{\otimes n}). \quad (8.65)$$

Согласно субаддитивности квантовой взаимной информации (свойство *iii*. в лемме 7.6.1), имеем

$$\max_{S_{12}} I(S_{12}, \Phi_1 \otimes \Phi_2) = \max_{S_1} I(S_1, \Phi_1) + \max_{S_2} I(S_2, \Phi_2),$$

откуда вытекает замечательное свойство аддитивности

$$\max_{S_A} I(S_A, \Phi^{\otimes n}) = n \max_{S_A} I(S_A, \Phi).$$

Окончательно, получаем (8.60). □

8.4.3 Доказательство прямого утверждения теоремы кодирования

Здесь будет доказано неравенство

$$C_{ea}(\Phi) \geq \max_{S_A} I(S_A, \Phi). \tag{8.66}$$

Сначала покажем, используя обобщение протокола сверхплотного кодирования, что

$$C_{ea}^{(1)}(\Phi^{\otimes n}) \geq I\left(\frac{P}{\dim P}, \Phi^{\otimes n}\right) \tag{8.67}$$

для произвольного проектора P в $\mathcal{H}_A^{\otimes n}$.

В самом деле, пусть $P = \sum_{k=1}^{m} |e_k\rangle\langle e_k|$, где $\{e_k; k = 1, \ldots, m = \dim P\}$ – ортонормированная система. Рассмотрим дискретную систему Вейля $\{W(x,y); \quad x,y = 1, \ldots, m\}$, определенную соотношениями (6.46) на подпространстве $\mathcal{H}_P = P\mathcal{H}_A^{\otimes n}$.

Поскольку $\dim \mathcal{H}_A \leq \dim \mathcal{H}_B$, мы можем предположить, что $\mathcal{H}_A \subseteq \mathcal{H}_B$. Пусть

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |e_k\rangle \otimes |e_k\rangle.$$

Задача 8.4.2 *Докажите, что система* $\{(W(x,y) \otimes I_B)\,|\psi_{AB}\rangle; x,y = 1, \ldots, m\}$ *является ортонормированным базисом в* $\mathcal{H}_P \otimes \mathcal{H}_P$. *В частности,*

$$\sum_{x,y=1}^{m} (W(x,y) \otimes I_B)\,|\psi_{AB}\rangle\langle\psi_{AB}|\,(W(x,y) \otimes I_B)^* = P \otimes P. \tag{8.68}$$

Операторы $W(x,y)$ будут играть роль, аналогичную матрицам Паули в протоколе сверхплотного кодирования для q-бита. Пусть теперь классический сигнал принимает значения $i = (x,y)$ с равными вероятностями $1/m^2$, в качестве сцепленного состояния возьмем $|\psi_{AB}\rangle\langle\psi_{AB}|$, а в качестве кодирующих отображений $\mathcal{E}_A^i[S] = W(x,y)SW(x,y)^*$. Тогда получим

$$C_{ea}^{(1)}(\Phi^{\otimes n}) \geq H\left(\frac{1}{m^2}\sum_{x,y}(\Phi \otimes \mathrm{Id}_B)^{\otimes n}[S_{AB}^{xy}]\right) - \frac{1}{m^2}\sum_{x,y} H\left((\Phi \otimes \mathrm{Id}_B)^{\otimes n}[S_{AB}^{xy}]\right),$$

где $S_{AB}^{xy} = \left(W(x,y) \otimes I_B^{\otimes n}\right)|\psi_{AB}\rangle\langle\psi_{AB}|\left(W(x,y) \otimes I_B^{\otimes n}\right)^*$. Согласно (8.68) первое слагаемое в правой части равно

$$H\left((\Phi \otimes \mathrm{Id}_B)\left[\frac{P}{m} \otimes \frac{P}{m}\right]\right) = H\left(\frac{P}{m}\right) + H\left(\Phi\left[\frac{P}{m}\right]\right).$$

Поскольку состояние S_{AB}^{xy} очищает смешанное состояние $\frac{P}{m}$, все энтропии во втором слагаемом одинаковы и равны $H\left(\frac{P}{m}, \Phi\right)$. Учитывая выражение для квантовой взаимной информации

$$I(S_A, \Phi) = H(S_A) + H(\Phi[S_A]) - H(S_A; \Phi)$$

получаем(8.67). Для будущего напомним, что последнее слагаемое – обменная энтропия – равно энтропии окружения $H(S'_E) = H(\Phi_E[S_A])$, где Φ_E канал из пространства системы $\mathcal{H}_A$ в пространство окружения $\mathcal{H}_E$ задаваемый формулой (7.37).

Пусть теперь $S_A = S$ произвольное состояние в $\mathcal{H}_A$, и пусть $\hat{P}^{n,\delta}$ *сильно типичный проектор* состояния $S^{\otimes n}$ в пространстве $\mathcal{H}_A^{\otimes n}$, определяемый ниже. Докажем, что *произвольного* канала Ψ из $\mathcal{H}_A$ в какое-либо гильбертово пространство $\tilde{\mathcal{H}}$ имеет место соотношение

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}H\left(\Psi^{\otimes n}\left[\frac{\hat{P}^{n,\delta}}{\dim \hat{P}^{n,\delta}}\right]\right) = H(\Psi[S]).$$

Отсюда, согласно выражениям для взаимной информации и обменной энтропии, будет следовать

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}I\left(\frac{\hat{P}^{n,\delta}}{\dim \hat{P}^{n,\delta}}, \Phi^{\otimes n}\right) = I(S, \Phi), \tag{8.69}$$

а значит, согласно (8.67), и доказываемое неравенство (8.66).

Определение 8.4.1 *Фиксируем положительное δ, и обозначим λ_j собственные числа, $|e_j\rangle$ собственные векторы оператора плотности S. Тогда собственные числа и собственные векторы оператора плотности $S^{\otimes n}$ суть $\lambda_J = \lambda_{j_1}\cdot\ldots\cdot\lambda_{j_n}$, $\quad |e_J\rangle = |e_{j_1}\rangle\otimes\ldots\otimes|e_{j_n}\rangle$, где $J = (j_1, ..., j_n)$. Последовательность J называется сильно типичной, если числа $N(j|J)$ появлений символов j в последовательности J удовлетворяют условию*

$$\left|\frac{N(j|J)}{n} - \lambda_j\right| < \delta, \quad j = 1, \ldots, d,$$

причем $N(j|J) = 0$, если $\lambda_j = 0$. Обозначим совокупность всех сильно типичных последовательностей $\hat{T}^{n,\delta}$. Сильно типичный проектор определяется как следующий спектральный проектор оператора плотности $S^{\otimes n}$:

$$\hat{P}^{n,\delta} = \sum_{J\in\hat{T}^{n,\delta}} |e_J\rangle\langle e_J|.$$

Пусть P^n распределение вероятностей, задаваемое собственными числами λ_J. Согласно закону больших чисел, $\mathsf{P}^n\left(\hat{T}^{n,\delta}\right) \to 1$ при $n\to\infty$. Для произвольной функции $f(j), j = 1, \ldots, d$, и последовательности $J = (j_1, ..., j_n) \in \hat{T}^{n,\delta}$

$$\left|\frac{f(j_1) + \cdots + f(j_n)}{n} - \sum_{j=1}^{d}\lambda_j f(j)\right| < \delta \max f. \tag{8.70}$$

В частности, любая сильно типичная последовательность является типичной в обычном (энтропийном) смысле: полагая $f(j) = -\log\lambda_j$, имеем

$$n[H(S) - \delta_1] < -\log\lambda_J < n[H(S) + \delta_1], \tag{8.71}$$

где $\delta_1 = \delta \max_{\lambda_j>0}(-\log\lambda_j)$ (обратное неверно – не всякая типичная последовательность сильно типична.)

В дальнейшем нам понадобится следующая комбинаторная лемма:

Задача 8.4.3 *Покажите, что размер множества $\hat{T}^{n,\delta}$ оценивается как*

$$2^{n[H(S)-\Delta_n(\delta)]} < \left|\hat{T}^{n,\delta}\right| < 2^{n[H(S)+\Delta_n(\delta)]}, \tag{8.72}$$

где $H(S) = -\sum_{j=1}^{d}\lambda_j\log\lambda_j$, и $\lim_{\delta\to 0}\lim_{n\to\infty}\Delta_n(\delta) = 0$.

Доказательство соотношения (8.69). Обозначим $d_{n,\delta} = \dim\hat{P}^{n,\delta} = \left|\hat{T}^{n,\delta}\right|$ и $\bar{S}^{n,\delta} = \frac{\hat{P}^{n,\delta}}{d_{n,\delta}}$ и докажем, что

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}H(\Psi^{\otimes n}\left(\bar{S}^{n,\delta}\right)) = H(\Psi[S]) \tag{8.73}$$

для произвольного канала Ψ. Имеем

$$nH(\Psi[S]) - H(\Psi^{\otimes n}\left[\bar{S}^{n,\delta}\right]) = H(\Psi[S]^{\otimes n}) - H(\Psi^{\otimes n}\left[\bar{S}^{n,\delta}\right])$$

$$= H\left(\Psi^{\otimes n}\left[\bar{S}^{n,\delta}\right];\Psi^{\otimes n}[S^{\otimes n}]\right] + \mathrm{Tr}\log\Psi[S]^{\otimes n}\left(\Psi^{\otimes n}\left[\bar{S}^{n,\delta}\right] - \Psi[S]^{\otimes n}\right). \tag{8.74}$$

Строго говоря, это преобразование имеет смысл, если оператор плотности $\Psi(S)^{\otimes n}$ невырожден, что мы сначала и предположим.

Для первого слагаемого, используя свойство монотонности относительной энтропии, имеем

$$H\left(\Psi^{\otimes n}\left[\bar{S}^{n,\delta}\right];\Psi^{\otimes n}[S^{\otimes n}]\right) \le H\left(\bar{S}^{n,\delta};S^{\otimes n}\right),$$

где правая часть

$$H\left(\bar{S}^{n,\delta};S^{\otimes n}\right) = \sum_{J\in\hat{T}^{n,\delta}}\frac{1}{d_{n,\delta}}\log\frac{1}{d_{n,\delta}\lambda_J} = -\log d_{n,\delta} - \sum_{J\in\hat{T}^{n,\delta}}\frac{1}{d_{n,\delta}}\log\lambda_J,$$

что, согласно (8.71), (8.72), меньше или равно величины $n\left(\delta_1 + \Delta_n(\delta)\right)$, стремящейся к нулю при $n\to\infty, \delta\to 0$.

Используя соотношение

$$\log\Psi[S]^{\otimes n} = \log\Psi[S]\otimes I\otimes\cdots\otimes I + \cdots + I\otimes\cdots\otimes I\otimes\log\Psi[S],$$

и вводя оператор $F = \Psi^*[\log\Psi[S]]$, где Ψ^* – сопряженный канал, мы можем переписать второе слагаемое в виде

$$n \operatorname{Tr} \frac{(F \otimes I \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes I \otimes F)}{n} \left(\bar{S}^{n,\delta} - S^{\otimes n}\right)$$
$$= \frac{n}{d_{n,\delta}} \sum_{J \in \hat{T}^{n,\delta}} \left[\frac{f(j_1) + \cdots + f(j_n)}{n} - \sum_{j=1}^{d} \lambda_j f(j) \right], \tag{8.75}$$

где $f(j) = \langle e_j|F|e_j\rangle$, что оценивается величиной $n\delta \max f$ в силу (8.70). Это доказывает соотношение (8.73) в случае невырожденного $\Psi[S]$.

В общем случае обозначим P_Ψ проектор на носитель оператора $\Psi[S]$. Тогда соответствующий проектор оператора $\Psi[S]^{\otimes n}$ есть $P_\Psi^{\otimes n}$, и носитель оператора $\Psi^{\otimes n}\left(\bar{S}^{n,\delta}\right)$ содержится в носителе $\Psi[S]^{\otimes n} = \Psi^{\otimes n}[S^{\otimes n}]$, поскольку носитель $\bar{S}^{n,\delta}$ содержится в носителе $S^{\otimes n}$. Поэтому второе слагаемое в (8.74) следует понимать как

$$\operatorname{Tr} P_\Psi^{\otimes n} \log \left[P_\Psi^{\otimes n} \Psi[S]^{\otimes n} P_\Psi^{\otimes n}\right] P_\Psi^{\otimes n} \left(\Psi^{\otimes n}\left[\bar{S}^{n,\delta}\right] - \Psi[S]^{\otimes n}\right),$$

где логарифм берется от невырожденного оператора в подпространстве $P_\Psi^{\otimes n}\mathcal{H}_A^{\otimes n}$. Теперь можно повторить рассуждения, определив F как $\Psi^*[P_\Psi(\log P_\Psi \Psi[S] P_\Psi)P_\Psi]$. Это завершает доказательство соотношения (8.69), откуда вытекает (8.66). □

8.5 Комментарии

1. Предложение 8.1.1, доказанное в работе Холево [35], было обобщено на квантовые каналы с памятью Кречманом и Вернером [119]. Обобщение на произвольные последовательности каналов в духе информационного спектра Хана и Верду дано Хайаши и Нагаока [93], [92].

2. На плодотворную связь между χ- пропускной способностью и относительной энтропией, в частности, свойство максимальной равноудаленности, указано в работе Шумахера и Вестморленда [138].

3. Проблема аддитивности упоминается в работе Беннета, Фукса и Смолина [61]. Формулировка в терминах четырех пропускных способностей была дана Беннетом и Шором [62]. Шор [144] определил классическую пропускную способность с использованием адаптивных измерений на выходе и показал, что она находится строго между $C_{1,1}$ и $C_{1,\infty}$.

Полная формулировка теоремы 8.3.2 состоит из нескольких утверждений, часть из которых носит индивидуальный характер, тогда как другие выполняются только глобально, например, "если минимальная выходная энтропия аддитивна для всех каналов, то это же верно и для χ-пропускной способности". Ряд авторов (Мацумото, Шимоно и Винтер, Ауденаэрт и Браунштейн, Померанский, Холево и Широков) внесли вклад в доказательство отдельных индивидуальных импликаций, тогда как наиболее неожиданные глобальные утверждения были получены Шором, который предложил оригинальную конструкцию расширения канала, позволяющую усилить его индивидуальные свойства. Подробности и ссылки см. в работе Шора [145]. Тот факт, что глобальное

выполнение свойства (8.23) влечет аддитивность (8.22) для всех пар каналов Φ_1, Φ_2, (а также аналогичное утверждение для минимальной выходной энтропии)) доказан Фукудой и Вольфом [87]. Предложение 8.3.3 доказано в работе Холево и Широкова [104], где была систематически исследована гипотеза (8.35) и установлена ее эквивалентность аддитивности χ-пропускной способности с ограничениями. Решение задачи 8.3.5 см. в [40].

Неравенство Линдблада-Озава установлено в [130]. Выполнение гипотезы аддитивности для каналов, разрушающих сцепленность, доказано Шором [143]. Другие важные достижения включают доказательство Кингом свойств (8.22) и (8.33) для q-битных унитальных [113] и деполяризующих [114] каналов, с использованием весьма специальных структурных свойств этих каналов и неравенства Либа-Тирринга, см. [65], раздел X.2. Подробный обзор прогресса в решении проблемы аддитивности до 2006 г. имеется в докладе Холево [106]. Из последних достижений следует отметить результаты Винтера [158] и Хайдена [94], которые показали, что гипотеза аддитивности не выполняется для минимальной выходной энтропии Реньи со значениями параметра, соответственно, $p > 2$ и $1 < p \leq 2$ при достаточно высоких размерностях. (Напомним, что энтропия фон Неймана соответствует значению $p = 1$). В [158] рассматривались каналы вида

$$\Phi(\rho) = \frac{1}{n} \sum_{j=1}^{n} U_j \rho U_j^*, \tag{8.76}$$

где $U_j; j = 1, \ldots, n$ – последовательность независимых случайных унитарных операторов, распределенных равномерно (т. е. в соответствии с нормированной мерой Хаара на группе унитарных операторов), и $n = O(d \log d)$, $d \to \infty$. Нарушение аддитивности с вероятностью, близкой к 1, следует из оценки больших уклонений для сумм случайных операторов.

Наконец совсем недавно, основываясь на методах этих работ, Хастингс [90] рассмотрел смеси случайных унитарных операторов со случайными весами и анонсировал построение примера, в котором нарушена аддитивность минимальной выходной энтропии при достаточно больших n, n/d. Подробное доказательство было проведено Фукудой, Кингом и Мозером [86]. Отсюда, в силу теоремы 8.3.2, должно следовать существование контрпримеров ко всем формам аддитивности, а также к равенству (8.24). Следует, однако, заметить, что эти методы, идейно родственные методу случайного кодирования Шеннона, не позволяют получить конструктивные контрпримеры.

4. Теорема 8.4.1 доказана Беннетом, Шором, Смолиным и Таплиалом [63]. Здесь приводится упрощенное доказательство, следующее работе Холево [101]. Относительно сильно типичных проекторов и доказательства результата задачи 8.4.3 см. [45].

9. Передача квантовой информации

Как уже неоднократно подчеркивалось, квантовое состояние само по себе представляет особого рода информационный ресурс постольку, поскольку имеет статистическую неопределенность. Статистическая механика изучает необратимые изменения состояния системы в результате естественной эволюции (следствием которых является и потеря информации). Теория информации ставит перед собой в некотором смысле обратную задачу: как свести к пренебрежимой величине эти необратимые изменения, используя некоторые специальные преобразования состояний – кодирование и декодирование – до и после эволюции, задаваемой каналом с шумом. Этому кругу вопросов и посвящена настоящая глава.

При этом мы начнем с рассмотрения методов кодирования квантовой информации, устойчивого по отношению к данному ограниченному набору ошибок (например, к произвольной ошибке в одном произвольном q-бите), когда требуется передача с абсолютной точностью. Затем мы перейдем к приближенной постановке, когда вводится некоторая мера точности воспроизведения квантового состояния. Тогда цель состоит в том, чтобы используя блочное кодирование и передачу через составные каналы $\Phi^{\otimes n}$, добиться асимптотически точной передачи квантовых состояний с максимально высокой скоростью при $n \to \infty$. Таким образом возникает аналог шенноновской теоремы кодирования для квантовой информации. Это приводит к понятию квантовой пропускной способности, которое тесно связано с так называемой когерентной информацией.

9.1 Квантовые коды, исправляющие ошибки

9.1.1 Постановка вопроса

При передаче сообщений по каналу с шумом желательно иметь код, который был бы устойчив относительно возможных ошибок. В случае классической информации, принципиальная осуществимость такого кодирования при скоростях передачи, меньших пропускной способности, вытекает из теоремы Шеннона. Однако эта теорема не дает конструктивного способа построения помехоустойчивого (пусть даже не-оптимального) кода,

и практическому решению этой проблемы посвящен специальный раздел теории информации.

Самый незамысловатый способ застраховаться от ошибок состоит в простом повторении сообщений (что, конечно, снижает скорость передачи). Рассмотрим двоичный сигнал(бит), принимающий значения $0, 1$, который может кодироваться последовательностью битов. Предположим, что вероятность ошибки – изменения значения одного бита в процессе передачи равна малой величине p, так что вероятность изменения двух битов p^2 — пренебрежимо мала. Рассмотрим код $0 \to 00$, $1 \to 11$. Тогда, получив 00 или 11, можно с высокой достоверностью утверждать, что был послан 0 или, соответственно, 1. Однако, получив 01 либо 10, можно лишь констатировать наличие ошибки, но исправить ее (т. е. восстановить с высокой степенью достоверности посланный сигнал) нельзя. От этого недостатка можно избавиться, если добавить еще один кодирующий разряд: $0 \to 000$, $1 \to 111$. Такой код будет уже помехоустойчивым по отношению к ошибке в любом одном бите, т. е. позволяющим восстановить посланный сигнал.

Прямолинейное обобщение такого рецепта на случай квантовой информации наталкивается на трудность – квантовую информацию невозможно размножить. Кроме того, при передаче через квантовый канал с шумом безошибочно должны передаваться не только базисные состояния, но и всевозможные их суперпозиции. На первый взгляд, задача может показаться неразрешимой, однако это не так. Рассмотрим пример кода Шора, который решает такую задачу в важном модельном случае.

Предположим, что ставится задача безошибочной передачи любого чистого состояния q-бита $\psi = a|\,0\rangle + b|\,1\rangle$, причем разрешается кодировать эти состояния состояниями нескольких q-битов. Кодирование в квантовом случае задается изометричным отображением. Следуя классической аналогии, рассмотрим сначала код, отображающий базисные векторы согласно правилу

$$|\,0\rangle \to |\,000\rangle, |\,1\rangle \to |\,111\rangle. \tag{9.1}$$

Такой код исправляет “переворот q-бита”, т. е. переход $|\,0\rangle \leftrightarrow |\,1\rangle$ в любом одном q-бите кода. Нас же интересует произвольное состояние $\psi = a|\,0\rangle + b|\,1\rangle$, которое при выбранном способе кодирования переходит в $a|\,000\rangle + b|\,111\rangle$. Пусть, например, произошла ошибка в первом q-бите кода:

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

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

Однако такой код не исправляет ошибки типа “переворота фазы” $|\,0\rangle \leftrightarrow |\,0\rangle, |\,1\rangle \leftrightarrow -|\,1\rangle$. В самом деле, в результате такой фазовой ошибки в одном q-бите получим $a|\,000\rangle - b|\,111\rangle$ вместо $a|\,000\rangle + b|\,111\rangle$, и эти состояния не ортогональны, т. е. не являются безошибочно различимыми в квантовой статистике.

Теперь заметим, что преобразование Адамара

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

отображает переворот фазы в переворот бита и наоборот. Преобразуя соответствующим образом код (9.1), получим код

$$\begin{aligned} |0\rangle &\to \frac{1}{2^{3/2}}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)(|0\rangle + |1\rangle), \\ |1\rangle &\to \frac{1}{2^{3/2}}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle)(|0\rangle - |1\rangle), \end{aligned} \tag{9.2}$$

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

Код Шора, который исправляет как переворот бита, так и переворот фазы в любом одном q-бите, получается комбинированием кодов (9.1), (9.2) и требует 9 q-битов

$$\begin{aligned} |0\rangle &\to \frac{1}{2^{3/2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ |1\rangle &\to \frac{1}{2^{3/2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \end{aligned} \tag{9.3}$$

Более того, оказывается, что этот код исправляет не только битовую и фазовую, но любую ошибку, возникающую в одном (любом) из q-битов, поскольку любая ошибка может быть сведена к этим двум видам ошибок. Это вытекает из выполнения общих условий, рассматриваемых в следующем разделе.

9.1.2 Общая формулировка

Пусть $\mathcal{M}$ – гильбертово пространство, которое играет роль пространства квантовых сигналов, S произвольное состояние в $\mathcal{M}$. *Кодом* называется изометричное отображение $V : \mathcal{M} \to \mathcal{N}$, переводящее состояния S в кодированные состояния VSV^* в гильбертовом пространстве кодированного сигнала $\mathcal{N}$. Система $\mathcal{N}$ может быть подвержена ошибкам, которые описываются вполне положительными отображениями Φ в $\mathfrak{T}(\mathcal{N})$, образующими в каждой конкретной задаче некоторый класс $\mathcal{E}$. Физический смысл необратимых эволюций имеют отображения, удовлетворяющими условию $\Phi(I) \leqslant I$ (см. раздел 6.5), однако в рассматриваемой ниже постановке это ограничение несущественно, т. к. эффект ошибки определяется с точностью до постоянного множителя.

Итак, преобразования квантовой информации описываются диаграммой

$$S \longrightarrow VSV^* \longrightarrow \Phi(VSV^*), \quad \Phi \in \mathcal{E}.$$

Определение 9.1.1 *Код V называется кодом, исправляющим ошибки из класса $\mathcal{E}$, если существует восстанавливающий канал $\mathcal{D}$, такой что*

$$\mathcal{D}[\Phi(VSV^*)] = c(\Phi)S, \tag{9.4}$$

для любого состояния S и любого $\Phi \in \mathcal{E}$, где $c(\Phi)$ — некоторая постоянная, зависящая только от Φ.

На самом деле код можно задавать самим подпространством $\mathcal{L} = V\mathcal{M} \subset \mathcal{N}$, не вводя явно $\mathcal{M}, V$.

Пример. Хранение квантовой информации в памяти квантового компьютера. Пусть $\mathcal{N} = \mathcal{H}_2^{\otimes n}$ – квантовый регистр, в котором предполагается хранить информацию из $\mathcal{M}$. Рассмотрим ошибки, при которых изменению может подвергнуться не более m q-битов регистра. Соответствующее множество $\mathcal{E}(n,m)$ состоит из отображений $\Phi = \Phi_1 \otimes \cdots \otimes \Phi_n$, где количество отображений $\Phi_k \neq \mathrm{Id}$ не превышает m, причем ошибка в k-м q-бите Φ_k может быть произвольным вполне положительным отображением. Операторами элементарных ошибок в каждом q-бите могут служить матрицы Паули, причем σ_x описывает переворот бита, σ_z переворот фазы, а $\sigma_y = i\sigma_x\sigma_z$ – их комбинацию. Вместе с единичным оператором I, который соответствует отсутствию ошибки, они образуют базис в алгебре наблюдаемых q-бита.

Код Шора демонстрирует возможность исправления ошибок из $\mathcal{E}(n,1)$, если n достаточно велико (можно доказать, что наименьшее значение n для кода, исправляющего одну ошибку, равно 5). Возможность исправления только одной ошибки является, конечно, серьезным ограничением. Однако известно, что существуют коды, исправляющие ошибки из $\mathcal{E}(n,m)$, где m может быть сколь угодно большим для достаточно больших размеров регистра n.

9.1.3 Необходимые и достаточные условия исправления ошибок

Класс ошибок $\mathcal{E}$ обычно имеет следующую структуру: он состоит из всевозможных отображений вида $\Phi(S) = \sum_j V_j S V_j^*$, где $V_j \in \mathrm{Lin}(B_1, \ldots, B_p)$, а B_j – фиксированные операторы *элементарных ошибок*.

Теорема 9.1.1 *Следующие утверждения эквивалентны:*

i. *код $\mathcal{L}$ исправляет ошибки класса $\mathcal{E}$;*
ii. *код $\mathcal{L}$ исправляет ошибку $\Phi[S] = \sum_{j=1}^{p} B_j S B_j^*$;*
iii. *для $\phi, \psi \in \mathcal{L}$, таких что $\langle\phi|\,\psi\rangle = 0$, имеет место $\langle\phi|\,B_i^* B_j\,|\psi\rangle = 0$, для всех $i, j = 1, \ldots p$.*
iv. *для любого ортонормированного базиса $\{|\,k\rangle\}$ в $\mathcal{L}$ выполняется*

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

v. $P_{\mathcal{L}} B_i^* B_j P_{\mathcal{L}} = b_{ij} P_{\mathcal{L}}$, *где* $P_{\mathcal{L}}$ *– проектор на* $\mathcal{L}$.

Доказательство. i. $\Rightarrow$ *ii.* очевидно.

ii. $\Rightarrow$ *iii.* Пусть существует восстанавливающий канал $\mathcal{D}[S] = \sum_r R_r S R_r^*$ для ошибки Φ. Рассмотрим чистое входное состояние $S = |\,\psi\rangle\langle\psi|\,, |\,\psi\rangle \in \mathcal{L}$. Тогда

$$\sum_r \sum_j R_r B_j |\,\psi\rangle\langle\psi|\, B_j^* R_r^* = c |\,\psi\rangle\langle\psi|\,.$$

Но это возможно, лишь если $R_r B_j |\,\psi\rangle = c_{jr} |\,\psi\rangle$, при этом $c = \sum_r \sum_j |\,c_{jr}|^2$.

Пусть $|\,\phi\rangle$, $|\,\psi\rangle \in \mathcal{L}$ два ортогональных вектора. Поскольку восстанавливающий канал сохраняет след, то $I = \sum_r R_r^* R_r$, и значит

$$\langle\phi|\, B_i^* B_j |\,\psi\rangle = \sum_r \langle\phi|\, B_i^* R_r^* R_r B_j |\,\psi\rangle = (\sum_r \bar{c_{ir}} c_{jr}) \langle\phi|\,\psi\rangle = 0.$$

iii. $\Rightarrow$ *iv.* Второе равенство очевидно, а первое получается из рассмотрения ортогональных векторов $|\,\phi\rangle = |\,k+l\rangle$, $|\,\psi\rangle = |\,k-l\rangle$.

iv. $\Rightarrow$ *v.* Обозначая $b_{ij} = \langle k|\, B_i^* B_j |\,k\rangle$ мы можем записать условие *iv.* в виде

$$\langle k|\, B_i^* B_j |\,l\rangle = \delta_{kl} b_{ij},$$

что эквивалентно *v.*

v. $\Rightarrow$ *i.* Матрица $[b_{ij}]$ эрмитова неотрицательно определенная, поэтому найдется унитарная матрица $[u_{ij}]$, такая что

$$\sum_{ij} \bar{u}_{ir} b_{ij} u_{js} = \delta_{rs} \lambda_r,$$

где $\lambda_r \geq 0$. Полагая $\tilde{B}_s = \sum_j u_{js} B_j$, имеем $\Phi[S] = \sum_{s=1}^p \tilde{B}_s S \tilde{B}_s^*$ и

$$P_{\mathcal{L}} \tilde{B}_r^* \tilde{B}_s P_{\mathcal{L}} = \delta_{rs} \lambda_r P_{\mathcal{L}}.$$

Таким образом, при $\lambda_r > 0$ имеем

$$\lambda_s^{-1/2} \tilde{B}_s P_{\mathcal{L}} = U_s P_{\mathcal{L}},$$

где U_s частично изометричные операторы, отображающие $\mathcal{L}$ на взаимно ортогональные подпространства $\mathcal{L}_s \subset \mathcal{H}$. Отсюда следует, что

$$\Phi[S] = \sum_{s=1}^p \lambda_s U_s S U_s^* \tag{9.5}$$

для любого состояния S с $\mathrm{supp} S \subset \mathcal{L}$.

Пусть P_s – проектор на подпространство $\mathcal{L}_s$, и обозначим $P_0 = I - \sum_s P_s$. Определим канал

$$\mathcal{D}[S] = \sum_s U_s^* P_s S P_s U_s + P_0 S P_0$$

и покажем, что он является восстанавливающим для всех ошибок из $\mathcal{E}$. Поскольку линейная комбинация элементарных ошибок B_j является также линейной комбинацией операторов $\tilde{B}_r$, то, принимая во внимание, что $U_s^* P_s \tilde{B}_r P_{\mathcal{L}} = \lambda_r^{1/2} \delta_{sr} P_{\mathcal{L}}$, мы получаем (9.4) для произвольного $\Phi \in \mathcal{E}$. □

Задача 9.1.1 *Проверьте выполнение условий iv. для кода Шора (9.3) и операторов элементарных ошибок, задаваемыми матрицами Паули в первом q-бите кода.*

9.1.4 Когерентная информация и точное исправление ошибок

Рассмотрим произвольный канал Φ, действующий из A в B и $S = S_A$ – входное состояние канала. Введем эталонное пространство R, так что $S_{AR} = |\psi_{AR}\rangle\langle\psi_{AR}|$ является очищением состояния S_A. Пусть $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ – изометрия представления Стайнспринга (6.8) для канала Φ, где E – пространство окружения.

Важной компонентой квантовой взаимной информации $I(S, \Phi)$ является *когерентная информация*

$$\begin{aligned} I_c(S, \Phi) &= H(\Phi[S]) - H(S; \Phi) \\ &= H(B) - H(E) \\ &= H(B) - H(RB) \\ &= -H(R|B). \end{aligned}$$

Как мы позже увидим, это понятие тесно связано с квантовой пропускной способностью канала Φ.

Особенностью когерентной информации является *отсутствие* некоторых "естественных" свойств, таких как выпуклость по S; субаддитивность; второе неравенство об обработке информации. Кроме того, эта величина может принимать отрицательные значения. Ее классический аналог вообще не бывает положительным: $H(Y) - H(XY) = -H(X|Y) \leq 0$. Однако $I_c(S, \Phi)$ является выпуклой по Φ и удовлетворяет первому неравенству об обработке информации:

$$I_c(S, \Phi_2 \circ \Phi_1) \leq I_c(S, \Phi_1), \tag{9.6}$$

что легко выводится из соотношения $I_c(S, \Phi) = I(S, \Phi) - H(S)$ и соответствующих свойств квантовой взаимной информации $I(S, \Phi)$. Покажем, почему второе неравенство об обработке информации $I_c(S, \Phi_2 \circ \Phi_1) \leq I_c(\Phi_1[S], \Phi_2)$, вообще говоря, не выполняется. Если предположить, что это верно, то

$$H(\Phi_2 \circ \Phi_1[S]) - H(S, \Phi_2 \circ \Phi_1) \leq H(\Phi_2 \circ \Phi_1[S]) - H(\Phi_1[S], \Phi_2),$$

т. е. $H(S, \Phi_2 \circ \Phi_1) \geq H(\Phi_1[S], \Phi_2)$. Это равносильно тому, что $H(E_1E_2) \geq H(E_2)$, где E_j – окружение для j-ого канала. Но свойство монотонности в общем случае для квантовой энтропии не выполняется.

Задача 9.1.2 *Используя идею доказательства субаддитивности квантовой взаимной информации (утверждение 7.6.1), покажите, что субаддитивность $I_c(S, \Phi)$ эквивалентна следующему неравенству*

$$H(B_1B_2) - H(B_1) - H(B_2) \leq H(E_1E_2) - H(E_1) - H(E_2),$$

которое может быть не выполнено в общем случае.

Существует тесная связь между точной передачей квантовой информации, исправлением ошибок и свойствами когерентной информации.

Определение 9.1.2 *Канал Φ называется точно обратимым на состоянии $S = S_A$, если найдется восстанавливающий канал $\mathcal{D}$ из B в A, такой, что*

$$(\mathcal{D} \circ \Phi \otimes \mathrm{Id}_R)[S_{AR}] = S_{AR}.$$

Предложение 9.1.2 *Следующие условия эквивалентны:*

i. канал Φ точно обратим на состоянии S;
ii. $I(R; E) = 0$, т. е. $S_{RE} = S_R \otimes S_E$;
iii. $I_c(S, \Phi) = H(S)$.
iv. канал Φ допускает представление (9.5) на носителе состояния S.

Условие *ii*. означает, что информация не уходит в окружение, т. е. канал является "секретным". Таким образом, точная обратимость канала равносильна его секретности. Условие *iii*. означает, что при обратимой передаче каналом Φ состояния S, когерентная информация $I_c(S, \Phi)$ должна достигать своего максимально возможного значения $H(S)$. В частности, беря хаотическое состояние $S = \bar{S}_A = I_A/d_A$, мы получаем условие $\log d_A = I_c(\bar{S}_A, \Phi)$. Это показывает, что когерентная информация должна быть связана с квантовой пропускной способностью канала Φ, характеризующей максимальную размерность обратимо передаваемых квантовых состояний. В следующем разделе мы увидим, что для асимптотически точной передачи через канал без памяти $\Phi^{\otimes n}$ такая связь действительно имеет место при $n \to \infty$.

Доказательство. i. $\Rightarrow$ *ii.* Вектор $|\psi_{BRE}\rangle = (V \otimes I_R)|\psi_{AR}\rangle$ является вектором чистого состояния S_{BRE} составной системы BRE. Из точной обратимости вытекает

$$(\mathcal{D} \otimes \mathrm{Id}_{RE})[S_{BRE}] = S_{ARE}.$$

Так как состояние S_{AR} – чистое, то согласно следствию 3.1.2 $S_{ARE} = S_{AR} \otimes S_E$. Взяв частичный след по пространству A, получаем

$$S_{RE} = S_R \otimes S_E, \tag{9.7}$$

то есть *ii*.

ii. $\Leftrightarrow$ *iii*. Имеем

$$H(S) - I_c(S, \Phi) = H(A) + H(E) - H(B) \tag{9.8}$$
$$= H(R) + H(E) - H(RE) = I(R; E) \geq 0; \tag{9.9}$$

причем равенство имеет место тогда и только тогда, когда $I(R; E) = 0$, т. е. $S_{RE} = S_R \otimes S_E$.

iii. $\Rightarrow$ *i*. Вектор $|\psi_{BRE}\rangle = (V \otimes I_R)|\psi_{AR}\rangle$ является вектором чистого состояния S_{BRE} составной системы BRE, очищающего состояние S_{RE}. Равенство (9.7) дает другое очищение $|\psi_{AR}\rangle \otimes |\psi_{EE'}\rangle$, где E' – очищающая система для S_E. Выбирая E', мы всегда можем считать, что пространство второго очищения имеет размерность не меньшую, чем первое; следовательно, согласно замечанию после теоремы 3.1.2, найдется изометрия $W : \mathcal{H}_B \to \mathcal{H}_A \otimes \mathcal{H}_{E'}$, такая что

$$(I_{RE} \otimes W)|\psi_{BRE}\rangle = |\psi_{AR}\rangle \otimes |\psi_{EE'}\rangle. \tag{9.10}$$

Беря частичные следы по пространствам E, E' от соответствующих операторов плотности, получаем точную обратимость

$$(\mathrm{Id}_R \otimes \mathcal{D} \circ \Phi)[S_{AR}] = S_{AR}, \tag{9.11}$$

где восстанавливающий канал определяется как

$$\mathcal{D}[S_B] = \mathrm{Tr}_{E'} W S_B W^*. \tag{9.12}$$

iii. $\Rightarrow$ *iv*. $\Rightarrow$ *i*. Далее будет доказано следствие 9.2.1, из утверждения *i*. $\Leftrightarrow$ *iii*. которого с заменой Φ на $\mathcal{D} \circ \Phi$) вытекает, что канал Φ точно обратим на состоянии S тогда и только тогда, когда

$$\mathcal{D} \circ \Phi[\tilde{S}] = \tilde{S}$$

для всех $\tilde{S}$ с supp $\tilde{S} \subset \mathcal{L} \equiv$ supp S, где supp S – носитель оператора плотности S (см. раздел 1.3). Мы можем выразить это же свойство, сказав, что Φ является *точно обратимым на подпространстве* $\mathcal{L}$. Другими словами, подпространство $\mathcal{L}$ является квантовым кодом, исправляющим ошибку Φ (а значит, и все ошибки ассоциированного с ней класса $\mathcal{E}$). Утверждение *iv*., а также импликация *iv*. $\Rightarrow$ *i*. следуют тогда из доказательства теоремы 9.1.1. □

Следующее предложение дает характеризацию в терминах комплементарного канала, которая лежит в основе теоремы о секретной классической пропускной способности в разделе 9.4.1.

Предложение 9.1.3 *Для заданного расширения Стайнспринга канала Φ следующие условия эквивалентны:*

i. канал Φ точно обратим на подпространстве $\mathcal{L}$;

ii. комплементарный канал (7.37) является полностью деполяризующим на $\mathcal{L}$, т.е.

$$\Phi_E[\tilde{S}] = S_E, \tag{9.13}$$

для любого состояния $\tilde{S}$ с supp $\tilde{S} \subset \mathcal{L}$*, где S_E – фиксированное состояние.*

Доказательство. i. $\Rightarrow$ ii. Пусть канал Φ точно обратим, тогда соотношение (9.10) выполнено для некоторой изометрии W и для всех $|\psi_{AR}\rangle$, имеющих тензорное разложение с A-компонентами в $\mathcal{L}$. Здесь $|\psi\rangle_{EE'}$ не зависит от входного вектора $|\psi_{AR}\rangle$, так как иначе правая часть не была бы линейной по $|\psi_{AR}\rangle$. Беря частичный след по RAE' от соответствующего оператора плотности, получаем (9.13).

ii. $\Rightarrow$ i. Имеет место представление Стайнспринга $\Phi_E[\tilde{S}] = \mathrm{Tr}_B\, V\tilde{S}V^*$, где B играет роль окружения для системы E. Пусть (9.13) выполнено и пусть $|\psi_{EE'}\rangle$ – очищение состояния S_E; тогда имеет место представление Стайнспринга для ограничения канала Φ_E на состояния с носителем в $\mathcal{L}$:

$$\Phi_E[\tilde{S}] = \mathrm{Tr}_{AE'}\, V'\tilde{S}V'^*, \tag{9.14}$$

где $V' = I_A \otimes |\psi_{EE'}\rangle$ – изометричный оператор, действующий из $\mathcal{H}_A$ в $\mathcal{H}_{AEE'} = \mathcal{H}_E \otimes \mathcal{H}_{AE'}$ по формуле $V'|\psi_A\rangle = |\psi_A\rangle \otimes |\psi_{EE'}\rangle$. В этом представлении роль окружения играет система AE'. По теореме 6.2.3, найдется частичная изометрия $W : \mathcal{H}_B \to \mathcal{H}_{AE'}$, такая что

$$(W \otimes I_{RE})V = V' = I_A \otimes |\psi_{EE'}\rangle.$$

Следовательно, соотношение (9.10) выполнено для всех векторов $|\psi\rangle_{RA}$ с тензорным разложением, у которого A-компоненты принадлежат $\mathcal{L}$. Расширяя, если необходимо, пространство $\mathcal{H}_{E'}$, мы можем заменить W на изометрию, для которой продолжает выполняться соотношение (9.10). Рассуждая как в доказательстве предыдущего утверждения, возьмем частичный след по REE' от соответствующего оператора плотности и получим точную обратимость на $\mathcal{L}$:

$$\mathcal{D} \circ \Phi[\tilde{S}] = \tilde{S}, \tag{9.15}$$

где восстанавливающий канал определяется, как в формуле (9.12). □

9.2 Точность воспроизведения квантовой информации

В качестве меры точности передачи состояния S каналом Φ можно использовать следовую норму $\| S - \Phi[S] \|_1$. Однако более удобными оказы-

ваются другие меры точности, которые играют ту же роль, что вероятность правильного решения в классическом случае. Дадим их описание и рассмотрим взаимосвязи между ними.

9.2.1 Точность воспроизведения чистых состояний

Лемма 9.2.1 *Пусть ψ – единичный вектор и S – произвольное состояние. Имеют место неравенства*

$$2\,[1 - \langle\psi|\,S\,|\psi\rangle] \leqslant \|\,|\psi\rangle\langle\psi| - S\|_{\,1} \leqslant 2\sqrt{1 - \langle\psi|\,S\,|\psi\rangle}. \tag{9.16}$$

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

$$\|\,|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|\,\|_{\,1} = 2\sqrt{1 - |\,\langle\psi|\,\phi\rangle\,|^{\,2}}. \tag{9.17}$$

Доказательство. Согласно (1.19), имеем

$$\|\,|\psi\rangle\langle\psi| - S\|_{\,1} = \max_U |\,\mathrm{Tr}(|\psi\rangle\langle\psi| - S)U|,$$

где максимум берется по всем унитарным операторам U. Полагая $U = 2|\psi\rangle\langle\psi| - I$, получаем первое неравенство.

Равенство (9.17) получается аналогично соотношению (2.46)). Для вычисления следовой нормы достаточно найти собственные значения оператора $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$, который имеет ранг 2 (см. задачу 2.3.2).

Пусть теперь S – произвольный оператор плотности; рассмотрим его спектральное разложение $S = \sum \lambda_i S_i$. Используя выпуклость нормы, вогнутость функции $\sqrt{\ }$ и соотношение (9.17), получаем

$$\|\,|\psi\rangle\langle\psi| - S\|_{\,1} \leq \sum_i \lambda_i \|\,|\psi\rangle\langle\psi| - S_i\|_{\,1}$$

$$= 2\sum_i \lambda_i \sqrt{1 - \langle\psi|\,S_i\,|\psi\rangle} \leq 2\sqrt{1 - \langle\psi|\,S\,|\psi\rangle}.$$

□

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

$$F(|\psi\rangle\langle\psi|, S) = \langle\psi|\,S\,|\psi\rangle \tag{9.18}$$

называется *точностью воспроизведения* (состояния $|\psi\rangle\langle\psi|$ состоянием S). Очевидно, что $F \leq 1$, причем равенство имеет место тогда и только тогда, когда $S = |\psi\rangle\langle\psi|$. Заметим, что рассматривая проблему сжатия

квантовых данных (раздел 5.5) мы уже фактически использовали точность воспроизведения (9.18). Там состояние $|\,\psi_i\rangle\langle\psi_i|$ появлялось с вероятностью p_i, и среднее значение точности воспроизведения определялось как $\bar{F} = \sum_i p_i \langle\psi_i|\, S_i\,|\psi_i\rangle$.

В связи с понятием кода, исправляющего ошибки, естественно определить точность воспроизведения следующим образом. Пусть задано подпространство (квантовый код) $\mathcal{L} \subset \mathcal{H}$ и канал Φ в $\mathcal{H}$, для которого входное и выходное пространства совпадают: $\mathcal{H}_B = \mathcal{H}_A = \mathcal{H}$). *Точность воспроизведения подпространства* $\mathcal{L}$ определяется как

$$\begin{aligned} F_s(\mathcal{L}, \Phi) &= \min_{\psi\in\mathcal{L}, \|\psi\|=1} F(|\psi\rangle\langle\psi|, \Phi[|\psi\rangle\langle\psi|]|) \\ &= \min_{\psi\in\mathcal{L}, \|\psi\|=1} \langle\psi|\,\Phi[\,|\,\psi\rangle\langle\psi|\,]\,|\,\psi\rangle. \end{aligned}$$

Согласно лемме 9.2.1,

$$2[1 - F_s(\mathcal{L}, \Phi)] \le \max_{\psi\in\mathcal{L}} \|\,|\,\psi\rangle\langle\psi| - \Phi[|\,\psi\rangle\langle\psi|]\,\|_1 \le 2\sqrt{1 - F_s(\mathcal{L}, \Phi)}. \quad (9.19)$$

Другой важной величиной является точность воспроизведения сцепленности. Рассмотрим очищение $|\,\psi_{AR}\rangle\langle\psi_{AR}|$ состояния $S = S_A$ в гильбертовом пространстве $\mathcal{H}_A \otimes \mathcal{H}_R$. Точность воспроизведения этого чистого состояния под действием тривиального расширения канала Φ равна

$$F_e(S, \Phi) = \langle\psi_{AR}|\,(\Phi \otimes \mathrm{Id}_R)[|\,\psi_{AR}\rangle\langle\psi_{AR}|\,]\,|\,\psi_{AR}\rangle,$$

и называется textitточностью воспроизведения сцепленности. Дадим удобное выражение для $F_e(S, \Phi)$, которое также показывает независимость этой величины от выбора очищения исходного состояния.

Пусть канал имеет представление Крауса

$$\Phi[S] = \sum_i V_i S V_i^*, \quad (9.20)$$

тогда

$$F_e(S, \Phi) = \sum_i |\,\mathrm{Tr} V_i S|\,^2. \quad (9.21)$$

В самом деле,

$$\begin{aligned} &\langle\psi_{AR}|\,\textstyle\sum_i (V_i \otimes I_R)\,|\,\psi_{AR}\rangle\langle\psi_{AR}|\,(V_i \otimes I_R)^*\,|\,\psi_{AR}\rangle \\ &= \textstyle\sum_i |\,\langle\psi_{AR}|\,(V_i \otimes I_R)\,|\,\psi_{AR}\rangle|\,^2 = \sum_i |\,\mathrm{Tr}\, V_i S|\,^2. \end{aligned}$$

Заметим, что для чистого состояния $S = |\psi\rangle\langle\psi|$,

$$F_e(S, \Phi) = \langle\psi|\Phi[|\psi\rangle\langle\psi|]|\psi\rangle = F(|\psi\rangle\langle\psi|, \Phi[|\psi\rangle\langle\psi|]).$$

Задача 9.2.1 *Докажите, что $F_e(S, \Phi)$ – выпуклая функция состояния S. Указание: Согласно (9.21), $F_e(S, \Phi)$ является суммой квадратов аффинных функций.*

9.2.2 Соотношения между мерами точности воспроизведения

Лемма 9.2.2 *Для произвольного состояния S*

$$1 - F_e(S, \Phi) \leq 4\sqrt{1 - F_s(\mathrm{supp}S, \Phi)}.$$

Доказательство. Имеем

$$1 - F_e(S, \Phi) = 1 - \langle\psi_{AR}|\,(\Phi \otimes \mathrm{Id}_R)(|\,\psi_{AR}\rangle\langle\psi_{AR}|\,)|\,\psi_{AR}\rangle$$
$$= \langle\psi_{AR}|(\mathrm{Id}_A - \Phi) \otimes \mathrm{Id}_R)[|\,\psi_{AR}\rangle\langle\psi_{AR}|]|\,\psi_{AR}\rangle.$$

Представляя $|\,\psi_{AR}\rangle = \sum_j |\,\psi_j\rangle \otimes |\,e_j\rangle$, где $\{|\,e_j\rangle\}$ – ортонормированный базис в $\mathcal{H}_R$, $\sum_j \|\psi_j\|^2 = 1$, $\psi_j \in \mathrm{supp}S$, получаем равносильное соотношение

$$\sum_{jk}\langle\psi_j|(\mathrm{Id} - \Phi)[|\,\psi_j\rangle\langle\psi_k|]|\,\psi_k\rangle = \sum_{jk}\mathrm{Tr}\,|\psi_k\rangle\langle\psi_j|\,(|\psi_j\rangle\langle\psi_k| - \Phi[|\psi_j\rangle\langle\psi_k|])\,.$$

Используя неравенство (1.18) и тот факт, что и следовая, и операторная нормы оператора $|\psi_j\rangle\langle\psi_k|$ равны $\|\psi_j\|\|\psi_k\|$, получаем, что это выражение не превосходит

$$\max_T \frac{\|T - \Phi[T]\|_1}{\|T\|_1},$$

где максимум берется по всем операторам $T \neq 0$, действующим в $\mathcal{L} = \mathrm{supp}S$. Рассматривая разложение $T = T_1 + iT_2$, где $T_1^* = T_1, T_2^* = T_2$ – эрмитовы операторы, также действующие в $\mathcal{L}$, принимая во внимание, что $\|T_{1,2}\|_1 \leq \|T\|_1$, и используя неравенство треугольника, получаем, что это выражение не превосходит величину

$$2\max_{T^*=T}\frac{\|T - \Phi[T]\|_1}{\|T\|_1} = 2\max_{S:\mathrm{supp}S\in\mathcal{L}}\|S - \Phi[S]\|_1. \tag{9.22}$$

Здесь неравенство $\geq$ получается при переходе к положительным T. С другой стороны,

$$\max_{T^*=T}\frac{\|T - \Phi[T]\|_1}{\|T\|_1} \leq \|S_+ - \Phi[S_+]\|_1\frac{p_+}{p_+ + p_-} + \|S_- - \Phi[S_-]\|_1\frac{p_-}{p_+ + p_-}, \tag{9.23}$$

где использованы обозначения $p_\pm = \mathrm{Tr}\,T_\pm; S_\pm = p_\pm^{-1}T_\pm$. Правая часть этого неравенства меньше или равна правой части в соотношении (9.22), которое таким образом установлено. Далее,

$$\max_{S:\mathrm{supp}S\in\mathcal{L}}\|S - \Phi[S]\|_1 = \max_{\psi:\psi\in\mathcal{L},\|\psi\|=1}\|\,|\,\psi\rangle\langle\psi| - \Phi[|\,\psi\rangle\langle\psi|]\|_1,$$

поскольку максимум непрерывной выпуклой функции на выпуклом множестве достигается в его крайней точке. Используя второе неравенство (9.19), получаем, что оцениваемое выражение не превосходит $4\sqrt{1 - F_s(\mathcal{L}, \Phi)}$.

□

Задача 9.2.2 *Используя свойство выпуклости из задачи 9.2.1, докажите неравенство*

$$1 - F_s(\mathcal{L}, \Phi) \leq 1 - \min_{S:\mathrm{supp}S\subset\mathcal{L}} F_e(S, \Phi). \tag{9.24}$$

Пусть теперь $\mathcal{L} = \mathcal{H}$, тогда

$$\max_T \frac{\|T - \Phi[T]\|_1}{\|T\|_1} = \|\mathrm{Id} - \Phi\|.$$

Из доказательства леммы 9.2.2 следует справедливость цепочки неравенств:

$$1 - \min_S F_e(S, \Phi) \leq \|\mathrm{Id} - \Phi\| \leq 2 \max_{\|\psi\|=1} \|\,|\psi\rangle\langle\psi| - \Phi[|\psi\rangle\langle\psi|]\|_1 \leq 2\|\mathrm{Id} - \Phi\|. \tag{9.25}$$

Вместе с соотношениями (9.24) это означает, что отклонение канала Φ от идеального канала Id эквивалентным образом описывается одной из величин

$$\|\mathrm{Id} - \Phi\|; \quad \max_S \|S - \Phi[S]\|_1; \quad 1 - F_s(\mathcal{H}, \Phi); \quad 1 - \min_S F_e(S, \Phi).$$

При рассмотрении квантовой пропускной способности канала нам понадобится очевидное следствие леммы 9.2.2:

$$1 - F_e(\bar{S}, \Phi) \leq 4\sqrt{1 - F_s(\mathcal{H}, \Phi)}, \tag{9.26}$$

где $\bar{S}$ – хаотическое состояние в $\mathcal{H}$.

Поучительно рассмотреть случай "идеальной" точности воспроизведения.

Следствие 9.2.1 *Следующие условия эквивалентны:*

i. $F_e(S, \Phi) = 1$*;*
ii. $F_s(\mathrm{supp}S, \Phi) = 1$*;*
iii. $\Phi[\tilde{S}] = \tilde{S}$ *для произвольного состояния* $\tilde{S}$ *с* $\mathrm{supp}\tilde{S} \subset \mathrm{supp}S$*;*
iv. $F_e(\tilde{S}, \Phi) = 1$ *для произвольного состояния* $\tilde{S}$ *с* $\mathrm{supp}\tilde{S} \subset \mathrm{supp}S$.

Доказательство. Докажем $i. \Rightarrow ii.$ Пусть $|\psi\rangle \in \mathrm{supp}S$, тогда

$$S = p|\psi\rangle\langle\psi| + (1-p)S',$$

где $p > 0$, а S' – оператор плотности. Согласно задаче 9.2.1, функция $S \to F_e(S, \Phi)$ выпукла, поэтому

$$pF_e(|\psi\rangle\langle\psi|, \Phi) + (1-p)F_e(S', \Phi) \geq F_e(S, \Phi) = 1,$$

и, значит, $F_e(|\psi\rangle\langle\psi|, \Phi) = 1$. Но отсюда следует, что

$$\langle\psi|\,\Phi[\,|\,\psi\rangle\langle\psi|\,]\,|\,\psi\rangle = 1$$

для всех $|\psi\rangle \in \mathrm{supp}S$ и $F_s(\mathrm{supp}S, \Phi) = 1$.

Из леммы 9.2.2 вытекает утверждение *ii.* $\Rightarrow$ *i.* Кроме того, так как из *ii.* очевидным образом следует аналогичное условие с supp $\tilde{S}$ вместо supp S, то мы также установили и справедливость свойства *iv.*, которое формально является усилением свойства *i.*

По определению F_s, свойство *ii.* равносильно *iii.* для чистых, а следовательно, и для смешанных состояний $\tilde{S}$. □

9.2.3 Точность воспроизведения и расстояние Бюреса

В этом разделе мы обсудим понятие точности воспроизведения для двух произвольных состояний. Этот материал понадобится нам лишь в разделе 9.4.4.

Для чистых состояний $S = |\phi\rangle\langle\phi|, T = |\psi\rangle\langle\psi|$, точность воспроизведения равна $F(T, S) = |\langle\phi|\psi\rangle|^2$, при этом связь между точностью воспроизведения и следовой нормой разности S и T выражается формулой (9.17). Пусть теперь S, T – произвольные операторы плотности.

Определение 9.2.1 *Точность воспроизведения определяется соотношением*

$$F(T, S) = \max |\langle\psi_T|\psi_S\rangle|^2, \tag{9.27}$$

где максимум берется по всем возможным очищениям ψ_S, ψ_T состояний S, T в одном и том же пространстве $\mathcal{H} \otimes \mathcal{H}'$.)

Используя теорему 3.1.3, видим, что

$$F(T, S) = \max_W |\langle\psi_T|(I \otimes W)\psi_S\rangle|^2, \tag{9.28}$$

где ψ_S, ψ_T – фиксированные очищения состояний S, T в некотором гильбертовом пространстве $\mathcal{H} \otimes \mathcal{H}'$, а максимум берется по всем унитарным операторам W из $\mathcal{H}'$. В частности, рассматривая очищения $\sqrt{S}, \sqrt{T}$ в $L^2(\mathcal{H}) \simeq \mathcal{H} \otimes \mathcal{H}^*$ получаем

$$F(T, S) = \max_W \left|\mathrm{Tr}\sqrt{S}\sqrt{T}W\right|^2,$$

где максимум берется по всем унитарным операторам W в $\mathcal{H}$. Используя полярное разложение оператора $\sqrt{S}\sqrt{T} = U|\sqrt{S}\sqrt{T}|$, можно видеть, что максимум достигается при $W = U^*$ и равен

$$F(T, S) = \left(\mathrm{Tr}\,|\sqrt{S}\sqrt{T}|\right)^2 = \left\|\sqrt{S}\sqrt{T}\right\|_1^2.$$

Эта формула переходит в (9.18), если $T = |\psi\rangle\langle\psi|$. Из этого рассуждения вытекает также, что в определении (9.27) можно ограничиться максимизацией по очищениям только одного состояния, при том, что очищение другого состояния фиксировано.

Задача 9.2.3 *Покажите, что $F(T,S) \leq 1$, причем равенство имеет место тогда и только тогда, когда $T = S$.*

С понятием точности воспроизведения (9.27) связана метрика на множестве всех состояний, называемая *расстоянием Бюреса*, которая определяется как

$$\beta(T,S) = \min \|\psi_T - \psi_S\|, \tag{9.29}$$

где минимум берется по всем возможным очищениям ψ_S, ψ_T состояний S, T. Рассматривая квадрат нормы и используя определение (9.27), получаем

$$\beta(T,S) = \sqrt{2\left(1 - \left\|\sqrt{S}\sqrt{T}\right\|_1\right)} = \sqrt{2\left(1 - \sqrt{F(T,S)}\right)}. \tag{9.30}$$

В частности, для чистых состояний $T = |\psi\rangle\langle\psi|, S = |\phi\rangle\langle\phi|$

$$\beta(T,S) = \sqrt{2\left(1 - |\langle\psi|\phi\rangle|\right)}.$$

Отсюда также видно, что как и при определении точности воспроизведения, в (9.29) можно брать максимум по всевозможным очищениям лишь одного состояния.

Неравенство треугольника для $\beta(T,S)$ следует из определения (9.29) и соответствующего неравенства для нормы. Для чистых состояний оно превращается в

$$\sqrt{1 - |\langle\varphi|\psi\rangle|} \leq \sqrt{1 - |\langle\varphi|\chi\rangle|} + \sqrt{1 - |\langle\chi|\psi\rangle|}, \tag{9.31}$$

где φ, χ, ψ – произвольные единичные векторы.

Лемма 9.2.3 *Для любых двух операторов плотности S_1, S_2 в $\mathcal{H}$*

$$\beta(S_1,S_2)^2 \leq \|S_1 - S_2\|_1 \leq 2\beta(S_1,S_2), \tag{9.32}$$

Рассматривая первое неравенство вместе с (9.30) и (9.28), получаем

Следствие 9.2.2 *Для любых двух операторов плотности S_1, S_2 в $\mathcal{H}$ и их очищений ψ_{S_1}, ψ_{S_2} в $\mathcal{H} \otimes \mathcal{H}'$ имеет место*

$$\|S_1 - S_2\|_1 \geq 2\left(1 - \max_W |\langle\psi_{S_1}|(I \otimes W)\psi_{S_2}\rangle|\right), \tag{9.33}$$

где максимум берется по всем унитарным операторам W в $\mathcal{H}'$.

Доказательство леммы 9.2.3. Так как $\operatorname{Tr}\sqrt{S_1}\sqrt{S_2} \leq \operatorname{Tr}|\sqrt{S_1}\sqrt{S_2}|$, то для доказательства первого неравенства (9.32) достаточно доказать, что

$$2\left(1 - \operatorname{Tr}\sqrt{S_1}\sqrt{S_2}\right) \leq \|S_1 - S_2\|_1.$$

Это вытекает из более общего факта

$$\operatorname{Tr}\left(\sqrt{S_1}-\sqrt{S_2}\right)^2 \le \|S_1-S_2\|_1\,,$$

справедливого для произвольных положительных S_1, S_2. Докажем последнее неравенство. Для эрмитова оператора K мы обозначаем через $1_+(K)$ (соответственно $1_-(K)$) спектральный проектор, соответствующий положительным (соответственно неотрицательным) собственным значениям. Тогда $\sigma(K) = 1_+(K) - 1_-(K)$ является унитарным эрмитовым оператором, причем $\sigma(K)K = K\sigma(K) = |K|$. Вводя операторы $K = \sqrt{S_1}-\sqrt{S_2}, L = \sqrt{S_1}+\sqrt{S_2}$, получаем $S_1 - S_2 = \frac{1}{2}(KL+LK) \equiv K\circ L$; следовательно,

$$\|S_1-S_2\|_1 \ge |\operatorname{Tr}\sigma(K)K\circ L| = \operatorname{Tr}|K|L$$

$$= \operatorname{Tr}|K|\left(1_+(K)L1_+(K) + 1_-(K)L1_-(K)\right).$$

Так как $L \ge K \ge -L$, то $1_+(K)L1_+(K) \ge 1_+(K)K$ и $1_-(K)L1_-(K) \ge -1_-(K)K$. Таким образом, $(1_+(K)L1_+(K) + 1_-(K)L1_-(K)) \ge |K|$, и в итоге получаем

$$\|S_1-S_2\|_1 \ge \operatorname{Tr}|K|^2 = \operatorname{Tr}\left(\sqrt{S_1}-\sqrt{S_2}\right)^2.$$

Для доказательства второго неравенства (9.32) рассмотрим представления $S_1 = T_1^*T_1, S_2 = T_2^*T_2$, где T_1, T_2 – некоторые (не обязательно эрмитовы) операторы. Вводя обозначения $K = T_1 - T_2, L = T_1 + T_2$, получаем

$$\|S_1-S_2\|_1 = \operatorname{Tr}\sigma(S_1-S_2)(S_1-S_2) = \frac{1}{2}\left[\operatorname{Tr}\sigma(S_1-S_2)K^*L + \operatorname{Tr}\sigma(S_1-S_2)L^*K\right],$$

Используя неравенство Коши-Буняковского для следа, имеем

$$|\operatorname{Tr}\sigma(K\circ L)K^*L| \le \left[\operatorname{Tr}\sigma(S_1-S_2)^2K^*K\cdot\operatorname{Tr}L^*L\right]^{\frac{1}{2}} = \left[\operatorname{Tr}K^*K\cdot\operatorname{Tr}L^*L\right]^{\frac{1}{2}};$$

аналогичная оценка имеет место и для второго слагаемого. Тогда $\|S_1-S_2\|_1 \le \left[\operatorname{Tr}K^*K\cdot\operatorname{Tr}L^*L\right]^{\frac{1}{2}}$. Используя нормировку операторов S_1, S_2, получаем

$$\operatorname{Tr}K^*K = 2\left(1-\Re\operatorname{Tr}T_1^*T_2\right), \quad \operatorname{Tr}L^*L = 2\left(1+\Re\operatorname{Tr}T_1^*T_2\right).$$

Полагая $T_1 = \sqrt{S_1}, T_2 = U\sqrt{S_2}$, где U – унитарный оператор из полярного разложения оператора $\sqrt{S_1}\sqrt{S_2}$, получаем $\Re\operatorname{Tr}T_1^*T_2 = \Re\operatorname{Tr}T_2^*T_1 = \operatorname{Tr}\left|\sqrt{S_1}\sqrt{S_2}\right|$. Таким образом

$$\operatorname{Tr}K^*K = 2\left(1-\left\|\sqrt{S_1}\sqrt{S_2}\right\|_1\right) = \beta(S_1,S_2)^2,$$

$$\operatorname{Tr}L^*L = 2\left(1+\left\|\sqrt{S_1}\sqrt{S_2}\right\|_1\right) \le 4,$$

откуда следует второе неравенство (9.32). □

9.3 Квантовая пропускная способность

9.3.1 Достижимые скорости

Пусть $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$ – некоторый квантовый канал. Рассмотрим составной канал $\Phi^{\otimes n} : \mathfrak{T}(\mathcal{H}_A^{\otimes n}) \to \mathfrak{T}(\mathcal{H}_B^{\otimes n})$.

Определение 9.3.1 *Величина $R \geq 0$ называется достижимой скоростью (для передачи квантовой информации), если существуют последовательность пространств $\mathcal{H}^{(n)}$, такая, что*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \dim \mathcal{H}^{(n)} = R,$$

и последовательности каналов $\mathcal{E}^{(n)} : \mathfrak{T}(\mathcal{H}^{(n)}) \to \mathfrak{T}(\mathcal{H}_A^{\otimes n})$ (кодирований) и $\mathcal{D}^{(n)} : \mathfrak{T}(\mathcal{H}_B^{\otimes n}) \to \mathfrak{T}(\mathcal{H}^{(n)})$ (декодирований), такие что

$$\lim_{n\to\infty} F_s(\mathcal{H}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) = 1. \tag{9.34}$$

Точная верхняя грань множества достижимых скоростей обозначается $Q(\Phi)$ и называется квантовой пропускной способностью канала Φ.

При доказательстве классической теоремы кодирования было удобно использовать среднюю вероятность ошибки вместо максимальной; равносильность соответствующих критериев была доказана в лемме 4.4.1. В квантовом случае роль средней вероятности ошибки играет величина $1 - F_e\left(\bar{S}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}\right)$, где $\bar{S}^{(n)} = \frac{1}{d_n} I_{\mathcal{H}^{(n)}}$ – хаотическое состояние в $\mathcal{H}^{(n)}$, тогда как аналогом максимальной ошибки является $1 - F_s\left(\mathcal{H}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}\right)$. В самом деле, из $F_s\left(\mathcal{H}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}\right) \to 1$ следует $F_e\left(\bar{S}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}\right) \to 1$ (согласно лемме 9.2.2). В обратном направлении справедлив следующий аналог леммы 4.4.1:

Лемма 9.3.1 *Пусть $\mathcal{H}$ – гильбертово пространство размерности $2d$, Φ – канал в $\mathcal{H}$. Тогда найдется подпространство $\mathcal{H}'$ размерности d, такое что*

$$1 - F_s\left(\mathcal{H}', \mathcal{D}' \circ \Phi\right) \leq 2(1 - F_e\left(\bar{S}, \Phi\right)),$$

где $\bar{S} = \frac{1}{2d} I_{\mathcal{H}}$, а $\mathcal{D}'[S] = P'SP' + \frac{P'}{d} \mathrm{Tr}(I - P')S$ для всех $S \in \mathfrak{S}(\mathcal{H})$, где P' – проектор на $\mathcal{H}'$.

Доказательство. На единичном шаре в $\mathcal{H}$ рассмотрим непрерывную функцию $|\psi\rangle \to f(\psi) = \langle\psi|\Phi\left[|\psi\rangle\langle\psi|\right]|\psi\rangle = F_e\left(|\psi\rangle\langle\psi|, \Phi\right)$. Пусть $|\psi_1\rangle$ – вектор, минимизирующий $f(\psi)$. Определим ортонормированный базис $\{|\psi_j\rangle; j = 1, \ldots, 2d\}$ в $\mathcal{H} \equiv \mathcal{H}_0$ с помощью следующей рекуррентной процедуры: $|\psi_{j+1}\rangle$ есть вектор в подпространстве $\mathcal{H}_j = \{|\psi_1\rangle, \ldots, \psi_j\rangle\}^{\perp}$, минимизирующий $f(\psi)$. Тогда $\mathcal{H}_j \supset \mathcal{H}_{j+1}$ и $\dim \mathcal{H}_j = 2d - j$. Используя выпуклость $F_e\left(S, \Phi\right)$ (см. задачу 9.2.1), получаем

$$
\begin{aligned}
1 - F_e\left(\bar{S}, \Phi\right) &\geq \frac{1}{2d} \sum_{j=1}^{2d} \left(1 - F_e\left(|\psi_j\rangle\langle\psi_j|, \Phi\right)\right) \\
&\geq \frac{1}{2d} \sum_{j=d+1}^{2d} \left(1 - f(\psi_j)\right) \\
&\geq \frac{1}{2}\left(1 - \min_{|\psi\rangle \in \mathcal{H}_d} f(\psi)\right) \\
&\geq \frac{1}{2}\left(1 - F_s\left(\mathcal{H}', \mathcal{D}' \circ \Phi\right)\right),
\end{aligned}
$$

где $\mathcal{H}' = \mathcal{H}_d$. □

Из доказанной леммы вместе с неравенством (9.26) следует, что в определении достижимых скоростей мы можем заменить требование (9.34) на

$$\lim_{n\to\infty} F_e\left(\bar{S}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}\right) = 1. \tag{9.35}$$

Теперь покажем, что в определении квантовой пропускной способности достаточно ограничиться *изометричными* кодированиями, т. е. кодированиями вида

$$\mathcal{E}[S] = \mathrm{Ad}V[S] = VSV^*, \tag{9.36}$$

где V – изометричное отображение пространства $\mathcal{H}^{(n)}$ во входное пространство канала $\mathcal{H}^{\otimes n}$. Мы докажем, что если существует кодирование общего вида, при котором достигается высокая точность воспроизведения данного состояния, то найдется и изометричное кодирование с аналогичным свойством. Чтобы пояснить это на интуитивном уровне, рассмотрим случай точного воспроизведения. Если композиция кодирование-канал-декодирование точно воспроизводит некоторое состояние S, то кодирующий канал $\mathcal{E}$ является точно обратимым на S, причем восстанавливающим каналом является композиция канал-декодирование. Тогда, согласно предложению 9.1.2 канал $\mathcal{E}$ на $\mathrm{supp}S$ представляется в виде (9.5) как выпуклая комбинация изометричных кодирований $\mathcal{U}_j = \mathrm{Ad}U_j$. Следовательно, каждый из каналов $\mathcal{U}_j$ является точно обратимым на S с тем же самым восстанавливающим каналом.

Лемма 9.3.2 *Пусть S – состояние в гильбертовом пространстве $\mathcal{L}$, $\mathcal{L} = \mathrm{supp}S$, пусть $\mathcal{E}$ – вполне положительное отображение из $\mathfrak{T}(\mathcal{L})$ в $\mathfrak{T}(\mathcal{H})$, для которого $\mathrm{Tr}\,\mathcal{E}[S] = 1$ и $\mathcal{A}$ – канал из $\mathfrak{T}(\mathcal{H})$ в $\mathfrak{T}(\mathcal{L})$. Предположим, что $\dim \mathcal{L} \leq \dim \mathcal{H}$, тогда найдется изометрия V из $\mathcal{L}$ в $\mathcal{H}$, такая что*

$$F_e(S, \mathcal{A} \circ \mathrm{Ad}V) > F_e(S, \mathcal{A} \circ \mathcal{E})^2. \tag{9.37}$$

Эта лемма будет использована в ситуации, когда $\mathcal{E}$ – кодирующее отображение, а $\mathcal{A}$ – композиция канал-декодирование. Для доказательства потребуется следующее обобщение полярного разложения: если X – оператор из $\mathcal{L}$ в $\mathcal{H}$, $\dim \mathcal{L} \le \dim \mathcal{H}$, то $X = V|X|$, где $|X| = \sqrt{X^*X}$ – положительный оператор в $\mathcal{L}$, а V – изометрия из $\mathcal{L}$ в $\mathcal{H}$. Отсюда вытекает следующее представление для произвольной квадратной матрицы X: $X = VDU$, где D – диагональная матрица с неотрицательными элементами, а V, U – унитарные матрицы.

Доказательство. Пусть $A_i : \mathcal{H} \to \mathcal{L}$ и $E_j : \mathcal{L} \to \mathcal{H}$ – компоненты разложения Крауса для отображений $\mathcal{A}$ и $\mathcal{E}$, соответственно. Обозначим X матрицу с элементами $X_{ij} = \operatorname{Tr} A_i E_j S$. Тогда

$$F_e(S, \mathcal{A} \circ \mathcal{E}) = \sum_{ij} |X_{ij}|^2.$$

Добавляя, если необходимо, нулевые компоненты в разложения Крауса, можно считать, что X – квадратная матрица. Из представления $X = VDU$ следует, что преобразуя разложения Крауса для $\mathcal{A}$ и $\mathcal{E}$, можно добиться, чтобы матрица X стала диагональной. Тогда $F_e(S, \mathcal{A} \circ \mathcal{E}) = \sum_k \operatorname{Tr}(A_k E_k S)^2$. Обозначим $\lambda_k = \operatorname{Tr} S E_k^* E_k$, и ограничимся рассмотрением $\lambda_k > 0$. Тогда $\sum_k \lambda_k (\operatorname{Tr}(A_k E_k S)^2/\lambda_k) = F_e(S, \mathcal{A} \circ \mathcal{E})$ и $\sum_k \lambda_k = 1$, так что найдется k, такое что $\operatorname{Tr}(A_k E_k S)^2/\lambda_k \ge F_e(S, \mathcal{A} \circ \mathcal{E})$. Обозначая $E = E_k/\sqrt{\lambda_k} : \mathcal{L} \to \mathcal{H}$, $A = A_k : \mathcal{H} \to \mathcal{L}$, имеем $A^*A \le I_{\mathcal{H}}$, $\operatorname{Tr} SE^*E = 1$ и

$$F_e(S, \mathcal{A} \circ \mathcal{E}) \le F_e(S, \operatorname{Ad}A \circ \operatorname{Ad}E) = |\operatorname{Tr} AES|^2.$$

Пусть $A^* = V|A^*|$ – полярное разложение оператора $A^* : \mathcal{L} \to \mathcal{H}$, где $V : \mathcal{L} \to \mathcal{H}$ – изометрия. Согласно некоммутативному неравенству Коши-Буняковского (2.25)

$$|\operatorname{Tr} AES|^2 = |\operatorname{Tr} SAE|^2 \le \operatorname{Tr} SAA^* \operatorname{Tr} SE^*E = \operatorname{Tr} S|A^*|^2. \tag{9.38}$$

Поскольку $A^*A \le I_{\mathcal{H}}$, то также $AA^* \le I_{\mathcal{L}}$, следовательно $|A^*|^2 \le |A^*|$ и поэтому

$$\operatorname{Tr} S|A^*|^2 \le \operatorname{Tr} S|A^*| = \operatorname{Tr} AVS.$$

Тогда

$$|\operatorname{Tr} AVS|^2 \le \sum_k |\operatorname{Tr} A_k VS|^2 = F_e(S, \mathcal{A} \circ \operatorname{Ad}V).$$

Полученная цепочка неравенств влечет (9.37). □

Пусть теперь даны последовательности пространств $\mathcal{H}^{(n)}$, кодирований $\mathcal{E}^{(n)}$ и декодирований $\mathcal{D}^{(n)}$, для которых выполнено (9.35), тогда согласно лемме 9.3.2 найдется последовательность изометричных кодирований $\mathcal{V}^{(n)} = \operatorname{Ad}V^{(n)}$, такая что

$$\lim_{n\to\infty} F_e(\bar{S}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{V}^{(n)}) = 1. \tag{9.39}$$

Отсюда следует, что в определении квантовой пропускной способности можно ограничиться изометричными кодированиями.

9.3.2 Квантовая пропускная способность и когерентная информация

Следующий фундаментальный результат является теоремой кодирования для квантовой пропускной способности и устанавливает связь квантовой пропускной способности с когерентной информацией.

Теорема 9.3.3

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} \max_S I_c(S, \Phi^{\otimes n}). \tag{9.40}$$

Существование предела может быть доказано так же, как и в случае классической пропускной способности – с помощью свойства субаддитивности последовательности $\bar{Q}_n = \max_S I_c(S, \Phi^{\otimes n})$. Простым следствием из (9.40) является неравенство, связывающее квантовую и классическую пропускные способности

$$Q(\Phi) \le C(\Phi), \tag{9.41}$$

которое доказывается так же, как и (8.57). Действительно, выбирая S в виде смеси произвольных чистых состояний S_j с вероятностями p_j, и применяя (7.42), получаем

$$I_c(S, \Phi) \le H\left(\sum_j p_j \Phi[S_j]\right) - \sum_j p_j H(\Phi[S_j]) = \chi\left(\{p_j\}, \{\Phi[S_j]\}\right).$$

Взяв максимум, получаем $\max_S I_c(S, \Phi) \le C_\chi(\Phi)$. Применение этого неравенства к $\Phi^{\otimes n}$ дает (9.41).

Доказательство неравенства $\le$. Обозначим $\bar{Q}(\Phi)$ правую часть соотношения (9.40). Более простой частью теоремы является доказательство неравенства $Q(\Phi) \le \overline{Q}(\Phi)$. Пусть $\mathcal{H}^{(n)}$ – входное пространство размерности $d_n = \dim \mathcal{H}^{(n)} = 2^{n(R+o(1))}$ и пусть $\mathcal{E}^{(n)}, \mathcal{D}^{(n)}$ – кодирующие и декодирующие каналы, такие что $1 - F_e(\bar{S}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) \le \varepsilon$, где $\bar{S}^{(n)} = \frac{1}{d_n} I_{\mathcal{H}^{(n)}}$ – хаотическое состояние в $\mathcal{H}^{(n)}$ с энтропией $H(\bar{S}^{(n)}) = \log d_n = nR$. Здесь мы используем требование (9.35) в определении достижимых скоростей. Согласно лемме 9.3.2, можно считать, что $\mathcal{E}^{(n)}$ – *изометричные* кодирования. Так как

$$F_e(\bar{S}^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)})$$

$$= \langle \Omega^{(n)}| \left(\mathrm{Id}_{\mathcal{H}^{(n)}} \otimes \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}\right) \left[|\Omega^{(n)}\rangle\langle\Omega^{(n)}|\right] |\Omega^{(n)}\rangle,$$

где

$$|\Omega^{(n)}\rangle = \frac{1}{\sqrt{d_n}} \sum_{m=1}^{d_n} |m\rangle \otimes |m\rangle \tag{9.42}$$

– максимально сцепленный вектор в пространстве $\mathcal{H}^{(n)} \otimes \mathcal{H}^{(n)}$, то из леммы 9.2.1 следует, что для максимально сцепленного состояния $|\Omega^{(n)}\rangle\langle\Omega^{(n)}|$ в $\mathcal{H}^{(n)} \otimes \mathcal{H}^{(n)}$ выполняется

$$\left\| |\Omega^{(n)}\rangle\langle\Omega^{(n)}| - (\mathrm{Id}_{\mathcal{H}^{(n)}} \otimes \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) \left[|\Omega^{(n)}\rangle\langle\Omega^{(n)}| \right] \right\|_1 \leq 2\sqrt{\varepsilon}. \tag{9.43}$$

Это означает, что скорость R достижима для асимптотически безошибочной передачи хаотического состояния в $\mathcal{H}^{(n)}$.

Обозначая $S^{(n)} = \mathcal{E}^{(n)} \left[\bar{S}^{(n)} \right]$ и используя цепное правило (9.6) для когерентной информации, получаем

$$I_c(S^{(n)}, \Phi^{\otimes n}) \geq I_c(S^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n}) = H(S_{B'}) - H(S_{B'R'}), \tag{9.44}$$

где $S_{B'R'} = \left(\mathrm{Id}_{\mathcal{H}^{(n)}} \otimes (\mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) \right) \left[|\Omega^{(n)}\rangle\langle\Omega^{(n)}| \right]$, так что $S_{B'} = (\mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) \left[\bar{S}^{(n)} \right]$. Тот факт, что $H(S_{B'R'}) = H(S^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n})$ следует из того, что в силу изометричности кодирования $\mathcal{E}^{(n)}$ состояние $\left(\mathrm{Id}_{\mathcal{H}^{(n)}} \otimes \mathcal{E}^{(n)} \right) \left[|\Omega^{(n)}\rangle\langle\Omega^{(n)}| \right]$ является чистым, и его можно рассматривать как очищение состояния $S^{(n)}$. Согласно неравенству (9.43)

$$\left\| \bar{S}^{(n)} - S_{B'} \right\|_1 \leq \left\| |\Omega^{(n)}\rangle\langle\Omega^{(n)}| - S_{B'R'} \right\|_1 \leq 2\sqrt{\varepsilon}.$$

Здесь первое неравенство вытекает из следующего простого наблюдения:

Задача 9.3.1 *Следовая норма монотонна по отношению к взятию частичного следа:*

$$\|\mathrm{Tr}_{\mathcal{H}_0} T\|_1 \leq \|T\|_1 ,$$

для любого положительного оператора $T \in \mathcal{H} \otimes \mathcal{H}_0$. Указание: Используйте выражение (1.19) для следовой нормы.

Дважды используя оценку (7.22) для модуля непрерывности энтропии, имеем

$$\begin{aligned} H(S_{B'}) - H(S_{B'R'}) &= H(\bar{S}^{(n)}) + [H(S_{B'}) - H(\bar{S}^{(n)})] \\ &\quad + [H(|\Omega^{(n)}\rangle\langle\Omega^{(n)}|) - H(S_{B'R'})] \\ &\geq H(\bar{S}^{(n)}) - 6 \log \dim \mathcal{H}^{(n)} \sqrt{\varepsilon} - \frac{2 \log e}{e} \\ &= nR \left(1 - 6\sqrt{\varepsilon}\right) - \frac{2 \log e}{e}. \end{aligned}$$

Поэтому, принимая во внимание (9.44),

$$\bar{Q}_n = \max_{S^{(n)}} I_c(S^{(n)}, \Phi^{\otimes n}) \geq nR\left(1 - 6\sqrt{\varepsilon}\right) - \frac{2\log e}{e},$$

откуда $R \leq \lim_{n\to\infty} \frac{1}{n}\bar{Q}_n = \overline{Q}(\Phi)$.

Это завершает доказательство неравенства $Q(\Phi) \leq \overline{Q}(\Phi)$, т. е. слабого обращения теоремы кодирования. Доказательство прямого утверждения будет дано в разделе 9.4.4 □

Следующее неравенство об обработке информации полезно для оценки квантовой пропускной способности:

Предложение 9.3.4 *Для любых двух каналов Φ_1, Φ_2*

$$Q(\Phi_2 \circ \Phi_1) \leq \min\left\{Q(\Phi_1), Q(\Phi_2)\right\}. \tag{9.45}$$

В частности, если один из этих каналов имеет нулевую пропускную способность, то это же верно для их композиции.

Доказательство. Чтобы доказать неравенство $Q(\Phi_2 \circ \Phi_1) \leq Q(\Phi_2)$, достаточно заметить, что пропускная способность $Q(\Phi_2 \circ \Phi_1)$ равна супремуму достижимых скоростей для канала Φ_2 при специальном выборе кодирующих каналов, включающих заключительную обработку каналом Φ_1; следовательно, эта величина не превосходит $Q(\Phi_2)$. Неравенство $Q(\Phi_2 \circ \Phi_1) \leq Q(\Phi_1)$ доказывается аналогично, путем рассмотрения декодирующих каналов для Φ_1, включающих в себя предварительную обработку каналом Φ_2.

□

9.3.3 Деградируемые каналы

Прежде всего получим важные соотношения, связывающие когерентную информацию и величину

$$\chi\left(\{\pi_x\}, \{S_x\}\right) = H(\sum_x \pi_x S_x) - \sum_x \pi_x H(S_x),$$

дающую оценку (5.15) для количества классической информации, и как следствие – соотношение между квантовой и классической пропускными способностями канала и его комплементарного.

Рассмотрим канал $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$, его расширение Стайнспринга с изометрией $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ и комплементарный канал

$$\tilde{\Phi}[S] = \mathrm{Tr}_{\mathcal{H}_B} VSV^*; \quad S \in \mathfrak{T}\left(\mathcal{H}_A\right) \tag{9.46}$$

(см. раздел 6.6).

Пусть $S = \sum_x \pi_x S_x$ – произвольное разложение по чистым состояниям (например, спектральное разложение). Тогда имеет место цепочка равенств

$$\begin{aligned} I_c(S,\Phi) &= H(\Phi[S]) - H(\tilde{\Phi}[S]) \\ &= \left[H(\Phi[S]) - \sum_x \pi_x H(\Phi[S_x]) \right] \\ &- \left[H(\tilde{\Phi}[S]) - \sum_x \pi_x H(\tilde{\Phi}[S_x]) \right] \qquad (9.47) \\ &= \chi\left(\{\pi_x\}, \{\Phi[S_x]\}\right) - \chi\left(\{\pi_x\}, \{\tilde{\Phi}[S_x]\}\right) \qquad (9.48) \\ &= \sum_x \pi_x \left[H(\Phi[S_x]; \Phi[S]) - H(\tilde{\Phi}[S_x]; \tilde{\Phi}[S]) \right]. \qquad (9.49) \end{aligned}$$

Здесь равенство (9.47) следует из того, что состояние $S_{BEx} = V S_x V^*$ является чистым и, значит,

$$H(\Phi[S_x]) = H\left(S_{Bx}\right) = H\left(S_{Ex}\right) = H(\tilde{\Phi}[S_x])$$

для всех x, а (9.49) следует из тождества (7.17). Беря в (9.48) верхнюю и нижнюю грани по всевозможным ансамблям чистых состояний, получаем

$$Q_1(\Phi) \geq C_\chi(\Phi) - C_\chi(\tilde{\Phi}) \geq -Q_1(\tilde{\Phi}), \qquad (9.50)$$

где введено обозначение

$$Q_1(\Phi) = \max_S I_c\left(S, \Phi\right). \qquad (9.51)$$

Применяя (9.50) к каналу $\Phi^{\otimes n}$, переходя к пределу при $n \to \infty$ и используя теоремы кодирования, получаем

$$Q(\Phi) \geq C(\Phi) - C(\tilde{\Phi}) \geq -Q(\tilde{\Phi}).$$

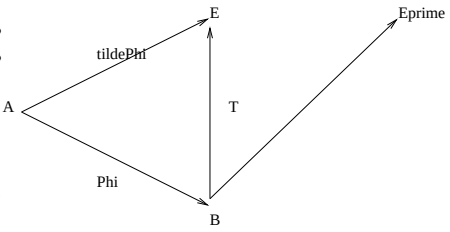


Рис. 9.1. *Деградируемый канал.*

Введем важный класс каналов, для которых квантовая пропускная способность дается “однобуквенным” выражением (9.51).

Определение 9.3.2 *Канал Φ называется деградируемым, если найдется канал T, такой что $\tilde{\Phi} = T \circ \Phi$, и анти-деградируемым, если найдется канал T', такой что $\Phi = T' \circ \tilde{\Phi}$.*

Очевидно, что Φ является деградируемым тогда и только тогда, когда $\tilde{\Phi}$ анти-деградируемый. Из соотношения (9.49), используя свойство монотонности относительной энтропии, получаем: если Φ анти-деградируемый канал, то

$$I_c(S, \Phi) \leq 0 \tag{9.52}$$

для любого состояния S (соответственно, $I_c(S, \Phi) \geq 0$ для деградируемого канала).

Предложение 9.3.5 *Если Φ – анти-деградируемый канал, то $Q(\Phi) = 0$. Если Φ – деградируемый, то*

$$Q(\Phi) = Q_1(\Phi) = \max_S I_c(S, \Phi). \tag{9.53}$$

Доказательство. Первое утверждение следует из (9.52) и теоремы кодирования (9.40).

Пусть Φ – деградируемый канал, и $W : \mathcal{H}_B \to \mathcal{H}_E \otimes \mathcal{H}_{E'}$ – изометрия Стайнспринга для канала $T : \mathfrak{T}(\mathcal{H}_B) \to \mathfrak{T}(\mathcal{H}_E)$. Тогда $I_c(S, \Phi) = H(B) - H(E) = H(EE') - H(E)$, так что

$$I_c(S, \Phi) = H(E'|E).$$

Для двух деградируемых каналов Φ_1, Φ_2 и произвольного состояния S_{12} в $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ имеем

$$I_c(S_{12}, \Phi_1 \otimes \Phi_2) = H(E_1' E_2' | E_1 E_2),$$

следовательно, по свойству субаддитивности условной энтропии

$$I_c(S_{12}, \Phi_1 \otimes \Phi_2) \leq I_c(S_1, \Phi_1) + I_c(S_2, \Phi_2),$$

откуда получаем

$$I_c\left(S^{(n)}, \Phi^{\otimes n}\right) = n \max_S I_c(S, \Phi) \tag{9.54}$$

и, следовательно, (9.53). □

Пример Рассмотрим канал, разрушающий сцепленность

$$\Phi[S] = \sum_{\alpha=1}^{\tilde{d}} |\varphi_\alpha\rangle\langle\psi_\alpha|S|\psi_\alpha\rangle\langle\varphi_\alpha|,$$

где $\{\psi_\alpha\}$ – переполненная система в $\mathcal{H}$, а $\{\varphi_\alpha\}$ – система единичных векторов в выходном пространстве $\mathcal{H}'$. Комплементарным является диагональный канал

$$\tilde{\Phi}[S] = \sum_{\alpha,\beta=1}^{\tilde{d}} c_{\alpha\beta}|e_\alpha\rangle\langle\psi_\alpha|S|\psi_\beta\rangle\langle e_\beta|,$$

где $c_{\alpha\beta} = \langle\varphi_\beta|\varphi_\alpha\rangle$, а $\{e_\alpha\}$ – стандартный базис в $\mathcal{H}_{\tilde{d}}$. Имеем $\Phi = T' \circ \tilde{\Phi}$, где

$$T'[S] = \sum_{\alpha=1}^{\tilde{d}} |\varphi_\alpha\rangle\langle e_\alpha|S|e_\alpha\rangle\langle\varphi_\alpha|.$$

Отсюда и из второго утверждения теоремы 9.3.5 вытекает

Следствие 9.3.1 *Всякий канал Φ, разрушающий сцепленность, является анти-деградируемым и следовательно, $Q(\Phi) = 0$.*

Пример Рассмотрим квантовый стирающий канал Φ_p. При $q \geq p$ имеем

$$\Phi_q = T_{(1-q)/(1-p)} \circ \Phi_p,$$

где $T_\alpha : \mathfrak{T}(\mathcal{H} \oplus \mathbb{C}) \to \mathfrak{T}(\mathcal{H} \oplus \mathbb{C})$ – канал, действующий по правилу

$$T_\alpha \left[\begin{bmatrix} S & 0 \\ 0 & s \end{bmatrix}\right] = \alpha \begin{bmatrix} S & 0 \\ 0 & s \end{bmatrix} + (1-\alpha) \begin{bmatrix} 0 & 0 \\ 0 & \operatorname{Tr} S + s \end{bmatrix}.$$

Используя тот факт, что $\tilde{\Phi}_p = \Phi_{1-p}$ (см. задачу 6.6.3), получаем, что канал Φ_p – деградируемый при $p \in [0, 1/2]$ и анти-деградируемый при $p \in [1/2, 1]$. Применяя теорему 9.3.5, получаем

$$Q(\Phi_p) = \begin{cases} (1-2p)\log d, & p \in [0, 1/2]; \\ 0, & p \in [1/2, 1]. \end{cases} \tag{9.55}$$

Здесь первое утверждение вытекает из (9.53) и того факта, что $H(\Phi_p[S]) = (1-p)H(S) + h_2(p)$, следовательно $I_c(S, \Phi_p) = H(\Phi_p[S]) - H(\tilde{\Phi}_p[S]) = (1-p)H(S) - pH(S) = (1-2p)H(S)$.

9.4 Секретная классическая пропускная способность и квантовая пропускная способность

9.4.1 Квантовый канал с перехватом

Рассмотрим ситуацию передачи классической информации, в которой имеется три участника: отправитель A, получатель B и перехватчик E. Математическая модель *квантового канала с перехватом* включает

три гильбертова пространства $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_E$ и изометрическое отображение $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$, при котором входное состояние S_A переходит в состояние $S_{BE} = \Phi_{BE}[S_A] \equiv V S_A V^*$ системы BE, с частичными состояниями

$$S_B = \Phi_B[S_A] \equiv \mathrm{Tr}_E V S_A V^*, \quad S_E = \Phi_E[S_A] \equiv \mathrm{Tr}_B V S_A V^*.$$

Заметим, что ранее буква E использовалась для обозначения окружения. Если все окружение доступно перехватчику, то такое обозначение является согласованным. Предположим, что A выбирает состояния $\{S_A^x\}$ с вероятностями $\{\pi_x\}$; тогда участники B и E получают, соответственно, состояния $\{S_B^x\}$ и $\{S_E^x\}$, верхними границами шенноновской информации для B и E являются величины $\chi(\{\pi_x\}, \{S_B^x\})$ и $\chi(\{\pi_x\}, \{S_E^x\})$. По аналогии с классическим каналом с перехватом (см. раздел 4.5), "секретность" передачи может быть охарактеризована величиной $\chi(\{\pi_x\}, \{S_B^x\}) - \chi(\{\pi_x\}, \{S_E^x\})$. Предполагая, что входные состояния S_A^x являются *чистыми*, и обозначая $\bar{S}_A = \sum_x \pi_x S_A^x$ среднее состояние входного ансамбля, из (9.48) получаем ключевое соотношение

$$I_c\left(\bar{S}_A, \Phi_B\right) = \chi(\{\pi_x\}, \{S_B^x\}) - \chi(\{\pi_x\}, \{S_E^x\}), \tag{9.56}$$

которое указывает на важную связь между когерентной информацией и секретной классической пропускной способностью; оно также подсказывает идею доказательства прямой теоремы кодирования для квантовой пропускной способности через рассмотрение канала с перехватом.

Для определения секретной классической пропускной способности рассмотрим блочное кодирование для квантового канала с перехватом. Нашей целью будет передать максимальное количество классической информации от A к B секретным образом, так, чтобы E смог получить лишь асимптотически исчезающее количество информации. *Код* $\left(\Sigma^{(n)}, M^{(n)}\right)$ длины n и размера N определяется так же, как в определении 8.1.1. Таким образом, код состоит из набора состояний $\Sigma^{(n)} = \left\{S_{A^{(n)}}^i; i = 1, \ldots, N\right\}$ в $\mathcal{H}_A^{\otimes n}$ вместе с наблюдаемой $M^{(n)} = \{M_j; j = 0, 1, \ldots, N\}$ в $\mathcal{H}_B^{\otimes n}$. Наряду с вероятностью ошибки $P_e\left(\Sigma^{(n)}, M^{(n)}\right)$, определяемой соотношением (8.1), введем новую величину

$$v_E\left(\Sigma^{(n)}\right) = \max_{i,k=1,\ldots,N} \left\| S_{E^{(n)}}^i - S_{E^{(n)}}^k \right\|_1,$$

характеризующую *изменчивость* состояния системы E, а, следовательно, и количество информации, которое доступно перехватчику E. Заметим, что из неравенства треугольника для нормы вытекает

$$\left\| \bar{S}_{E^{(n)}} - S_{E^{(n)}}^i \right\|_1 \leq v_E\left(\Sigma^{(n)}\right) \quad \text{для всех } i,$$

где $\bar{S}_{E^{(n)}} = \frac{1}{N} \sum_{i=1}^N S_{E^{(n)}}^i$. Применяя неравенство (7.21), получаем

$$\frac{1}{n}\chi\left(\left\{\frac{1}{N}\right\},\{S^i_{E^{(n)}}\}\right)=\frac{1}{nN}\sum_{i=1}^{N}\left[H\left(\bar{S}_{E^{(n)}}\right)-H\left(S^i_{E^{(n)}}\right)\right] \tag{9.57}$$
$$\leq \log\dim\mathcal{H}_E\cdot v_E\left(\Sigma^{(n)}\right)+\frac{1}{n}\eta\left(v_E\left(\Sigma^{(n)}\right)\right),$$

в предположении, что величина $v_E\left(\Sigma^{(n)}\right)$ достаточно мала. Поэтому если изменчивость мала, то мала и шенноновская взаимная информация между A и E для равновероятных сообщений.

Мы будем называть R *достижимой скоростью* для канала с перехватом, если найдется последовательность кодов $\left(\Sigma^{(n)},M^{(n)}\right)$ размера $N=2^{nR}$ таких, что

$$\lim_{n\to\infty}P_e\left(\Sigma^{(n)},M^{(n)}\right)=0$$

и

$$\lim_{n\to\infty}v_E\left(\Sigma^{(n)}\right)=0. \tag{9.58}$$

Точная верхняя грань множества достижимых скоростей называется *секретной классической пропускной способностью* $C_p\left(\Phi_{BE}\right)$ канала с перехватом. Согласно (9.57), условие (9.58) влечет асимптотическое исчезновение взаимной информации между A и E. Можно показать, что на самом деле эти утверждения равносильны, однако условие (9.58) более удобно математически.

Теорема 9.4.1

$$C_p\left(\Phi_{BE}\right)=\lim_{n\to\infty}\frac{1}{n}\max_{\pi^{(n)},\Sigma^{(n)}}\left[\chi\left(\{\pi_i^{(n)}\},\{S^i_{B^{(n)}}\}\right)-\chi\left(\{\pi_i^{(n)}\},\{S^i_{E^{(n)}}\}\right)\right], \tag{9.59}$$

где максимум берется по всем конечным наборам состояний $\Sigma^{(n)}=\left\{S^i_{A^{(n)}}\right\}$ *в* $\mathcal{H}_A^{\otimes n}$ *и распределениям вероятностей* $\pi^{(n)}=\{\pi_i^{(n)}\}$ *(мы используем обозначения* $S^i_{B^{(n)}}=\Phi_B^{\otimes n}\left[S^i_{A^{(n)}}\right], S^i_{E^{(n)}}=\Phi_E^{\otimes n}\left[S^i_{A^{(n)}}\right]$*).*

Из соотношений (9.40), (9.56) и (9.59) вытекает важное неравенство между квантовой и классической секретной пропускными способностями

$$Q\left(\Phi_B\right)\leq C_p\left(\Phi_{BE}\right).$$

Оно следует из того, что при вычислении $C_p\left(\Phi_{BE}\right)$ учитываются все ансамбли состояний, а при вычислении $Q\left(\Phi_B\right)$ – только ансамбли чистых состояний для A. В общем случае здесь возможно строгое неравенство, поэтому особый интерес представляет следующее утверждение.

Предложение 9.4.2 *Если канал* Φ_B *деградируемый, то*

$$C_p\left(\Phi_{BE}\right)=Q\left(\Phi_B\right)=Q_1\left(\Phi_B\right). \tag{9.60}$$

Доказательство. Используя равенство (9.48), имеем

$$\begin{aligned}
&\chi\left(\{\pi_i^{(n)}\},\{S^i_{B^{(n)}}\}\right)-\chi\left(\{\pi_i^{(n)}\},\{S^i_{E^{(n)}}\}\right)\\
&=H\left(\sum_i \pi_i^{(n)} S^i_{B^{(n)}}\right)-H\left(\sum_i \pi_i^{(n)} S^i_{E^{(n)}}\right)\\
&-\sum_i \pi_i^{(n)}\left[H\left(S^i_{B^{(n)}}\right)-H\left(S^i_{E^{(n)}}\right)\right]\\
&=I_c\left(\sum_i \pi_i^{(n)} S^i_{A^{(n)}},\Phi_B^{\otimes n}\right)-\sum_i \pi_i^{(n)} I_c\left(S^i_{A^{(n)}},\Phi_B^{\otimes n}\right)\\
&\le I_c\left(\sum_i \pi_i^{(n)} S^i_{A^{(n)}},\Phi_B^{\otimes n}\right),
\end{aligned}$$

поскольку канал $\Phi_B^{\otimes n}$, как и Φ_B, деградируем, а значит $I_c\left(S^i_{A^{(n)}},\Phi_B^{\otimes n}\right)\ge 0$. Следовательно,

$$\max_{\pi^{(n)},\Sigma^{(n)}}\left[\chi\left(\{\pi_i^{(n)}\},\{S_B^i\}\right)-\chi\left(\{\pi_i^{(n)}\},\{S^i_{E^{(n)}}\}\right)\right]=\max_{S^{(n)}} I_c\left(S^{(n)},\Phi_B^{\otimes n}\right),$$

что равно $nQ_1(\Phi_B)$ в силу (9.54). Подставляя в (9.59), получаем (9.60). □

9.4.2 Доказательство теоремы о секретной пропускной способности

Сначала докажем неравенство $\le$ в (9.59), т. е. слабое обращение. Пусть R – достижимая скорость, тогда используя (5.29) и границу (5.15), получаем

$$\begin{aligned}
\bar{P}_e\left(\Sigma^{(n)},M^{(n)}\right)&\ge 1-\frac{\chi\left(\{1/N\},\{S^i_{B^{(n)}}\}\right)}{nR}-\frac{1}{nR}\\
&=1-\frac{\chi\left(\{1/N\},\{S^i_{B^{(n)}}\}\right)-\chi\left(\{1/N\},\{S^i_{E^{(n)}}\}\right)}{nR}\\
&-\frac{1+\chi\left(\{1/N\},\{S^i_{E^{(n)}}\}\right)}{nR}.
\end{aligned}\tag{9.61}$$

Согласно (9.57)

$$\chi\left(\{1/N\},\{S^i_{E^{(n)}}\}\right)\le v_E\left(\Sigma^{(n)}\right)\left[n\log d_E-\log v_E\left(\Sigma^{(n)}\right)\right],$$

где $d_E=\dim\mathcal{H}_E$, поэтому последнее слагаемое в правой части (9.61) стремится к нулю при $n\to\infty$; ошибка $\bar{P}_e$ также стремится к нулю. Отсюда получаем $0\ge 1-\bar{C}_p/R$, где $\bar{C}_p$ – правая часть в (9.59), следовательно, $C_p\left(\Phi_{BE}\right)\le\bar{C}_p$.

Для доказательства прямого утверждения достаточно показать, что для произвольного ансамбля $\{\pi_x\}, \{S_A^x\}$ и $\delta > 0$ скорость $\chi(\{\pi_x\}, \{S_B^x\}) - \chi(\{\pi_x\}, \{S_E^x\}) - \delta$ является достижимой. Тогда достижимость скорости $\bar{C}_p - \delta$ получается путем дополнительной группировки в блоки. Рассмотрим c-q канал $x \to S_B^x$ и случайный набор кодовых слов $W^{(n)}$ размера $N = 2^{n[\chi(\{\pi_x\},\{S_B^x\})-c\delta]}$ с независимыми словами, имеющими распределение вероятностей

$$\mathsf{P}\{w = (x_1, \ldots, x_n)\} = \pi_{x_1} \cdot \ldots \cdot \pi_{x_n}. \tag{9.62}$$

Константа c будет определена ниже. Согласно замечанию после доказательства прямого утверждения квантовой теоремы кодирования в конце раздела 5.6, для достаточно больших n найдутся наблюдаемые $M^{(n)}$ в системе B такие, что

$$\mathsf{E}\bar{P}_e(W^{(n)}, M^{(n)}) \leq 2^{-n\beta}, \quad \beta > 0.$$

Теперь рассмотрим новый случайный код с независимыми словами, равномерно распределенными на множестве $\hat{T}^{n,\delta}$ сильно $\delta-$типичных слов:

$$\tilde{\mathsf{P}}(w) = \begin{cases} \frac{1}{\left|\hat{T}^{n,\delta}\right|}, & w \in \hat{T}^{n,\delta}; \\ 0, & w \notin \hat{T}^{n,\delta}. \end{cases}$$

Согласно задаче 8.4.3, $\left|\hat{T}^{n,\delta}\right| \geq 2^{n[H(\pi)-\Delta_n(\delta)]}$, причем $\lim_{\delta\to 0}\lim_{n\to\infty}\Delta_n(\delta) = 0$. Так как $\mathsf{P}(w) \geq 2^{-n[H(\pi)+c\delta]}$ для $w \in \hat{T}^{n,\delta}$, где c – константа, зависящая от $\{\pi_x\}$, то $\tilde{\mathsf{P}}(w) \leq 2^{n[c\delta+\Delta_n(\delta)]}\mathsf{P}(w)$ для всех w, и, следовательно,

$$\tilde{\mathsf{E}}\bar{P}_e(W^{(n)}, M^{(n)}) \leq 4^{n[c\delta+\Delta_n(\delta)]}2^{-n\beta} \leq \varepsilon$$

для достаточно больших n, поскольку оценка (5.46) для $\bar{P}_e(W^{(n)}, M^{(n)})$ включает не более двух независимых слов. Следовательно,

$$\tilde{\mathsf{P}}\left\{\bar{P}_e(W^{(n)}, M^{(n)}) \geq \sqrt{\varepsilon}\right\} \leq \sqrt{\varepsilon} \tag{9.63}$$

по неравенству Чебышева.

Пока что была дана лишь некоторая модификация случайного кодирования для передачи классической информации от A к B. Чтобы сделать код секретным, передатчик A должен пожертвовать $n[\chi(\{\pi_x\}, \{S_E^x\}) + c\delta/2]$ битами информации, дополнительно рандомизуя входные состояния. Положим

$$N_E = 2^{n[\chi(\{\pi_x\},\{S_E^x\})+c\delta/2]}, N_B = 2^{n[\chi(\{\pi_x\},\{S_B^x\})-\chi(\{\pi_x\},\{S_E^x\})-3c\delta/2]},$$

так что $N_E N_B = N$; представим набор кодовых слов $W^{(n)}$ в виде прямоугольной таблицы с N_B строками и N_E столбцами. Тогда

$$W^{(n)} = \left\{w^{mj}; m = 1, \ldots, N_B; j = 1, \ldots, N_E\right\},$$
$$M^{(n)} = \left\{M^{mj}; m = 1, \ldots, N_B; j = 1, \ldots, N_E\right\}.$$

Пусть для каждого значения m передатчик A выбирает значение j случайным образом с равными вероятностями, что приводит к входным состояниям

$$m \to S^m_{A^{(n)}} = \frac{1}{N_E} \sum_{j=1}^{N_E} S^{w^{mj}}_{A^{(n)}},$$

преобразуемым в выходные состояния $S^m_{B^{(n)}}$ для B и $S^m_{E^{(n)}}$ для E. (Напомним, что $S^w_{A^{(n)}} = S^{x_1}_A \otimes \cdots \otimes S^{x_n}_A$ для слова $w = (x_1, \ldots, x_n)$.) Такая рандомизация приводит к тому, что почти вся переданная информация оказывается скрытой для E: для каждого значения m набор кодовых слов

$$\left\{w^{mj}; j = 1, \ldots, N_E\right\}$$

с высокой вероятностью содержит почти максимально возможную информацию от A к E, при условии, что E применяет оптимальное декодирование. Поэтому взаимная информация между наборами кодовых слов с различными значениями m должна быть близка к нулю, так что рандомизация внутри каждого набора уничтожает почти всю информацию, передаваемую от A к E.

Решающим шагом в строгом обосновании этой идеи является применение следующей оценки, основанной на квантовой версии неравенства Бернштейна, которая доказывается в следующем разделе:

Предложение 9.4.3 *Пусть $S^{w^{mj}}_{E^{(n)}}; m = 1, \ldots, N_B; j = 1, \ldots, N_E$ – случайные независимые одинаково распределенные операторы плотности; величины N_B, N_E определены выше. Тогда найдется число $\beta_1 > 0$, такое, что для достаточно больших n*

$$\tilde{\mathsf{P}}\left\{\left\|\frac{1}{N_E}\sum_{j=1}^{N_E} S^{w^{mj}}_{E^{(n)}} - \theta\right\|_1 \geq \varepsilon\right\} \leq 2^{-2^{n\beta_1}}, \tag{9.64}$$

для $m = 1, \ldots, N_B$, где θ – некоторый неслучайный оператор (в действительности можно взять $\theta = \tilde{\mathsf{E}} S^{w^{mj}}_{E^{(n)}}$, но нам это не потребуется).

Из этой оценки вытекает, что

$$\tilde{\mathsf{P}}\left\{\left\|S^m_{E^{(n)}} - \theta\right\|_1 < \varepsilon; \quad m = 1, \ldots, N_B\right\} \geq 1 - N_B 2^{-2^{n\beta_1}}$$
$$= 1 - 2^{n[\chi(\{\pi_x\},\{S^x_B\}) + c\delta/2] - 2^{n\beta_1}},$$

то есть вероятность может быть сколь угодно близка к 1 для достаточно больших n. Учитывая (9.63), получаем, что найдется реализация $W^{(n)} = \left\{w^{mj}\right\}$ набора кодовых слов, для которых

$$\bar{P}_e(W^{(n)}, M^{(n)}) < \sqrt{\varepsilon};$$

и

$$\left\| S^m_{E^{(n)}} - S^l_{E^{(n)}} \right\|_1 < 2\varepsilon; \quad m, l = 1, \ldots, N_B. \tag{9.65}$$

Определяя $\Sigma^{(n)} = \left\{ S^m_{A^{(n)}}; \quad m = 1, \ldots, N_B \right\}, \tilde{M}^{(n)} = \left\{ \tilde{M}_m; m = 0, 1, \ldots, N_B \right\}$, где $\tilde{M}_m = \sum_{j=1}^{N_E} M_{mj}$, имеем $\bar{P}_e \left(\Sigma^{(n)}, \tilde{M}^{(n)} \right) \leq \bar{P}_e(W^{(n)}, M^{(n)}) < \sqrt{\varepsilon}$. Рассуждения, аналогичные лемме 4.4.1, показывают, что можно выбрать подкод, для которого максимальная ошибка $P_e(W^{(n)}, M^{(n)}) < 2\sqrt{\varepsilon}$ при $v_E \left(\Sigma^{(n)} \right) < \varepsilon'$, где $\varepsilon' \to 0$ при $\varepsilon \to 0$, что и завершает доказательство теоремы. □

Доказательство предложения 9.4.3.

Рассуждая как в разделе 5.6, рассмотрим оператор плотности

$$\bar{S}_\pi = \sum_x \pi_x S^x_E.$$

Для краткости будем использовать обозначения

$$H(\bar{S}_\pi) = H(E), \quad \sum_x \pi_x H(S^x_E) = H(E|A),$$

так что $\chi\left(\{\pi_x\}, \{S^x_E\}\right) = H(E) - H(E|A)$.

Обозначим λ_j собственные значения оператора $\bar{S}_\pi$. Фиксируем малое положительное δ, и пусть $P = P^{n,\delta}$ – сильно типичный проектор оператора плотности $\bar{S}^{\otimes n}_\pi$, соответствующий собственным значениям $\lambda_J = \lambda_{j_1} \cdot \ldots \cdot \lambda_{j_n}$, для которых последовательность $J = (j_1, \ldots, j_n)$ является сильно типичной, т. е. частоты $N(j|J)$ появления символа j в J удовлетворяют условию

$$\left| \frac{N(j|J)}{n} - \lambda_j \right| < \delta, \quad j = 1, \ldots, d_E,$$

и $N(j|J) = 0$ если $\lambda_j = 0$ (см. определение 8.4.1). Обозначим через $\lambda_j(x)$ собственные значения оператора плотности S^x_E. Для кодового слова $w = (x_1, \ldots, x_n)$ длины n, положим $S_w = S^{x_1}_E \otimes \cdots \otimes S^{x_n}_E$ и обозначим P_w типичный проектор оператора S_w, соответствующий собственным значениям $\lambda_J(w) = \lambda_{j_1}(x_1) \cdot \ldots \cdot \lambda_{j_n}(x_n)$, для которых

$$\left| \frac{N(j, x|J, w)}{n} - \frac{N(x|w)}{n} \lambda_j(x) \right| < \delta, \quad j = 1, \ldots, d_E,$$

и $N(j, x|J, w) = 0$ если $\lambda_j(x) = 0$. Нам понадобятся следующие свойства, которые являются переформулировкой соответствующих свойств типичных последовательностей:

i. Для $w \in \hat{T}^{n,\delta}$, *собственные значения оператора* $P_w S_w P_w$ *лежат в интервале* $\left(2^{-n[H(E|A)+c\delta]}, 2^{-n[H(E|A)-c\delta]}\right)$, *где* c – *некоторая константа, зависящая от* $\{\lambda_j(x)\}$ (для упрощения обозначений мы используем одно и то же обозначение c для всех констант; в действительности, мы можем выбрать максимальную из всех констант). В частности, $P_w S_w P_w \leq 2^{-n[H(E|A)-c\delta]} P_w$.

ii. Для $\varepsilon_1 > 0$ *и достаточно больших* n *выполняется* $\mathrm{Tr} P_w S_w \geq 1 - \varepsilon_1$. (Это является следствием закона больших чисел для распределения вероятностей, задаваемого собственными значениями оператора S_w).

Следующее свойство является менее очевидным.

iii. Для данных $\varepsilon_1, \delta > 0$ *найдется* $\delta_1 > 0$ *такое, что для* $w \in T^{n,\delta_1}$ *и достаточно больших* n $\mathrm{Tr} P S_w \geq 1 - \varepsilon_1$, *где* $P = P^{n,\delta}$ – *типичный проектор оператора* $\bar{S}_\pi^{\otimes n}$.

Доказательство. Обозначим $p^{(k)}(j) = \langle e_j | S_E^{x_k} | e_j \rangle$, где $|e_j\rangle$ – собственные векторы оператора $\bar{S}_\pi$, и заметим, что $p^{(k)}(j) = 0$, если $\lambda_j = 0$, так как $\mathrm{supp} S_E^{x_k} \subset \mathrm{supp} \bar{S}_\pi$. Имеем

$$\begin{aligned} \mathrm{Tr} P^{n,\delta} S_w &= \sum_{J \in B^{n,\delta}} p^{(1)}(j_1) \cdot \dots \cdot p^{(n)}(j_n) \\ &= \mathsf{P}\left\{ \left| \frac{N(j|J)}{n} - \lambda_j \right| < \delta, \text{если } \lambda_j > 0 \right\}, \end{aligned} \tag{9.66}$$

где P – распределение вероятностей, приписывающее вероятность $p^{(1)}(j_1) \cdot \dots \cdot p^{(n)}(j_n)$ последовательности $J = (j_1, \dots, j_n)$, $B^{n,\delta}$ – множество сильно типичных последовательностей. Тот факт, что $w \in T^{n,\delta_1}$ выражается аналогичными неравенствами

$$\left| \frac{N(x|w)}{n} - \pi_x \right| < \delta_1, \quad \text{если } \pi_x > 0,$$

и $N(x|w) = 0$, если $\pi_x = 0$, откуда следует, для $\bar{\lambda}_j = \frac{1}{n} \sum_{k=1}^n p^{(k)}(j)$:

$$\left| \bar{\lambda}_j - \lambda_j \right| = \left| \langle e_j | \left[\frac{1}{n} \sum_{k=1}^n S_E^{x_k} - \bar{S}_\pi \right] | e_j \rangle \right| = \left| \sum_x \left[\frac{N(x|w)}{n} - \pi_x \right] \langle e_j | S_E^x | e_j \rangle \right| < \delta_1 K,$$

где K – количество различных символов x. Поэтому, выбирая $\delta_1 = \delta / 2K$, имеем

$$\begin{aligned} &\mathsf{P}\left\{ \left| \frac{N(j|J)}{n} - \lambda_j \right| < \delta; \text{если } \lambda_j > 0 \right\} \\ &\geq \mathsf{P}\left\{ \left| \frac{N(j|J)}{n} - \bar{\lambda}_j \right| < \delta/2; \text{ для всех } j : \lambda_j > 0 \right\} \\ &\geq 1 - \sum_{j : \lambda_j > 0} \mathsf{P}\left\{ \left| \frac{N(j|J)}{n} - \bar{\lambda}_j \right| \geq \delta/2 \right\}. \end{aligned}$$

Для распределения вероятностей P, случайная величина $N(j|J)$ имеет среднее $\sum_{k=1}^{n} p^{(k)}(j) = n\bar{\lambda}_j$ и дисперсию $\sum_{k=1}^{n} p^{(k)}(j)\left(1-p^{(k)}(j)\right) \leq n/4$. Применяя неравенство Чебышева и используя (9.66), получаем *iii*. □

Далее нам понадобится

Лемма 9.4.4 *Для любого положительного оператора X и проектора P*

$$\|X - PXP\|_1 \leq 3\sqrt{\mathrm{Tr}X\mathrm{Tr}(I-P)X}.$$

Доказательство. Согласно неравенству треугольника и свойствам следовой нормы

$$\|X - PXP\|_1 \leq \|(I-P)X(I-P)\|_1 + 2\|PX(I-P)\|_1 \leq 3\|X(I-P)\|_1.$$

Пусть U унитарный оператор, такой что $\|X(I-P)\|_1 = \mathrm{Tr}X(I-P)U$. По неравенству Коши - Шварца для следа (при $S = I/d$ в (2.25))

$$\left[\mathrm{Tr}\sqrt{X}\sqrt{X}(I-P)U\right]^2 \leq \mathrm{Tr}X\ \mathrm{Tr}U^*(I-P)X(I-P)U = \mathrm{Tr}X\mathrm{Tr}(I-P)X,$$

отсюда и следует утверждение леммы. □

Для упрощения обозначений в (9.64), обозначим $S_{E^{(n)}}^{w^{mj}} = S_{w^j}$, где кодовые слова $w^j; j = 1, \ldots, N_E$, являются независимыми и равномерно распределенными на множестве T^{n,δ_1}. Мы выведем соотношение (9.64), применяя операторное неравенство типа Бернштейна (9.73), доказываемое в следующем разделе, к случайным операторам

$$X_j = 2^{n[H(E|A)-\delta]} \Pi P P_{w^j} S_{w^j} P_{w^j} P \Pi,$$

где Π – проектор на собственное подпространство оператора

$$\theta' = \tilde{\mathsf{E}} P P_{w^j} S_{w^j} P_{w^j} P$$

соответствующий собственным значениям $\geq 2^{-n[H(E)+3c\delta/2]}$. По построению, $0 \leq X_j \leq I$,

$$M = \tilde{\mathsf{E}} X_j = 2^{n[H(E|A)-c\delta]} \tilde{\mathsf{E}}(\Pi P P_{w^j} S_{w^j} P_{w^j} P \Pi) = 2^{n[H(E|A)-c\delta]} \Pi \theta' \Pi$$

и $\mu = 2^{n[H(E|A)-H(E)-5c\delta/2]} = 2^{-n[\chi(\{\pi_x\},\{S_E^x\})+5c\delta/2]}$, так что $N_E\mu = 2^{-2nc\delta}$ и из (9.73) следует, что

$$\tilde{\mathsf{P}}\left\{\omega : \left\| \frac{1}{N_E}\sum_{j=1}^{N_E} X_j(\omega) - M \right\|_1 \geq \varepsilon_2 2^{n[H(E|A)-c\delta]}\right\}$$
$$\leq 2d_E^n \exp\left(-2^{-2nc\delta}\frac{\varepsilon_2^2}{4}\right) \leq 2^{-2^{-n\beta_1}} \qquad (9.67)$$

для некоторого $\beta_1 > 0$ и достаточно большого n. Здесь мы использовали, что $\mathrm{Tr}\, M \leq 2^{n[H(E|A)-c\delta]}$. Теперь получим (9.64) из (9.67).

Обозначая $\theta = \Pi\theta'\Pi$, имеем

$$\left\| \frac{1}{N_E}\sum_{j=1}^{N_E} S_{w^j} - \theta \right\|_1 \leq \frac{1}{N_E}\sum_{j=1}^{N_E} \|S_{w^j} - PS'_{w^j}P\|_1 \tag{9.68}$$

$$+ \left\| \frac{1}{N_E}\sum_{j=1}^{N_E} PS'_{w^j}P - \frac{1}{N_E}\sum_{j=1}^{N_E} \Pi PS'_{w^j}P\Pi \right\|_1 \tag{9.69}$$

$$+ \left\| \frac{1}{N_E}\sum_{j=1}^{N_E} \Pi PS'_{w^j}P\Pi - \theta \right\|_1 . \tag{9.70}$$

Из оценки (9.67) вытекает, что

$$\tilde{\mathsf{P}}\left\{ \left\| \frac{1}{N_E}\sum_{j=1}^{N_E} \Pi PS'_{w^j}P\Pi - \theta \right\|_1 \leq \varepsilon_2 \right\} > 1 - 2^{-2^{-n\beta_1}}.$$

Неравенство

$$\left\| \frac{1}{N_E}\sum_{j=1}^{N_E} \Pi PS'_{w^j}P\Pi - \theta \right\|_1 \leq \varepsilon_2 \tag{9.71}$$

дает оценку слагаемого (9.70).

Для слагаемого (9.68) используем неравенство треугольника

$$\|S_{w^j} - PS'_{w^j}P\|_1 \leq \|S_{w^j} - P_{w^j}S'_{w^j}P_{w^j}\|_1 + \|S'_{w^j} - PS'_{w^j}P\|_1 .$$

Первое слагаемое в правой части можно сделать малым для достаточно больших n (согласно свойству $ii.$ и лемме 9.4.4). Далее отметим, что для $w^j \in T^{n,\delta_1}$

$$\mathrm{Tr}PS'_{w^j} = \mathrm{Tr}S'_{w^j} - \mathrm{Tr}(I-P)S'_{w^j} \geq \mathrm{Tr}P_{w^j}S_{w^j} - \mathrm{Tr}(I-P)S_{w^j} \geq 1 - 2\varepsilon_1 \tag{9.72}$$

в силу $ii.$ и $iii.$ Применяя лемму, мы можем сделать второе слагаемое в правой части малым для достаточно больших n.

Слагаемое (9.69) имеет вид $\|S - \Pi S\Pi\|_1$, поэтому для доказательства его близости к нулю, согласно лемме 9.4.4 достаточно доказать, что величина

$$\mathrm{Tr}\Pi S = \mathrm{Tr}\frac{1}{N_E}\sum_{j=1}^{N_E} \Pi PS'_{w^j}P\Pi$$

близка к 1. Но в соответствии с (9.71), это выражение не меньше, чем $\mathrm{Tr}\theta - \varepsilon_2 = \mathrm{Tr}\Pi\theta' - \varepsilon_2$. Теперь заметим, что $\mathrm{Tr}\Pi\theta' = \mathrm{Tr}\theta' - \mathrm{Tr}(I-\Pi)\theta' \geq$

$\mathrm{Tr}\theta' - 2^{-nc\delta/2}$, так как $\mathrm{Tr}(I-\Pi)\theta'$ является суммой собственных значений θ', которые не превосходят $2^{-n[H(E)+3c\delta/2]}$, тогда как общее количество положительных собственных значений, не превосходит $\dim \mathrm{supp}\theta' \leq \dim \mathrm{supp}P = 2^{n[H(E)+c\delta]}$. Остается доказать, что $\mathrm{Tr}\theta'$ близко к 1. Мы знаем, что $\theta' = \tilde{\mathsf{E}} P S'_{w^j} P$ и $\mathrm{Tr}\theta' = \tilde{\mathsf{E}}\mathrm{Tr} P S'_{w^j} \geq 1 - 2\varepsilon_1$ в силу (9.72). Это означает, что при заданном ε можно выбрать $\varepsilon_1, \varepsilon_2$ такими, что (9.64) будет выполнено для достаточно большого n. □

9.4.3 Большие уклонения для случайных операторов

Теорема 9.4.5 *(Операторное неравенство типа Бернштейна) Пусть* $X_1(\omega), \ldots, X_N(\omega)$ *– независимые одинаково распределенные операторно-значные случайные величины в* $\mathcal{H}$, $\dim \mathcal{H} = d$, *такие что* $0 \leq X_i(\omega) \leq I$, $M = \mathsf{E}X_i \geq \mu I$, *где* $0 < \mu \leq \frac{1}{2}$. *Тогда для* $0 \leq \varepsilon \leq 1$,

$$\mathsf{P}\left\{\omega : \left\| \frac{1}{N}\sum_{j=1}^{N} X_j(\omega) - M \right\|_1 \geq \varepsilon \,\mathrm{Tr}\, M \right\} \leq 2d \exp\left(-\frac{N\mu\varepsilon^2}{4}\right). \qquad (9.73)$$

Доказательство. Доказательство состоит из нескольких шагов. Сначала докажем операторное обобщение неравенства Маркова: пусть X – эрмитова операторно-значная случайная величина, $t > 0$, тогда

$$\mathsf{P}\{\omega : X(\omega) \not\leq 0\} \leq \mathrm{Tr}\,\mathsf{E}\exp tX,$$

где $\{\omega : X(\omega) \not\leq 0\} = \{\omega : X(\omega) \leq 0\}^c$. Действительно,

$$\begin{aligned}\mathsf{P}\{\omega : X(\omega) \not\leq 0\} &= \mathsf{P}\{\omega : \exp tX(\omega) \not\leq I\} \\ &\leq \mathsf{E}1_{\{\exp tX \not\leq I\}} \mathrm{Tr}\exp tX \\ &\leq \mathrm{Tr}\,\mathsf{E}\exp tX,\end{aligned}$$

где первое неравенство вытекает из того, что $0 \leq y \not\leq I$ влечет $1 \leq \mathrm{Tr}\, y$.

Во-вторых, обозначая $X(\omega) = \sum_{i=1}^{N} X_i(\omega)$, где X_i – независимые одинаково распределенные эрмитово операторно-значные случайные величины, получаем

$$\begin{aligned}\mathsf{P}\left\{\omega : \sum_{i=1}^{N} X_i(\omega) \not\leq 0\right\} &\leq \mathrm{Tr}\,\mathsf{E}\exp\left(t\sum_{i=1}^{N} X_i\right) \\ &\leq \mathsf{E}\,\mathrm{Tr}\exp\left(t\sum_{i=1}^{N-1} X_i\right)\exp tX_N \\ &= \mathrm{Tr}\,\mathsf{E}\exp\left(t\sum_{i=1}^{N-1} X_i\right)\mathsf{E}\exp tX_N,\end{aligned}$$

где использовано неравенство Голдена-Томпсона

$$\operatorname{Tr}\exp(A+B)\le\operatorname{Tr}\exp A\exp B.$$

Согласно неравенству 1.18, это не превосходит

$$\operatorname{Tr}\mathsf{E}\exp\left(t\sum_{i=1}^{N-1}X_i\right)\|\mathsf{E}\exp tX_N\|.$$

Применяя эту оценку последовательно, получаем

$$\mathsf{P}\left\{\omega:\sum_{i=1}^{N}X_i(\omega)\not\le 0\right\}\le d\,\|\mathsf{E}\exp tX_i\|^N. \tag{9.74}$$

В-третьих, пусть $X_1(\omega),\dots,X_N(\omega)$ – независимые операторно-значные случайные величины, такие что $0\le X_i(\omega)\le I$, $\mathsf{E}X_i=\mu I$, и $0\le\mu\le a\le 1$, тогда

$$\mathsf{P}\left\{\omega:\frac{1}{N}\sum_{i=1}^{N}X_i(\omega)\not\le aI\right\}\le d\exp\left(-Nh(a;\mu)\right), \tag{9.75}$$

где $h(a;\mu)=a\ln\frac{a}{\mu}+(1-a)\ln\frac{1-a}{1-\mu}$. Действительно, используя (9.74) для величин $X_i(\omega)-aI$, получаем

$$\mathsf{P}\left\{\omega:\frac{1}{N}\sum_{i=1}^{N}X_i(\omega)\not\le aI\right\}\le d\,\|\mathsf{E}\exp tX_i\|^N\exp(-atN).$$

Однако $\exp tx\le 1+x(\exp t-1)$ для $x\in[0,1]$; следовательно, $\exp tX_i\le I+X_i(\exp t-1)$, и правая часть не превосходит

$$d\left[1+\mu(\exp t-1)\right]^N\exp(-atN).$$

Полагая $\exp t=\frac{a}{\mu}\frac{1-\mu}{1-a}\ge 1$, получаем величину в правой части (9.75). Аналогично, для $0\le a\le\mu\le 1$,

$$\mathsf{P}\left\{\omega:\frac{1}{N}\sum_{i=1}^{N}X_i(\omega)\not\ge aI\right\}\le d\exp\left(-Nh(a;\mu)\right). \tag{9.76}$$

Заметим, что $h(\mu;\mu)=0$, $\frac{\partial}{\partial a}h(a;\mu)|_{\mu=a}=0$, $\frac{\partial^2}{\partial^2 a}h(a;\mu)=\frac{1}{a(1-a)}$. Отсюда следует, что

$$\frac{\partial^2}{\partial^2\varepsilon}h(\mu(1\pm\varepsilon);\mu)=\frac{\mu}{(1\pm\varepsilon)(1-\mu(1\pm\varepsilon))}\ge\frac{\mu}{2},$$

если $\mu\le\frac{1}{2}$, $|\varepsilon|\le 1$, следовательно,

$$h(\mu\,(1\pm\varepsilon)\,;\mu)\geq\frac{\mu\varepsilon^2}{4}. \tag{9.77}$$

Наконец, пусть $X_i\,(\omega)$ удовлетворяют условиям теоремы 9.4.5. Рассмотрим операторные случайные величины $Y_i\,(\omega)=\mu M^{-1/2}X_i\,(\omega)\,M^{-1/2}$, так что $0\leq Y_i\,(\omega)\leq I$ и $\mathsf{E}Y_i=\mu I$. Имеем

$$\begin{aligned}
&\left\{\omega:\left\|\frac{1}{N}\sum_{i=1}^{N}X_i\,(\omega)-M\right\|_1<\varepsilon\operatorname{Tr}M\right\}\\
&\supseteq\left\{\omega:(1-\varepsilon)M\leq\frac{1}{N}\sum_{i=1}^{N}X_i\,(\omega)\leq(1+\varepsilon)M\right\}\\
&=\left\{\omega:(1-\varepsilon)\mu\leq\frac{1}{N}\sum_{i=1}^{N}Y_i\,(\omega)\leq(1+\varepsilon)\mu\right\}\\
&=\left\{\omega:(1-\varepsilon)\mu\leq\frac{1}{N}\sum_{i=1}^{N}Y_i\,(\omega)\right\}\cap\left\{\omega:\frac{1}{N}\sum_{i=1}^{N}Y_i\,(\omega)\leq(1+\varepsilon)\mu\right\}.
\end{aligned}$$

Беря дополнения и применяя оценки (9.75), (9.76) вместе с (9.77), получаем (9.73). □

9.4.4 Прямая теорема кодирования для квантовой пропускной способности

Доказательство неравенства $Q(\Phi)\geq\overline{Q}(\Phi)$ будет основано на "когерентной" версии доказательства для классической секретной пропускной способности, т. е. на версии, использующей то же случайное кодирование, однако с заменой смесей состояний их суперпозициями. Достаточно доказать, что для данного входного состояния S величина $R=I_c(S,\Phi)-\delta$ является достижимой скоростью; достижимость величины $\overline{Q}(\Phi)-\delta$ тогда доказывается применением тех же рассуждений к каналу $\Phi^{\otimes n}$ и последующим переходом к пределу $n\to\infty$. В этом разделе мы покажем, как следует строить кодирующие и декодирующие каналы, чтобы выполнялось

$$F_e\left(\bar{S}^{(n)},\mathcal{D}^{(n)}\circ\Phi^{\otimes n}\circ\mathcal{E}^{(n)}\right)\to 1, \tag{9.78}$$

где $\bar{S}^{(n)}$ – хаотическое состояние в $\mathcal{H}^{(n)},d_n=\dim\mathcal{H}^{(n)}=2^{nR}$. Применяя рассуждения, следующие за определением 9.3.1, получаем, что скорость R является достижимой для асимптотически безошибочной передачи квантовой информации.

Для канала Φ рассмотрим представление Стайнспринга (6.8) (теорема 6.2.1), а именно

$$\Phi\,[S]=\operatorname{Tr}_E VSV^*,$$

где V – изометричное отображение входного пространства $\mathcal{H}_A$ в тензорное произведение $\mathcal{H}_B \otimes \mathcal{H}_E$ выходного пространства и пространства окружения.

Рассмотрим спектральное разложение $S = \sum_x \pi_x S_A^x$, где $S_A^x = |\phi_x\rangle_A\langle\phi_x|$, и $\{|\phi_x\rangle_A\}$ – ортонормированный базис. Обозначим $|\phi'_x\rangle_{BE} = V|\phi_x\rangle_A$ векторы выходных состояний составной системы BE. Тогда состояния систем B и E равны, соответственно, $S_B^x = \mathrm{Tr}_E\,|\phi'_x\rangle_{BE}\langle\phi'_x|$ и $S_E^x = \mathrm{Tr}_B\,|\phi'_x\rangle_{BE}\langle\phi'_x|$. Напомним, что согласно (9.56)

$$I_c\left(S,\Phi\right) = \chi\left(\{\pi_x\},\{S_B^x\}\right) - \chi\left(\{\pi_x\},\{S_E^x\}\right). \tag{9.79}$$

Нашей целью является построение последовательности пространств $\mathcal{H}^{(n)}$, таких что $\dim\mathcal{H}^{(n)} = 2^{nR}$, $R = I_c\left(S,\Phi\right) - \delta$, а также соответствующих кодирований и декодирований, для которых выполняется (9.78).

Рассмотрим блочный канал $\Phi^{\otimes n}$ и набор кодовых слов

$$W^{(n)} = \left\{w^{mj}; m = 1,\ldots,N_B; j = 1,\ldots,N_E\right\},$$

используемый в конструкции секретного кода, с соответствующим разложением единицы $M^{(n)} = \{M_{mj}\}$ в $\mathcal{H}_B^{\otimes n}$, такие что

$$P_e(W^{(n)},M^{(n)}) \equiv \max_{mj}\left(1 - \mathrm{Tr}\,S_{B^{(n)}}^{mj}M_{mj}\right) < \varepsilon, \tag{9.80}$$

где $S_{B^{(n)}}^{mj} = S_B^{x_1}\otimes\cdots\otimes S_B^{x_n}$, если $w^{mj} = (x_1,\ldots,x_n)$,

$$\left\|\frac{1}{N_E}\sum_{j=1}^{N_E} S_{E^{(n)}}^{mj} - \theta\right\|_1 < \sqrt{\varepsilon}; \quad m = 1,\ldots,N_B. \tag{9.81}$$

Введем пространство $\mathcal{H}^{(n)}$ с ортонормированным базисом $\{|m\rangle; m = 1,\ldots,N_B\}$, где

$$N_B = 2^{n[\chi(\{\pi_x\},\{S_B^x\})-\chi(\{\pi_x\},\{S_B^x\})-\delta]} = 2^{n[I_c(S,\Phi)-\delta]},$$

и зададим кодирующий канал $\mathcal{E}^{(n)}$ следующим изометрическим отображением $\mathcal{H}^{(n)}$ в $\mathcal{H}_A^{\otimes n}$:

$$|m\rangle \to \frac{1}{\sqrt{N_E}}\sum_{j=1}^{N_E} e^{i\alpha_{mj}}|\phi_{mj}\rangle_A,$$

где $|\phi_{mj}\rangle_A = |\phi_{x_1}\rangle_A\otimes\cdots\otimes|\phi_{x_n}\rangle_A$, если $w^{mj} = (x_1,\ldots,x_n)$, и $\{\alpha_{mj}\} = \alpha$ – набор фаз, которые будут выбраны позднее. Изометричность этого отображения вытекает из того, что $\{|\phi_{mj}\rangle_A\}$ – ортонормированная система, потому что она построена из ортонормированных векторов $|\phi_x\rangle_A$.

Это и есть то, что называется когерентной версией секретного кода для классической информации. Действие кодирующего канала на вторые компоненты максимально сцепленного вектора (9.42) дает вектор

$$\frac{1}{\sqrt{N_B}}\sum_{m=1}^{N_B}|m\rangle\otimes\left[\frac{1}{\sqrt{N_E}}\sum_{j=1}^{N_E}e^{i\alpha_{mj}}|\phi_{mj}\rangle_A.\right]$$

Кодирующее отображение с последующим действием канала $\Phi^{\otimes n}$ дает

$$|m\rangle\to\frac{1}{\sqrt{N_B}}\sum_{m=1}^{N_B}|m\rangle\otimes\left[\frac{1}{\sqrt{N_E}}\sum_{j=1}^{N_E}e^{i\alpha_{mj}}|\phi'_{mj}\rangle_{BE}\right],$$

где векторы $|\phi'_{mj}\rangle_{BE}=V|\phi_{mj}\rangle_A$ вновь образуют ортонормированную систему. Заметим, что $S^{mj}_{B^{(n)}}=\mathrm{Tr}_E\,|\phi'_{mj}\rangle_{BE}\langle\phi'_{mj}|$, $S^{mj}_{E^{(n)}}=\mathrm{Tr}_B\,|\phi'_{mj}\rangle_{BE}\langle\phi'_{mj}|$, где для упрощения записи мы обозначаем через Tr_E частичный след по пространству $\mathcal{H}_E^{\otimes n}$ и т. д.

Для построения декодирующего отображения $\mathcal{D}^{(n)}$ мы сначала используем наблюдаемую $M^{(n)}=\{M_{mj}\}$, которая дает значения mj с асимптотически исчезающей ошибкой. Согласно описанию процесса измерения, приводящему к соотношению (6.34), мы имеем: вспомогательную систему $\mathcal{H}_0^{(n)}=\mathcal{H}^{(n)}\otimes\mathcal{H}_1^{(n)}$ с базисом $\{|m\rangle\otimes|j\rangle_1\}$, единичный вектор $|\psi_0\rangle\in\mathcal{H}_0^{(n)}$ и унитарный оператор U в $\mathcal{H}_B^{\otimes n}\otimes\mathcal{H}_0^{(n)}$ такие, что

$$M_{mj}=\mathrm{Tr}_0\left(I_{B^{(n)}}\otimes|\psi_0\rangle\langle\psi_0|\right)U^*\left(I_{B^{(n)}}\otimes|mj\rangle_0\langle mj|\right)U,$$

где используется обозначение $|mj\rangle_0=|m\rangle\otimes|j\rangle_1$. Определяя

$$|\psi_{mj}\rangle_{BE0}=\left(I_{E^{(n)}}\otimes U\right)\left(|\phi'_{mj}\rangle_{BE}\otimes|\psi_0\rangle\right),$$

получаем следующий вектор состояния системы $\mathcal{H}^{(n)}\otimes\mathcal{H}_B^{\otimes n}\otimes\mathcal{H}_E^{\otimes n}\otimes\mathcal{H}_0^{(n)}$

$$|\Upsilon(\alpha)\rangle=\frac{1}{\sqrt{N_B}}\sum_{m=1}^{N_B}|m\rangle\otimes\left[\frac{1}{\sqrt{N_E}}\sum_{j=1}^{N_E}e^{i\alpha_{mj}}|\psi_{mj}\rangle_{BE0}\right],$$

как результат действия на максимально сцепленный вектор композиции кодирующего канала, блочного канала и, наконец, измерения. Здесь $\{|\psi_{mj}\rangle_{BE0}\}$ вновь является ортонормированной системой. Соотношение (9.80) означает, что проекции векторов этой системы на $\mathcal{H}_0^{(n)}$ близки к $\{|mj\rangle_0\}$, а именно,

$${}_0\langle mj|\left(\mathrm{Tr}_{BE}\,|\psi_{mj}\rangle_{BE0}\langle\psi_{mj}|\right)|mj\rangle_0>1-\varepsilon\quad\text{для всех } mj.$$

Из следующей леммы вытекает, что найдутся единичные векторы $|\chi_{mj}\rangle_{BE}\in\mathcal{H}_B^{\otimes n}\otimes\mathcal{H}_E^{\otimes n}$, такие что для $|\tilde{\psi}_{mj}\rangle_{BE0}=|\chi_{mj}\rangle_{BE}\otimes|mj\rangle_0$ выполняется

$$\left|{}_{BE0}\langle\psi_{mj}|\tilde{\psi}_{mj}\rangle_{BE0}\right|^2>1-\varepsilon. \tag{9.82}$$

Лемма 9.4.6 *Пусть* $|\varphi\rangle_0 \in \mathcal{H}_0, |\psi\rangle \in \mathcal{H}_0 \otimes \mathcal{H}$ *– единичные векторы, тогда*

$$ {}_0\langle\varphi| \left(\mathrm{Tr}_{\mathcal{H}} |\psi\rangle\langle\psi|\right) |\varphi\rangle_0 = \max_{\chi} |\langle\psi|(|\varphi\rangle_0 \otimes |\chi\rangle)|^2 , \tag{9.83} $$

где максимум берется по всем единичным векторам $|\chi\rangle \in \mathcal{H}$.

Доказательство. Фиксируя ортонормированный базис $\{|e_j\rangle\}$ в $\mathcal{H}$, имеем $|\chi\rangle = \sum_j c_j |e_j\rangle$ с $\sum_j |c_j|^2 = 1$ и

$$ |\langle\psi|(|\varphi\rangle_0 \otimes |\chi\rangle)|^2 = \left| \sum_j c_j \langle\psi|(|\varphi\rangle_0 \otimes |e_j\rangle) \right|^2 . $$

Максимизируя по c_j, получаем (9.83). □

Заметим, что $\left\{|\tilde{\psi}_{mj}\rangle_{BE0}\right\}$ также является ортонормированной системой, аппроксимирующей систему $\{|\psi_{mj}\rangle_{BE0}\}$ в смысле (9.82). Определяя вектор

$$ |\Gamma(\alpha)\rangle = \frac{1}{\sqrt{N_B}} \sum_{m=1}^{N_B} |m\rangle \otimes \left[\frac{1}{\sqrt{N_E}} \sum_{j=1}^{N_E} e^{i\alpha_{mj}} |\tilde{\psi}_{mj}\rangle_{BE0} \right] , $$

имеем

$$ \int \ldots \int \langle\Upsilon(\alpha)|\Gamma(\alpha+\alpha')\rangle d\mu(\alpha) = \frac{1}{N_B N_E} \sum_{m=1}^{N_B} \sum_{j=1}^{N_E} e^{i\alpha'_{mj}} {}_{BE0}\langle\psi_{mj}|\tilde{\psi}_{mj}\rangle_{BE0}, $$

где $d\mu(\alpha)$ – равномерное распределение. Всегда можно выбрать набор фаз $\alpha' = \left\{\alpha'_{mj}\right\}$ таким образом, что правая часть будет равна

$$ \frac{1}{N_B N_E} \sum_{m=1}^{N_B} \sum_{j=1}^{N_E} \left| {}_{BE0}\langle\psi_{mj}|\tilde{\psi}_{mj}\rangle_{BE0} \right| > \sqrt{1-\varepsilon} $$

согласно (9.82). Следовательно, найдется α, такое что $\Re\langle\Upsilon(\alpha)|\Gamma(\alpha+\alpha')\rangle > \sqrt{1-\varepsilon}$. Вводя дополнительную общую фазу в α'_{mj}, всегда можно сделать скалярное произведение положительным, так что

$$ \langle\Upsilon(\alpha)|\Gamma(\alpha+\alpha')\rangle > \sqrt{1-\varepsilon}. \tag{9.84} $$

В декодирующий канал следует добавить еще одно преобразование, которое приведет систему E в состояние, почти не зависящее от m. Чтобы добиться этого, заметим, что комбинация (9.82) с (9.17) дает

$$ \left\| |\psi_{mj}\rangle_{BE0}\langle\psi_{mj}| - |\tilde{\psi}_{mj}\rangle_{BE0}\langle\tilde{\psi}_{mj}| \right\|_1 < 2\sqrt{\varepsilon}. \tag{9.85} $$

Теперь заметим, что $S^{mj}_{E^{(n)}} = \mathrm{Tr}_{B0}\,|\psi_{mj}\rangle_{BE0}\langle\psi_{mj}|$ и обозначим

$$\tilde{S}^{mj}_{E^{(n)}} = \mathrm{Tr}_{B0}\,|\tilde{\psi}_{mj}\rangle_{BE0}\langle\tilde{\psi}_{mj}| = \mathrm{Tr}_{B}\,|\chi_{mj}\rangle_{BE}\langle\chi_{mj}|.$$

Тогда из (9.85) и задачи 9.3.1 вытекает

$$\left\| S^{mj}_{E^{(n)}} - \tilde{S}^{mj}_{E^{(n)}} \right\|_1 < 2\sqrt{\varepsilon}, \tag{9.86}$$

что в сочетании с (9.81) дает

$$\left\| \frac{1}{N_E} \sum_{j=1}^{N_E} \tilde{S}^{mj}_{E^{(n)}} - \theta \right\|_1 < O\left(\sqrt{\varepsilon}\right); \quad m = 1, \ldots, N_B. \tag{9.87}$$

Рассмотрим векторы

$$|\varphi_m\rangle_{BE1} = \frac{1}{\sqrt{N_E}} \sum_{j=1}^{N_E} e^{i\left(\alpha_{mj}+\alpha'_{mj}\right)} |\chi_{mj}\rangle_{BE} \otimes |j\rangle_1,$$

являющиеся очищениями состояний $\frac{1}{N_E}\sum_{j=1}^{N_E} \tilde{S}^{mj}_{E^{(n)}}$. Пусть $|\phi_\theta\rangle \in \mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_E^{\otimes n} \otimes \mathcal{H}_1^{(n)}$ – очищение состояния θ, тогда из (9.87) следует, что найдутся унитарные операторы W_m в $\mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_1^{(n)}$, такие что

$$\left| {}_{BE1}\langle\phi_\theta| \left(I_{\mathcal{H}_E^{\otimes n}} \otimes W_m \right) \varphi_m\rangle_{BE1} \right| > 1 - O\left(\varepsilon\right) \text{ для всех } m. \tag{9.88}$$

Это вытекает из неравенства (9.33): если два состояния близки в следовой норме, то найдутся очищения этих состояний, для которых точность воспроизведения близка к 1.

Определяя “контролируемый унитарный оператор”

$$W = \sum_{m=1}^{N_B} e^{i\beta_m} W_m \otimes |m\rangle\langle m|$$

в $\mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_0^{(n)}$, где β_m – фазы, которые будут подобраны ниже, получаем

$$\left(I_{\mathcal{H}_E^{\otimes n}} \otimes W \right) \left(|\varphi_m\rangle_{BE1} \otimes |m\rangle \right) = e^{i\beta_m} \left(I_{\mathcal{H}_E^{\otimes n}} \otimes W_m \right) |\varphi_m\rangle_{BE1} \otimes |m\rangle,$$

следовательно,

$$\left(I_{\mathcal{H}^{(n)}} \otimes I_{\mathcal{H}_E^{\otimes n}} \otimes W \right) |\Gamma(\alpha + \alpha')\rangle$$

$$= \frac{1}{\sqrt{N_B}} \sum_{m=1}^{N_B} |m\rangle \otimes |m\rangle \otimes e^{i\beta_m} \left(I_{\mathcal{H}_E^{\otimes n}} \otimes W_m \right) |\varphi_m\rangle_{BE1}.$$

Обозначая $|\Omega^{(n)}\rangle = \frac{1}{\sqrt{N_B}} \sum_{m=1}^{N_B} |m\rangle \otimes |m\rangle$ максимально сцепленный вектор в $\mathcal{H}^{(n)} \otimes \mathcal{H}^{(n)}$, имеем

$$\left(\langle\Omega^{(n)}|\otimes\langle\phi_\theta|\right)\left(I_{\mathcal{H}^{(n)}}\otimes I_{\mathcal{H}_E^{\otimes n}}\otimes W\right)|\Gamma(\alpha+\alpha')\rangle$$

$$=\frac{1}{N_B}\sum_{m=1}^{N_B}e^{i\beta_m}\langle\phi_\theta|\left(I_{\mathcal{H}_E^{\otimes n}}\otimes W_m\right)|\varphi_m\rangle_{BE1},$$

что может быть сделано равным

$$\frac{1}{N_B}\sum_{m=1}^{N_B}\left|\langle\phi_\theta|\left(I_{\mathcal{H}_E^{\otimes n}}\otimes W_m\right)|\varphi_m\rangle_{BE1}\right|$$

при соответствующем выборе фаз β_m. Последнее выражение больше, чем $1-O(\varepsilon)$ согласно оценке (9.88). Сравнивая это с (9.84), получаем

$$\left(\langle\Omega^{(n)}|\otimes\langle\phi_\theta|\right)\left(I_{\mathcal{H}^{(n)}}\otimes I_{\mathcal{H}_E^{\otimes n}}\otimes W\right)|\Upsilon(\alpha)\rangle>1-O(\varepsilon),$$

что следует из неравенства треугольника (9.31).

Вводя оператор плотности

$$S'=\mathrm{Tr}_{BE1}\left(I_{\mathcal{H}^{(n)}}\otimes I_{\mathcal{H}_E^{\otimes n}}\otimes W\right)|\Upsilon(\alpha)\rangle\langle\Upsilon(\alpha)|\left(I_{\mathcal{H}^{(n)}}\otimes I_{\mathcal{H}_E^{\otimes n}}\otimes W\right)$$

в $\mathcal{H}^{(n)}\otimes\mathcal{H}^{(n)}$, имеем

$$\begin{aligned}F\left(|\Omega^{(n)}\rangle\langle\Omega^{(n)}|,S'\right)&\equiv\langle\Omega^{(n)}|S'|\Omega^{(n)}\rangle\\&\geq\left|\left(\langle\Omega^{(n)}|\otimes\langle\phi_\theta|\right)\left(I_{\mathcal{H}^{(n)}}\otimes I_{\mathcal{H}_E^{\otimes n}}\otimes W\right)|\Upsilon(\alpha)\rangle\right|^2\\&>1-O(\varepsilon).\end{aligned}\tag{9.89}$$

С другой стороны,

$$S'=\left(\mathrm{Id}_{\mathcal{H}^{(n)}}\otimes\mathcal{D}^{(n)}\circ\Phi^{\otimes n}\circ\mathcal{E}^{(n)}\right)\left[|\Omega^{(n)}\rangle\langle\Omega^{(n)}|\right],$$

где декодирующий канал $\mathcal{D}^{(n)}$ определяется как

$$\mathcal{D}^{(n)}\left[S_{B^{(n)}}\right]=\mathrm{Tr}_{B1}\,WU\left(S_{B^{(n)}}\otimes|\psi_0\rangle\langle\psi_0|\right)U^*W^*$$

(напомним, что $|\psi_0\rangle\in\mathcal{H}_0^{(n)}=\mathcal{H}^{(n)}\otimes\mathcal{H}_1^{(n)}$). Тогда неравенство (9.89) означает то же, что и

$$F_e\left(\bar{S}^{(n)},\mathcal{D}^{(n)}\circ\Phi^{\otimes n}\circ\mathcal{E}^{(n)}\right)>1-O(\varepsilon),$$

и утверждение (9.78) доказано. □

9.5 Комментарии

1. Примеры квантовых кодов, исправляющих ошибки, были построены независимо в работах Шора [68] и Стина [148]. Многие авторы сделали вклад в последующее развитие теории, фрагменты которой представлены в этой главе, см. обзор в книге Нильсена и Чанга [22]. Необходимые и достаточные условия исправления ошибок в теореме 9.1.1 были предложены Книллом и Лафламмом [116]. Проверку утверждения задачи 9.1.1 см. в [22], п. 10.2.

Возможность исправления ошибок является принципиальной для проблемы создания квантового компьютера, см. Валиев и Кокин [2]. Была предложена архитектура квантового компьютера, исправляющего ошибки не только в квантовой памяти, но и в самой схеме, исправляющей ошибки, при условии, что вероятность ошибки не превосходит некоторого порогового значения. По поводу квантовых вычислений, устойчивых к ошибкам, см. Китаев [11], Стин [148].

На роль когерентной информации в точном исправлении ошибок указано Барнумом, Нильсеном и Шумахером [56].

2. Соотношения между мерами точности исследовались Барнумом, Нильсеном и Шумахером [56] а также Вернером и Кречманом [118]. Точность воспроизведения для произвольных состояний исследовалась Ульманом [150]. Лемма 9.2.3 доказана в работе Холево [32].

3. Теорема кодирования для квантовой пропускной способности была впервые сформулирована Ллойдом [127] в виде $Q(\Phi) = \max_S I_c(S, \Phi)$, но затем было показано, что величина $\bar{Q}_n$ может быть строго супераддитивна (Ди Винченцо, Шор и Смолин [80], Смит и Смолин [147]), следовательно, использование регуляризованного предельного выражения в (9.40) в общем случае действительно необходимо.

Неравенство $Q(\Phi) \leq \overline{Q}(\Phi)$ (т. е. обращение теоремы кодирования) было доказано Барнумом, Нильсеном и Шумахером [56], Барнумом, Книллом и Нильсеном [55]. Дополнительные аргументы в пользу эвристического подхода Ллойда – использования случайных подпространств – были приведены Шором. Доказательство, основанное на этом подходе, было дано впоследствии Хайденом [95].

Соотношение (9.56) получено Шумахером и Вестморлендом [137], которые, основываясь на идеях классической теории информации [51], связали "секретность" передачи классической информации через квантовый канал Φ с величиной $\chi\left(\{\pi_x\}, \{S_B^x\}\right) - \chi\left(\{\pi_x\}, \{S_E^x\}\right)$.

Понятие деградируемого канала введено в работе Деветака и Шора [79], где было показано, что квантовая пропускная способность деградируемого канала дается однобуквенным выражением $Q(\Phi) = \max_S I_c(S, \Phi)$. Это напоминает понятие "ухудшенного" (stochastically degraded) канала [72] для классических широковещательных каналов, причем роль второго приемника играет окружение. Для таких каналов в теории информации

получено однобуквенное выражение для области пропускных способностей. Анти-деградируемые каналы рассматривались Карузо, Джованнетти и Холево [69]. Следствие 9.3.1 получено в работе Кубитта, Рускаи и Смита [73].

4. Полное доказательство прямой теоремы кодирования $\big(Q(\Phi) \geq \overline{Q}(\Phi)\big)$ дал Деветак [78]. Оно основано на идее, связанной с соотношением между квантовой и секретной классической пропускными способностями, см. (9.56). Именно это доказательство (а также более простое доказательство обратного утверждения) излагается нами (с несколькими упрощениями). Оценка вероятностей больших уклонений типа Бернштейна, использующая операторную версию (9.75) неравенства Хеффдинга ([98], теорема 1), принадлежит Винтеру [157]. Другое доказательство, основанное на методе случайных подпространств, идейно близком к первоначальным соображениям Ллойда, дано Хайденом, Шором и Винтером [96].

Как уже отмечалось, квантовый канал характеризуется целым спектром пропускных способностей. Помимо величин $C, C_\chi, C_{ea}, C_p, Q$, подробно рассмотренных в настоящей книге, в квантовой теории информации изучаются классическая и квантовая пропускные способности с обратной связью (обозначаемые, соответственно, C_b, Q_b), а также классическая и квантовая пропускные способности с дополнительным независимым классическим двусторонним каналом (соответственно, C_2, Q_2). Отметим, что в классическом случае $C = C_b = C_2 = C_{Shan}$ и шенноновская пропускная способность остается единственной характеристикой. Для квантового канала имеет место следующая иерархия

$$\begin{array}{ccccccc} C_\chi & \leq & C_b & \leq & C_2 & \leq & C_{ea} \\ \mathrm{VI} & & \mathrm{VI} & & \mathrm{VI} & & \mathrm{VI} \\ Q & \leq & Q_b & \leq & Q_2 & \leq & Q_{ea} \end{array},$$

где $\leq$ следует понимать как “меньше или равно для всех каналов и строго меньше для некоторых”, см. Беннет, Деветак, Шор, Смолин [59]. Известно также, что $C_{ea} = 2Q_{ea}$ и для ряда остальных пар возможны неравенства как в ту, так и в другую стороны. Далее, оказывается возможным построить так называемый “материнский” протокол передачи через квантовый канал, который при использовании различных частных дополнительных ресурсов (таких, например, как обратная связь либо сцепленность) позволяет реализовать всевозможные способы передачи, включая упомянутые выше [49].

10. Каналы с ограничениями на входе

Одним из главных вопросов, мотивировавших появление квантовой теории информации, был вопрос о пропускной способности оптического квантового канала с ограничением на среднюю энергию входного сигнала. С математической точки зрения, необходимость введения ограничений возникает, когда квантовая система, несущая информацию, описывается бесконечномерным гильбертовым пространством, что влечет необходимость использования неограниченных операторов с присущими усложнениями. Важный класс таких систем, иногда называемых системами с "непрерывными переменными", а также каналов, называемых (бозонными) гауссовскими, будет рассмотрен в следующей главе. Одной из конечных целей настоящей главы является получение общих выражений для пропускных способностей бесконечномерных каналов с ограничениями на входе, удобных для использования в случае квантовых гауссовских каналов с ограничением на среднюю энергию сигнала.

Особенностью бесконечномерного случая является разрывность и неограниченность энтропии квантовых состояний, что заставляет уделить серьезное внимание изучению непрерывности энтропийных характеристик. Другой характерной чертой является естественное появление "непрерывных"ансамблей, которые описываются вероятностными мерами на множестве квантовых состояний. Различные характеристики квантового канала, такие как пропускная способность, выражаются в виде точных верхних, либо нижних, граней функционалов на множестве обобщенных ансамблей, удовлетворяющих соответствующему ограничению. Эти точные грани оказываются достижимыми при выполнении определенных условий, а именно – непрерывности рассматриваемой характеристики и компактности множества, определяемого ограничением. Эти две математические проблемы и составляют основной предмет настоящей главы.

10.1 О сходимости квантовых состояний

Всюду далее $\mathcal{H}$ – сепарабельное гильбертово пространство. Линейный оператор A в пространстве $\mathcal{H}$ называется *ограниченным*, если он отображает единичный шар в ограниченное по норме подмножество пространства $\mathcal{H}$. Многие определения главы 2 переносятся на ограниченные опера-

торы; там, где необходимы модификации, будут даны соответствующие пояснения. Так, норма оператора $\|A\|$ определяется выражением (1.16), в котором максимум следует заменить на супремум. Оператор A ограничен тогда и только тогда, когда $\|A\| < \infty$. Множество всех ограниченных операторов образует банахову алгебру $\mathfrak{B}(\mathcal{H})$. Следовая (ядерная) норма $\|T\|_1$ определяется выражением (1.14), и множество ограниченных операторов, для которых $\|T\|_1 < \infty$, образует банахово пространство $\mathfrak{T}(\mathcal{H})$ *ядерных* операторов. Для любых операторов $T \in \mathfrak{T}(\mathcal{H}), A \in \mathfrak{B}(\mathcal{H})$ имеет место неравенство

$$|\mathrm{Tr}TA| \leq \|T\|_1 \|A\|; \tag{10.1}$$

банахово пространство $\mathfrak{B}(\mathcal{H})$ является сопряженным к $\mathfrak{T}(\mathcal{H})$, т. е. любой непрерывный линейный функционал на $\mathfrak{T}(\mathcal{H})$ имеет вид $T \to \mathrm{Tr}\,TA$ при некотором $A \in \mathfrak{B}(\mathcal{H})$, а его норма равна $\|A\|$. Заметим также, что

$$\|A\| \leq \|A\|_1 \tag{10.2}$$

и что $\mathfrak{T}(\mathcal{H})$ является собственным подпространством $\mathfrak{B}(\mathcal{H})$.

Оператором плотности (состоянием) называется положительный ядерный оператор с единичным следом (ср. определение 2.1.1). Пространство состояний $\mathfrak{S}(\mathcal{H})$ – замкнутое выпуклое подмножество банахова пространства $\mathfrak{T}(\mathcal{H})$, которое является полным сепарабельным метрическим пространством с метрикой, определяемой следовой нормой: $\rho(S_1, S_2) = \|S_1 - S_2\|_1$.

Последовательность ограниченных операторов $\{A_n\}$ в $\mathcal{H}$ сходится *слабо* к оператору A, если $\lim_{n\to\infty}\langle\psi|\, A_n|\phi\rangle \to \langle\psi|\, A|\phi\rangle$ для всех $\phi, \psi \in \mathcal{H}$. Хотя в пространстве $\mathfrak{T}(\mathcal{H})$ такая сходимость слабее сходимости по следовой норме, на подмножестве $\mathfrak{S}(\mathcal{H})$ эти два вида сходимости оказываются эквивалентными.

Лемма 10.1.1 *Пусть $\{S_n\}$ – последовательность операторов плотности в $\mathcal{H}$, слабо сходящаяся к оператору плотности S, тогда $\|S_n - S\|_1 \to 0$.*

Доказательство. Для любого конечномерного проектора P

$$\begin{aligned}\|S_n - S\|_1 \leq \|P(S_n - S)P\|_1 + 2\,\|PS_n(I-P)\|_1 + 2\,\|PS(I-P)\|_1 \\ + \|(I-P)S_n(I-P)\|_1 + \|(I-P)S(I-P)\|_1\,.\end{aligned}$$

Первое слагаемое в правой части стремится к нулю при любом выборе проектора P, поскольку $PS_nP \to PSP$ в силу слабой сходимости и эквивалентности всех видов сходимости в конечномерном случае. Для двух последних слагаемых имеем

$$\|(I-P)S(I-P)\|_1 = \mathrm{Tr}(I-P)S(I-P) = \mathrm{Tr}(I-P)S = 1 - \mathrm{Tr}PS,$$

что можно сделать сколь угодно малым путем выбора проектора P, и

$$\|(I-P)S_n(I-P)\|_1 = 1 - \mathrm{Tr}PS_n \to 1 - \mathrm{Tr}PS$$

в силу слабой сходимости. Среднее слагаемое можно оценить следующим образом:

$$\|PS_n(I-P)\|_1 = \mathrm{Tr}U^*PS_n(I-P) = \mathrm{Tr}U^*P\sqrt{S_n}\sqrt{S_n}(I-P),$$

где U – унитарный оператор из полярного разложения оператора $PS_n(I-P)$. В силу операторного неравенства Коши-Буняковского для следа имеем

$$\begin{aligned}\mathrm{Tr}U^*P\sqrt{S_n}\sqrt{S_n}(I-P) &\le \sqrt{\mathrm{Tr}U^*P\sqrt{S_n}\sqrt{S_n}PU}\sqrt{\mathrm{Tr}(I-P)\sqrt{S_n}\sqrt{S_n}(I-P)} \\ &= \sqrt{\mathrm{Tr}PS_n}\sqrt{1-\mathrm{Tr}PS_n} \to \sqrt{\mathrm{Tr}PS}\sqrt{1-\mathrm{Tr}PS}\end{aligned}$$

что также можно сделать сколь угодно малым путем выбора проектора P. □

В дальнейшем, говоря о сходимости состояний, мы будем иметь в виду сходимость, о которой идет речь в этой лемме.

Теорема 10.1.2 *Замкнутое по следовой норме подмножество K множества состояний $\mathfrak{S}(\mathcal{H})$ компактно тогда и только тогда, когда для любого $\varepsilon > 0$ найдется конечномерный проектор P_ε такой, что $\mathrm{Tr}\,P_\varepsilon S > 1-\varepsilon$ для всех $S \in K$.*

Доказательство. Пусть K – компактное подмножество $\mathfrak{S}(\mathcal{H})$. Предположим, что найдется $\varepsilon > 0$ такое, что для любого конечномерного проектора P существует состояние $S \in K$, для которого $\mathrm{Tr}PS \le 1-\varepsilon$. Пусть $\{P_n\}$ – монотонно возрастающая последовательность конечномерных проекторов в $\mathcal{H}$, слабо сходящаяся к единичному оператору I, и пусть S_n – соответствующая последовательность состояний из K. В силу компактности множества K найдется ее подпоследовательность $\{S_{n_k}\}$, сходящаяся к состоянию $S_0 \in K$. По построению $\mathrm{Tr}P_{n_l}S_{n_k} \le \mathrm{Tr}P_{n_k}S_{n_k} \le 1-\varepsilon$ при $k > l$. Следовательно,

$$\mathrm{Tr}S_0 = \lim_{l\to+\infty}\mathrm{Tr}P_{n_l}S_0 = \lim_{l\to+\infty}\lim_{k\to+\infty}\mathrm{Tr}P_{n_l}S_{n_k} \le 1-\varepsilon,$$

что противоречит тому, что $S_0 \in K \subseteq \mathfrak{S}(\mathcal{H})$.

Обратно, пусть K – замкнутое подмножество множества $\mathfrak{S}(\mathcal{H})$, удовлетворяющее условию теоремы, и пусть $\{S_n\}$ – произвольная последовательность из K. Поскольку единичный шар в $\mathfrak{B}(\mathcal{H})$ слабо компактен, найдется подпоследовательность $\{S_{n_k}\}$, слабо сходящаяся к ограниченному положительному оператору S_0. Имеем

$$\mathrm{Tr}S_0 \le \liminf_{k\to\infty}\mathrm{Tr}S_{n_k} = 1,$$

поэтому для доказательства того, что S_0 – состояние, достаточно установить, что $\mathrm{Tr}S_0 \geq 1$. Пусть $\varepsilon > 0$ и P_ε – соответствующий проектор. Имеем

$$\mathrm{Tr}S_0 \geq \mathrm{Tr}P_\varepsilon S_0 = \lim_{k\to\infty} \mathrm{Tr}P_\varepsilon S_{n_k} > 1 - \varepsilon,$$

где равенство следует из конечномерности проектора P_ε. Таким образом, S_0 имеет единичный след и является состоянием. Лемма 10.1.1 влечет сходимость подпоследовательности $\{S_{n_k}\}$ к состоянию S_0 по следовой норме. Поэтому множество K компактно в топологии следовой нормы. □

Далее нам потребуется следующая частичная бесконечномерная версия спектрального разложения (1.5). Пусть $\{|e_j\rangle\}$ – ортонормированный базис в $\mathcal{H}$, $\{f_j\}$ – последовательность вещественных чисел, ограниченная снизу, тогда формула

$$F|\psi\rangle = \sum_j f_j |e_j\rangle\langle e_j|\psi\rangle \tag{10.3}$$

определяет самосопряженный полуограниченный снизу (см. раздел 10.8) оператор F на плотной области

$$\mathcal{D}(F) = \{\psi : \sum_j |f_j|^2 |\langle e_j|\psi\rangle|^2 < \infty\}, \tag{10.4}$$

для которого $|e_j\rangle$ являются собственными векторами, а f_j – собственными числами.

Определение 10.1.1 *Оператор с областью определения (10.4), действующий по формуле (10.3), будем называть оператором типа* $\mathfrak{F}$.

В приложениях F является оператором энергии (системы осцилляторов). В этой связи важную роль будет играть оператор $\exp(-\theta F), \theta > 0$, который определяется соотношением

$$\exp(-\theta F)|\psi\rangle = \sum_j \exp(-\theta f_j)|e_j\rangle\langle e_j|\psi\rangle, \tag{10.5}$$

который является ограниченным положительным оператором типа $\mathfrak{F}$.

Пусть F – оператор типа $\mathfrak{F}$. Для произвольного оператора плотности S определим математическое ожидание

$$\mathrm{Tr}SF = \sum_{j=1}^{\infty} f_j \langle e_j|S|e_j\rangle \leq +\infty. \tag{10.6}$$

Лемма 10.1.3 *Функционал* $S \to \mathrm{Tr}SF$ *является аффинным и полунепрерывным снизу на множестве* $\mathfrak{S}(\mathcal{H})$.

Доказательство. Аффинность очевидна; далее используем полунепрерывность снизу точной верхней грани семейства непрерывных функций. Имеем

$$\mathrm{Tr}SF = \sup_n \sum_{j=1}^{n} f_j \langle e_j|S|e_j\rangle,$$

где все конечные суммы непрерывны. □

Лемма 10.1.4 *Пусть спектр оператора F типа $\mathfrak{F}$ состоит из собственных значений f_n конечной кратности и $\lim_{n\to\infty} f_n = +\infty$, тогда множество*

$$\mathfrak{S}_E = \{S : \mathrm{Tr}SF \le E\} \tag{10.7}$$

компактно.

Доказательство. Без ограничения общности можно считать, что последовательность f_n монотонно возрастает. Пусть P_n – конечномерный проектор на собственное подпространство, соответствующее первым n собственным значениям, тогда $P_n \uparrow I$. Множество $\mathfrak{S}_E$ замкнуто, поскольку $\mathrm{Tr}SF \le \liminf_{N\to\infty} \mathrm{Tr}S_N F$ для любой последовательности $\{S_N\}$, сходящейся к S, в силу леммы 10.1.3. Поскольку $f_{n+1}(I - P_n) \le F$, имеем $\mathrm{Tr}S(I - P_n) \le f_{n+1}^{-1}\mathrm{Tr}SF \le f_{n+1}^{-1}E < \varepsilon$ для достаточно больших n и для всех $S \in \mathfrak{S}_E$. В силу критерия компактности множество $\mathfrak{S}_E$ компактно. □

В приложениях, где F является оператором энергии системы осцилляторов, полученный результат означает компактность подмножества квантовых состояний с ограниченной средней энергией.

10.2 Квантовая энтропия и относительная энтропия

Энтропия $H(S)$ квантового состояния S по-прежнему определяется формулой (5.7), однако в бесконечномерном случае она может принимать значение $+\infty$. Относительная энтропия корректно определяется формулой (7.2). Большинство свойств, полученных в главе 4.1, обобщается на бесконечномерный случай, однако вместо непрерывности имеет место лишь полунепрерывность снизу.

Теорема 10.2.1 *Квантовая энтропия и относительная энтропия полунепрерывны снизу на множестве состояний $\mathfrak{S}(\mathcal{H})$: пусть $\{S_n\}$ (соответственно $\{S'_n\}$) – последовательность операторов плотности в $\mathcal{H}$, сходящаяся к оператору плотности S (соответственно S'), тогда*

$$H(S) \le \lim\inf_{n\to\infty} H(S_n),$$

$$H(S;S') \le \lim\inf_{n\to\infty} H(S_n;S'_n).$$

Доказательство. В силу неравенства (10.2) имеем $\|S_n - S\| \to 0$. Функция $\eta(x) = -x\log x$ непрерывна на интервале $[0,1]$, следовательно $\|\eta(S_n) - \eta(S)\| \to 0$. Действительно, функцию $\eta(x)$ можно приблизить многочленами равномерно на $[0,1]$ и затем использовать следующую оценку:

Задача 10.2.1 *Для многочлена p и произвольных операторов A, B, норма которых не превосходит единицу, имеет место неравенство*

$$\|p(A) - p(B)\| \le c_p \|A - B\|,$$

в котором константа c_p зависит только от p. Указание: использовать разложение

$$A^k - B^k = \sum_{l=0}^{k-1} A^{k-l-1}(A-B)B^l.$$

Таким образом, для любого конечномерного проектора P в силу неравенства (10.1) имеем

$$|\mathrm{Tr}P(\eta(S_n) - \eta(S))| \le \mathrm{Tr}P\,\|\eta(S_n) - \eta(S)\| \to 0. \tag{10.8}$$

С другой стороны, в силу того же неравенства, для эрмитова положительного оператора A

$$\mathrm{Tr}PA \le \|P\|\,\mathrm{Tr}A$$

и, следовательно, $\mathrm{Tr}A = \sup_P \mathrm{Tr}PA$, где P пробегает множество всех конечномерных проекторов. Таким образом,

$$H(S) = \sup_P \mathrm{Tr}P\eta(S) \le \lim\inf_{n\to\infty} \sup_P \mathrm{Tr}P\eta(S_n) = \lim\inf_{n\to\infty} H(S_n).$$

Аналогично, используя бесконечномерный аналог представления (7.19) для относительной энтропии, получаем

$$H(S;S') = \sup_{\lambda>0} \frac{1}{\lambda}\left[H(\lambda S + (1-\lambda)S') - \lambda H(S) - (1-\lambda)H(S')\right],$$

следовательно,

$$\begin{aligned} H(S;S') &= \sup_{P,\lambda>0} \frac{1}{\lambda}\mathrm{Tr}P\left[\eta(\lambda S + (1-\lambda)S') - \lambda\eta(S) - (1-\lambda)\eta(S')\right] \\ &\le \lim\inf_{n\to\infty} \sup_{P,\lambda>0} \frac{1}{\lambda}\mathrm{Tr}P\left[\eta(\lambda S_n + (1-\lambda)S'_n) - \lambda\eta(S_n) - (1-\lambda)\eta(S'_n)\right] \\ &= \lim\inf_{n\to\infty} H(S_n;S'_n). \end{aligned}$$

□

Лемма 10.2.2 *Пусть F – оператор типа $\mathfrak{F}$, удовлетворяющий условию*

$$\mathrm{Tr}\exp(-\theta F) < \infty \qquad \textit{для} \quad \theta > 0, \tag{10.9}$$

тогда квантовая энтропия $H(S)$ ограничена и непрерывна на множестве $\mathfrak{S}_E$, определенном соотношением (10.7).

Заметим, что из (10.9) следует, что оператор F удовлетворяет условиям леммы 10.1.4, а значит множество $\mathfrak{S}_E$ компактно.

Доказательство. Вводя оператор плотности $S_\theta = \exp(-\theta F - c(\theta))$, где $c(\theta) = \log \mathrm{Tr}\exp(-\theta F)$, имеем

$$H(S) = -H(S; S_\theta) + \theta \mathrm{Tr} SF + c(\theta), \tag{10.10}$$

где $H(S; S_\theta)$ – относительная энтропия, следовательно,

$$H(S) \le -H(S; S_\theta) + \theta E + c(\theta), \tag{10.11}$$

если $S \in \mathfrak{S}_E$. Поэтому энтропия $H(S)$ ограничена на $\mathfrak{S}_E$. Поскольку функция $H(S)$ полунепрерывна снизу в силу теоремы 10.2.1, достаточно доказать ее полунепрерывность сверху на множестве $\mathfrak{S}_E$. Пусть $\{S_n\} \subset \mathfrak{S}_E$ – последовательность состояний, сходящаяся к состоянию S. Из (10.11), в силу полунепрерывности снизу относительной энтропии, следует

$$\begin{aligned} \limsup_{n\to\infty} H(S_n) &\le -H(S; S_\theta) + \theta E + c(\theta) \\ &\le H(S) + \theta E. \end{aligned} \tag{10.12}$$

Устремляя $\theta \to 0$, получаем полунепрерывность сверху функции $H(S)$. □

Заметим, что если F – оператор энергии системы, то S_θ – оператор плотности гиббсовского равновесного состояния при обратной температуре θ, а функция $-\theta^{-1}c(\theta)$ равна свободной энергии.

10.3 C-q канал с бесконечным алфавитом

Пусть $\mathcal{X}$ – некоторое (бесконечное) множество; отображение $x \to S_x$ множества $\mathcal{X}$ в пространство состояний $\mathfrak{S}(\mathcal{H})$ будем называть *классическим-квантовым (c-q) каналом* с входным алфавитом $\mathcal{X}$. Пусть $f(x)$ – функция, определенная на $\mathcal{X}$ и принимающая значения в $[0, +\infty]$. Введем класс $\mathcal{P}_E$ распределений вероятностей с конечным носителем $\pi = \{\pi_x\}$ на $\mathcal{X}$, удовлетворяющих условию

$$\sum_x f(x)\pi_x \le E, \tag{10.13}$$

где E – положительное число. Будем считать, что множество $\mathcal{P}_E$ непусто, и предполагать, что канал удовлетворяет следующему условию:

$$\sup_{\pi\in\mathcal{P}_E} H\left(\sum_x \pi_x S_x\right) < \infty. \tag{10.14}$$

Определение 10.3.1 *Определим код (W, M) размера N и длины n так же, как в определении 5.1.1, с дополнительным требованием, что все кодовые слова $w = (x_1, \ldots, x_n) \in W$ удовлетворяют аддитивному ограничению*

$$f(x_1) + \ldots + f(x_n) \le nE, \tag{10.15}$$

Средняя вероятность ошибки кода $\bar{P}_e(W, M)$ и минимальная средняя вероятность ошибки $\bar{p}_e(n, N)$ даются, соответственно, формулами (5.5) и (5.6). Классическая пропускная способность канала $x \to S_x$ определяется в соответствии с определением 5.1.2, т. е. как точная верхняя грань допустимых скоростей R, для которых $\lim_{n\to\infty} \bar{p}_e(n, 2^{nR}) = 0$.

Теорема 10.3.1 *Классическая пропускная способность C c-q канала $x \to S_x$, удовлетворяющего условию (10.14), с входным ограничением (10.15) равна величине*

$$C_\chi = \sup_{\pi\in\mathcal{P}_E} \left[H\left(\sum_x \pi_x S_x\right) - \sum_x \pi_x H\left(S_x\right)\right]. \tag{10.16}$$

Доказательство. Обозначим $\mathcal{P}_E^{(n)}$ класс распределений вероятностей на $\mathcal{X}^n$, удовлетворяющих условию

$$\sum_{x_1,\ldots,x_n} [f(x_1) + \ldots + f(x_n)]\pi_{x_1,\ldots,x_n} \le nE \tag{10.17}$$

и определим величину C_χ^n выражением (5.26), в котором супремум по π берется по множеству $\mathcal{P}_E^{(n)}$, т. е.

$$C_\chi^n = \sup_{\pi\in\mathcal{P}_E^{(n)}} \chi\left(\{\pi_w\}; \{S_w\}\right).$$

Лемма 10.3.2 *Последовательность $\left\{C_\chi^n\right\}$ аддитивна, т. е. $C_\chi^n = nC_\chi$.*

Доказательство. В силу (5.27) имеем

$$\chi_n(\pi) \le \sum_{k=1}^n \chi(\pi^{(k)}),$$

где $\pi^{(k)}$ – k-е маргинальное распределение π на $\mathcal{X}$. Далее

$$\sum_{k=1}^{n} \chi(\pi^{(k)}) \leq n\chi(\bar{\pi}), \tag{10.18}$$

где $\bar{\pi} = \frac{1}{n}\sum_{k=1}^{n} \pi^{(k)}$, поскольку $\chi(\pi)$ – вогнутая функция π (это следует из вогнутости энтропии фон Неймана). Неравенство (10.17) можно переписать следующим образом

$$\frac{1}{n}\sum_{k=1}^{n}\sum_{x} f(x)\pi^{(k)}(x) \leq E,$$

откуда следует, что $\bar{\pi} \in \mathcal{P}_E$ если $\pi \in \mathcal{P}_E^{(n)}$. Беря супремум по $\pi \in \mathcal{P}_E^{(n)}$ в соотношении (10.18), получаем $C_\chi^n \leq nC_\chi$. Обратное неравенство очевидно. □

Для доказательства обращения теоремы кодирования используем следствие неравенства Фано (см. (5.29)):

$$\overline{P}_e(W, M) \geq 1 - \frac{\sup_{\pi\in\mathcal{P}_E^{(n)}} \sup_M \mathcal{I}_n(\pi, M)}{nR} - \frac{1}{nR}, \tag{10.19}$$

которое выводится следующим образом. Пусть слова в коде (W, M) берутся с входным распределением $\pi^{(N)}$, приписывающим равную вероятность $1/N$ каждому слову. Рассмотрим неравенство (4.35). Поскольку кодовые слова удовлетворяют условию (10.15), имеем $\pi^{(N)} \in \mathcal{P}_E^{(n)}$, откуда следует (10.19).

Применяя неравенство (10.19) и теорему 5.3.2, получаем

$$\bar{p}_e(n, 2^{nR}) \geq 1 - \frac{C_\chi}{R} - \frac{1}{nR},$$

где C_χ определяется выражением (10.16), откуда следует, что $\bar{p}_e(n, 2^{nR}) \not\to 0$ для $R > C_\chi$.

В классической теории информация прямая теорема кодирования для каналов с аддитивными ограничениями может быть доказана с помощью случайного кодирования с распределением вероятностей (5.42), модифицированного множителем, сконцентрированным на словах, для которых ограничение выполняется в виде приближенного равенства. Аналогичный подход можно применить и к c-q каналу. Пусть π – распределение на $\mathcal{X}$, удовлетворяющее условию (10.13), и пусть P – распределение на множестве N слов, при котором слова независимы и имеют распределение вероятностей (5.42). Обозначим $\nu_n = \mathsf{P}(\frac{1}{n}\sum_{k=1}^{n} f(x_k) \leq E)$ и определим модифицированное распределение вероятностей $\tilde{\mathsf{P}}$, при котором слова по-прежнему независимы, но

$$\tilde{\mathsf{P}}(w=(x_1,\dots,x_n))=\begin{cases}\nu_n^{-1}\pi_{x_1}\cdot\ldots\cdot\pi_{x_n}, & \text{если } \sum_{k=1}^n f(x_k)\le nE,\\ 0, & \text{в противном случае.}\end{cases} \tag{10.20}$$

Заметим, что поскольку $\pi\in\mathcal{P}_E$, то $\mathsf{E}f\le E$ (где E – математическое ожидание относительно распределения P) и, следовательно, в силу центральной предельной теоремы

$$\lim_{n\to\infty}\nu_n\ge 1/2.$$

Поэтому $\tilde{\mathsf{E}}\xi\le 2^m\mathsf{E}\xi$ (где $\tilde{\mathsf{E}}$ – математическое ожидание относительно распределения $\tilde{\mathsf{P}}$) для любой неотрицательной случайной величины ξ, зависящей от m различных слов.

Вероятность ошибки $\bar{P}_e(W,M)$ ограничена сверху величиной (5.46). Оценим математическое ожидание этой границы относительно распределения $\tilde{\mathsf{P}}$. Поскольку каждое слагаемое в правой части (5.46) зависит не более, чем от двух различных слов, имеем

$$\tilde{\mathsf{E}}\inf_M\bar{P}_e(W,M)\le 4\mathsf{E}\inf_M\bar{P}_e(W,M),$$

и математическое ожидание относительно распределения P можно сделать сколь угодно малым при $N=2^{nR}, n\to\infty$, если $R<C_\chi-3\delta$. Поэтому $\tilde{\mathsf{E}}\bar{P}_e(W,M)$ также можно сделать сколь угодно малым при тех же условиях. Поскольку распределение $\tilde{\mathsf{P}}$ сконцентрировано на словах, удовлетворяющих условию (10.15), можно выбрать код, для которого величину $\bar{P}_e(W,M)$ можно сделать сколь угодно малой для достаточно больших n. □

10.4 C-q канал с непрерывным алфавитом

Пусть теперь входной алфавит $\mathcal{X}$ является полным сепарабельным метрическим пространством с $\sigma-$алгеброй борелевских подмножеств. Рассмотрим c-q канал, который задается *непрерывным* отображением $x\to S_x$ алфавита $\mathcal{X}$ в множество квантовых состояний $\mathfrak{S}(\mathcal{H})$ (в силу леммы 10.1.1 для этого необходимо и достаточно, чтобы все матричные элементы $\langle\psi|\,S_x\,|\phi\rangle;\psi,\phi\in\mathcal{H}$, были непрерывны на $\mathcal{X}$). Нашей целью будет получение удобного интегрального выражения для классической пропускной способности C канала $x\to S_x$ при ограничении (10.15).

Предположим, что f – неотрицательная борелевская функция на $\mathcal{X}$. Рассмотрим множество $\mathcal{P}_E^B$ борелевских вероятностных мер π на $\mathcal{X}$, удовлетворяющих условию

$$\int_{\mathcal{X}}f(x)\pi(dx)\le E. \tag{10.21}$$

Последовательность борелевских вероятностных мер $\{\pi_n\}$ *слабо сходится* к мере π, если

$$\int_{\mathcal{X}} \varphi(x)\pi_n(dx) \to \int_{\mathcal{X}} \varphi(x)\pi(dx)$$

для любой ограниченной непрерывной функции φ. Нам понадобятся следующие вспомогательные результаты (см. Комментарии):

Задача 10.4.1 *Пусть f – полунепрерывная снизу функция, тогда функционал $\pi \to \int_{\mathcal{X}} f(x)\pi(dx)$ полунепрерывен снизу по отношению к слабой сходимости вероятностных мер. Указание: используйте тот факт, что f является точной верхней гранью семейства всех ограниченных непрерывных функций $\varphi \le f$.*

Лемма 10.4.1 *Пусть f – полунепрерывная снизу функция, тогда множество $\mathcal{P}_E$ всех вероятностных мер с конечным носителем, удовлетворяющих условию (10.21), слабо плотно в $\mathcal{P}_E^B$.*

Приведем также удобное достаточное условие слабой компактности множества $\mathcal{P}_E^B$.

Лемма 10.4.2 *Если функция f полунепрерывна снизу и для любого положительного числа k множество $\{x : f(x) \le k\} \subset \mathcal{X}$ компактно, то подмножество вероятностных мер*

$$\mathcal{P}_E^B = \left\{\pi : \int_{\mathcal{X}} f(x)\pi(dx) \le E\right\}$$

слабо компактно.

Доказательство. Из полунепрерывности снизу функции f следует, что функционал $\pi \to \int_{\mathcal{X}} f(x)\pi(dx)$ полунепрерывен снизу по отношению к слабой сходимости вероятностных мер (задача 10.4.1), откуда следует слабая замкнутость множества $\mathcal{P}_E^B$. Известно, что слабо замкнутое подмножество $\mathcal{P}$ борелевских вероятностных мер на $\mathcal{X}$ слабо компактно тогда и только тогда, когда для любого $\varepsilon > 0$ существует компакт $K \subset \mathcal{X}$ такой, что $\pi(K^c) \le \varepsilon$ для всех $\pi \in \mathcal{P}$. Для данного $\varepsilon > 0$ рассмотрим компактное множество $K = \{x : f(x) \le E/\varepsilon\}$. Тогда $\pi(K^c) \le \frac{\varepsilon}{E}\int_{K^c} f(x)\pi(dx) \le \varepsilon$ для $\pi \in \mathcal{P}_E^B$, следовательно, множество $\mathcal{P}_E^B$ слабо компактно. □

Для произвольной борелевской вероятностной меры π на $\mathcal{X}$ введем среднее состояние

$$\bar{S}_\pi = \int_{\mathcal{X}} S_x\pi(dx), \tag{10.22}$$

где, по предположению, подынтегральная функция S_x непрерывна и интеграл определяет оператор плотности в $\mathcal{H}$.

Предполагая, что $H(\bar{S}_\pi) < \infty$, рассмотрим функционал

$$\chi(\pi) = H(\bar{S}_\pi) - \int_{\mathcal{X}} H(S_x)\pi(dx), \tag{10.23}$$

где функция $H(S_x)$ – неотрицательна и полунепрерывна снизу в силу леммы 10.2.1, следовательно, интеграл от нее определен.

Теорема 10.4.3 *Пусть существует оператор F типа $\mathfrak{F}$, удовлетворяющий условию (10.9), такой, что*

$$f(x) \geq \mathrm{Tr} S_x F; \qquad x \in \mathcal{X}. \tag{10.24}$$

Тогда выполнено условие (10.14), пропускная способность C канала $x \to S_x$ с ограничением (10.15) конечна и дается выражением (10.16).

Если, дополнительно, функция f удовлетворяет условиям леммы 10.4.2, то

$$C = \max_{\pi \in \mathcal{P}_E^B} \chi(\pi). \tag{10.25}$$

Доказательство. Интегрируя (10.24), имеем

$$\mathrm{Tr}\bar{S}_\pi F \leq E \tag{10.26}$$

для $\pi \in \mathcal{P}_E$, следовательно, согласно лемме 10.2.2,

$$H(\bar{S}_\pi) \leq \theta \mathrm{Tr}\bar{S}_\pi F + c(\theta) \leq \theta E + c(\theta),$$

поэтому условие (10.14) выполнено, величина C конечна и дается выражением (10.16).

Если функция f удовлетворяет условиям леммы 10.4.2, то множество $\mathcal{P}_E^B$ слабо компактно. Покажем, что функционал $\pi \to \chi(\pi)$ непрерывен и, следовательно, достигает своего максимума на $\mathcal{P}_E^B$.

Рассмотрим первое слагаемое в формуле (10.23) для $\chi(\pi)$. Поскольку матричные элементы семейства $\{S_x\}$ непрерывно зависят от x, отображение $\pi \to \bar{S}_\pi$ в множество $\mathfrak{S}_E$, снабженное слабой операторной топологией, непрерывно относительно слабой сходимости мер. В силу леммы 10.1.1, слабая операторная топология на $\mathfrak{S}(\mathcal{H})$ эквивалентна топологии ядерной нормы. Квантовая энтропия непрерывна на $\mathfrak{S}_E$ в силу леммы 10.2.2. Поэтому функционал $\pi \to H(\bar{S}_\pi)$ непрерывен на множестве $\mathcal{P}_E^B$.

Покажем, что второе слагаемое в (10.23) также непрерывно. Для любой $\pi \in \mathcal{P}_E^B$ и $k > 0$ рассмотрим разложение

$$\int_{\mathcal{X}} H(S_x)\pi(dx) = \int_{f(x)\leq k} H(S_x)\pi(dx) + \int_{f(x)>k} H(S_x)\pi(dx). \tag{10.27}$$

Непрерывность будет следовать из того, что первое слагаемое непрерывно на $\mathcal{P}_E^B$, а второе может быть сделано равномерно малым на $\mathcal{P}_E^B$ за счет выбора достаточно большого k. В силу условия (10.24), неравенство $f(x) \leq k$ влечет $S_x \in \mathfrak{S}_k$, а в силу леммы 10.2.2, энтропия ограничена

и непрерывна на $\mathfrak{S}_k$, поэтому $H(S_x)$ ограничена и непрерывна на компактном множестве $\{x : f(x) \leq k\}$ и следовательно первое слагаемое в разложении (10.27) непрерывно зависит от π. С другой стороны, подставляя $S = S_x$ в (10.10) и вновь используя (10.24), получаем

$$H(S_x) \leq \theta f(x) + c(\theta).$$

Фиксируем $\varepsilon > 0$, тогда для $\pi \in \mathcal{P}_E^B$ получаем

$$\begin{aligned}\int_{f(x)>k} H(S_x)\pi(dx) &\leq \theta \int_{f(x)>k} f(x)\pi(dx) + c(\theta) \int_{f(x)>k} \pi(dx) \\ &\leq \theta E + c(\theta) E k^{-1} \leq \varepsilon,\end{aligned}$$

выбирая сначала достаточно малое θ, а затем достаточно большое k. Итак, второе слагаемое, а значит и весь функционал $\chi(\pi)$, непрерывны на $\mathcal{P}_E^B$.

Из леммы 10.4.1 и выражения (10.16) тогда вытекает, что $C = \sup_{\pi \in \mathcal{P}_E^B} \chi(\pi)$, причем супремум достигается в силу непрерывности $\chi(\pi)$ и компактности $\mathcal{P}_E^B$. □

10.5 Квантовый канал с ограничением

Пусть $\mathcal{H}_A, \mathcal{H}_B$ сепарабельные гильбертовы пространства, которые будут называться, соответственно, входным и выходным пространством.

Определение 10.5.1 *Каналом называется линейное, ограниченное, сохраняющее след, вполне положительное отображение $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$. Как и в конечномерном случае, полная положительность означает, что для любого $n = 1, 2, \ldots$ отображение $\Phi \otimes \mathrm{Id}_n$ положительно.*

Задача 10.5.1 *Произвольное аффинное отображение $\Phi_s : \mathfrak{S}(\mathcal{H}_A) \to \mathfrak{S}(\mathcal{H}_B)$ однозначно продолжается до линейного, ограниченного, положительного, сохраняющего след отображения $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$. Указание: для построения линейного продолжения используйте конструкцию задачи 2.2.1. Для доказательства ограниченности – оценки из леммы 9.2.2.*

В этом разделе мы изучим классическую пропускную способность канала Φ при аддитивном ограничении на входе. Пусть F – оператор типа $\mathfrak{F}$ в $\mathcal{H}_A$, представляющий величину, среднее значение которой ограничено (в приложениях это обычно средняя энергия сигнала на входе канала). Предполагается, что состояния $S^{(n)}$ на входе блочного канала $\Phi^{\otimes n}$ подчинены аддитивному ограничению

$$\mathrm{Tr} S^{(n)} F^{(n)} \leq nE, \tag{10.28}$$

где

$$F^{(n)} = F \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes F,$$

а E – положительная постоянная. Модифицируя определение 8.1.1 и доказательство предложения 8.1.1 на случай кодовых состояний с ограничениями (10.28), можно доказать

Предложение 10.5.1 *Пусть канал Φ удовлетворяет условию*

$$\sup_{S:\mathrm{Tr}SF\leq E} H(\Phi[S]) < \infty. \tag{10.29}$$

Тогда классическая пропускная способность этого канала при ограничении (10.28) конечна и дается выражением

$$C(\Phi, F, E) = \lim_{n\to\infty} \frac{1}{n} C_\chi(\Phi^{\otimes n}, F^{(n)}, nE), \tag{10.30}$$

где

$$C_\chi(\Phi, F, E) = \sup_{\pi:\mathrm{Tr}\,\bar{S}_\pi F\leq E} \chi\left(\{\pi_i\}; \{\Phi[S_i]\}\right). \tag{10.31}$$

Здесь $\bar{S}_\pi = \sum_i \pi_i S_i$ – среднее состояние ансамбля $\pi = \{\pi_i, S_i\}$.

Если для канала Φ выполняется свойство аддитивности

$$C_\chi(\Phi^{\otimes n}, F^{(n)}, nE) = nC_\chi(\Phi, F, E), \tag{10.32}$$

то имеет место равенство

$$C(\Phi, F, E) = C_\chi(\Phi, F, E).$$

Это тесно связано со свйством супераддитивности выпуклого замыкания выходной энтропии (8.35), из которого следует аддитивность χ-пропускной способности при линейных ограничениях (8.39) (см. раздел 8.3.2).

В любом случае величина (10.31) дает нижнюю границу для классической пропускной способности $C(\Phi, F, E)$. Получим более удобное выражение для $C_\chi(\Phi, F, E)$, опираясь на результаты предыдущего раздела. Рассмотрим c-q канал с алфавитом $\mathcal{X} = \mathfrak{S}(\mathcal{H}_A)$, определяемый отображением $S \to \Phi[S]$. Ограничение будет задаваться функцией $f(S) = \mathrm{Tr}SF$, которая является аффинной и полунепрерывной снизу в силу леммы 10.1.3, а условие (10.14) переходит в (10.29).

Определение 10.5.2 *Обобщенным ансамблем называется произвольная борелевская вероятностная мера π на $\mathfrak{S}(\mathcal{H}_A)$. Среднее состояние обобщенного ансамбля π задается барицентром*

$$\bar{S}_\pi = \int\limits_{\mathfrak{S}(\mathcal{H}_A)} S\pi(dS).$$

Обычные ансамбли соответствуют мерам с конечным носителем.

Из леммы 10.2.1 следует, в частности, что неотрицательная функция $S \mapsto H(\Phi[S])$ измерима. Следовательно, при условии $H(\Phi(\bar{S}_\pi)) < \infty$ определен функционал

$$\chi_\Phi(\pi) = H(\Phi(\bar{S}_\pi)) - \int\limits_{\mathfrak{S}(\mathcal{H}_A)} H(\Phi(S))\pi(dS). \tag{10.33}$$

Задача 10.5.2 *Величина $\chi_\Phi(\pi)$ является вогнутым функционалом от π.*

Следствие 10.5.1 *Пусть существует положительный оператор F' в $\mathcal{H}_B$, такой что*

$$\operatorname{Tr}\exp(-\theta F') < +\infty, \qquad \theta > 0 \tag{10.34}$$

и

$$\operatorname{Tr}\Phi[S]F' \le \operatorname{Tr} SF; \qquad S \in \mathfrak{S}(\mathcal{H}). \tag{10.35}$$

Тогда выполнено условие (10.29) и

$$C_\chi(\Phi, F, E) = \sup_{\pi:\operatorname{Tr}\bar{S}_\pi F \le E} \chi_\Phi(\pi).$$

Более того, энтропия $H(\Phi[S])$ непрерывна на компактном множестве $\mathfrak{S}_E = \{S : \operatorname{Tr} SF \le E\}$.

Если, дополнительно, оператор F удовлетворяет условиям леммы 10.1.4, то существует оптимальный обобщенный ансамбль, т. е.

$$C_\chi(\Phi, F, E) = \max_{\pi:\operatorname{Tr}\bar{S}_\pi F \le E} \chi_\Phi(\pi). \tag{10.36}$$

Доказательство. Утверждение следует из теоремы 10.4.3, применительно к c-q каналу $S \to \Phi[S]$ с функцией-ограничением $f(S) = \operatorname{Tr} SF$, при этом роль условия (10.24) играет неравенство (10.35), а оператор F' играет роль F. Функция $f(S) = \operatorname{Tr} SF$ удовлетворяет условиям леммы 10.4.2 в силу лемм 10.1.3, 10.1.4.

Непрерывность энтропии $H(\Phi[S])$ следует из непрерывности отображения $S \to S' = \Phi[S]$ и леммы 10.2.2, которая гарантирует непрерывность $H(S')$ на компактном множестве $\{S' : \operatorname{Tr} S'F' \le E\}$. □

10.6 Передача классической информации с помощью сцепленного состояния через канал с ограничениями

Пусть системы A и B находятся в сцепленном чистом состоянии S_{AB}. Предположим, что имеющаяся в их распоряжении сцепленность неограничена, но конечна, т. е. состояние S_{AB} должно удовлетворять условию $H(S_A) = H(S_B) < \infty$. Обобщая результат раздела 8.4, можно доказать

Предложение 10.6.1 *Пусть Φ – канал, удовлетворяющий условию (10.29) с оператором F, для которого выполнено условие (10.9), тогда пропускная способность протокола передачи классической информации с помощью сцепленного состояния по каналу Φ с ограничением (10.28) конечна и дается выражением*

$$C_{ea}(\Phi) = \sup_{S:\mathrm{Tr}SF\leq E} I\left(S,\Phi\right), \tag{10.37}$$

в котором

$$I(S,\Phi) = H(S) + H(\Phi[S]) - H(S;\Phi), \tag{10.38}$$

где $H(S;\Phi)$ – обменная энтропия.

Исследуем вопрос о достижимости точной верхней грани в правой части (10.37). Заметим, что из условия (10.9) следует, что оператор F удовлетворяет условию леммы 10.1.4 и, следовательно, множество $\mathfrak{S}_E = \{S : \mathrm{Tr}SF \leq E\}$ компактно.

Предложение 10.6.2 *Пусть ограничивающий оператор F удовлетворяет условию (10.9), и пусть существует оператор F' типа $\mathfrak{F}$, для которого выполнены условия (10.9) и (10.35). Тогда*

$$C_{ea}(\Phi) = \max_{S:\mathrm{Tr}SF\leq E} I\left(S,\Phi\right). \tag{10.39}$$

Более того, если канал Φ таков, что

$$\sup_S I\left(S,\Phi\right) = \infty, \tag{10.40}$$

то максимум в (10.39) достигается на операторе плотности S, для которого ограничение выполнено в форме равенства $\mathrm{Tr}SF = E$.

Доказательство. Рассмотрим отдельно каждое слагаемое в формуле (10.38). Заметим, что в силу леммы 10.2.1, квантовая энтропия полунепрерывна снизу. Поскольку обменную энтропию можно представить в виде $H(S;\Phi) = H(\Phi_E\left[S\right])$, где Φ_E – комплементарный канал из пространства системы $\mathcal{H}_A$ в пространство окружения $\mathcal{H}_E$, она также полунепрерывна снизу и поэтому последнее слагаемое в (10.38) полунепрерывно сверху. Что касается первого слагаемого $H(S)$, то оно непрерывно на множестве $\mathfrak{S}_E = \{S : \mathrm{Tr}SF \leq E\}$ в силу леммы 10.2.2, поскольку ограничивающий оператор F удовлетворяет условию (10.9). Из следствия 10.5.1 вытекает непрерывность второго слагаемого в (10.38), т.е. $H(\Phi[S])$. Более того, это следствие также гарантирует выполнение условия (10.29), а значит имеет место формула (10.37). Как следует из сказанного выше, взаимная информация (10.38) полунепрерывна сверху на компактном множестве $\mathfrak{S}_E$, а значит, достигает своего максимума.

Для доказательства второго утверждения обозначим

$$f_0(E) = \max_{\mathrm{Tr}SF=E} I\left(S, \Phi\right),$$

и предположим, что существует S_1, такое что

$$\mathrm{Tr}\, S_1 F < E, \quad I(S_1, \Phi) > f_0(E).$$

Из условия (10.40) следует, что найдется S_2, такое что

$$\mathrm{Tr}\, S_2 F > E, \quad I(S_2, \Phi) > f_0(E).$$

Тогда, полагая $\lambda = \frac{\mathrm{Tr}S_2F - E}{\mathrm{Tr}S_2F - \mathrm{Tr}S_1F}$, имеем $0 < \lambda < 1$, $\mathrm{Tr}(\lambda\ S_1 + (1-\lambda)\ S_2)F = E$ и

$$I(\lambda S_1 + (1-\lambda)S_2, \Phi) \leq f_0(E) < \lambda I(S_1, \Phi) + (1-\lambda) I(S_2, \Phi),$$

что противоречит вогнутости $I\left(S, \Phi\right)$. □

10.7 Каналы, разрушающие сцепленность

Состояние $S \in \mathfrak{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ называется *разделимым* (несцепленным), если оно принадлежит выпуклому замыканию (т. е. замыканию выпуклой оболочки) множества всех состояний-произведений $S_1 \otimes S_2 \in \mathfrak{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$.

Предложение 10.7.1 *Состояние S является разделимым тогда и только тогда, когда оно допускает представление*

$$S = \int_{\mathcal{X}} S_1(x) \otimes S_2(x)\, \pi(dx), \tag{10.41}$$

где $\pi(dx)$ – борелевская вероятностная мера, а $S_j(x), j = 1, 2,$ – борелевские $\mathfrak{S}(\mathcal{H}_j)$-значные функции на некотором полном сепарабельном метрическом пространстве $\mathcal{X}$.

Доказательство. Состояние S разделимо тогда и только тогда, когда $S = \lim_{n\to\infty} S_n$, где

$$S_n = \int_{\mathfrak{S}(\mathcal{H}_1)} \int_{\mathfrak{S}(\mathcal{H}_2)} S_1 \otimes S_2\, \pi_n(dS_1 dS_2), \tag{10.42}$$

и $\{\pi_n\}$ – последовательность борелевских вероятностных мер с конечными носителями.

Пусть S представляется в виде (10.41), тогда, производя замену переменной интегрирования $x \to S_1(x) \otimes S_2(x)$, можно свести представление (10.41) к

$$S = \int_{\mathfrak{S}(\mathcal{H}_1)} \int_{\mathfrak{S}(\mathcal{H}_2)} S_1 \otimes S_2\, \pi(dS_1 dS_2), \tag{10.43}$$

где мы используем то же обозначение π для образа исходной меры при данной замене переменной. Отображение $\pi \to S$ непрерывно относительно слабой сходимости вероятностных мер; это очевидно, если множество операторов плотности снабжено слабой операторной топологией, а в силу леммы 10.1.1 последняя совпадает на этом множестве с топологией ядерной нормы. В силу леммы 10.4.1, найдется последовательность $\{\pi_n\}$ борелевских вероятностных мер с конечными носителями, которая слабо сходится к π. Используя непрерывность отображения $\pi \to S$, получаем, что $S = \lim_{n\to\infty} S_n$, где S_n дается соотношением (10.42), следовательно, S разделимо.

Обратно, пусть S разделимо, тогда $S = \lim_{n\to\infty} S_n$, где S_n дается соотношением (10.42). Если мы докажем, что последовательность мер $\{\pi_n\}$ слабо относительно компактна, то отсюда будет следовать представление (10.43), где π – частичный предел этой последовательности. Сходящаяся последовательность $\{S_n\}$ относительно компактна, поэтому в силу теоремы 10.1.2 для любого $\varepsilon > 0$ найдется конечномерный проектор P, такой что $\operatorname{Tr} S_n(I-P) \le \varepsilon$ для всех n. Для $m = 1, 2, \ldots$ обозначим P_m проектор, такой что $\operatorname{Tr} S_n(I - P_m) \le 4^{-m}$, и следовательно

$$\int \operatorname{Tr}(S_1 \otimes S_2)(I - P_m)\pi_n\,(dS_1 dS_2) \le 4^{-m}. \tag{10.44}$$

Введем следующие подмножества в прямом произведении $\mathfrak{S}(\mathcal{H}_1)\times\mathfrak{S}(\mathcal{H}_2)$:

$$K_m = \left\{(S_1, S_2) : \operatorname{Tr}(S_1 \otimes S_2)(I - P_m) \le \delta^{-1}2^{-m}\right\}; \quad \mathcal{K}_\delta = \cap_{m\ge 1} K_m,$$

где $\delta > 0$. В силу той же теоремы 10.1.2, множество $\mathcal{K}_\delta$ компактно по построению. Для его дополнения имеем

$$\pi_n\,(\mathcal{K}_\delta^c) \le \sum_m \pi_n\,(K_m^c) \le \delta \sum_m 2^m \int \operatorname{Tr}(S_1 \otimes S_2)(I - P_m)\pi_n\,(dS_1 dS_2) \le \delta$$

согласно (10.44). Таким образом, последовательность мер $\{\pi_n\}$ удовлетворяет известному критерию слабой относительной компактности (см. Комментарии), что и завершает доказательство теоремы. □

Определение 10.7.1 *Канал $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$ называется каналом, разрушающим сцепленность, если для любого гильбертова пространства $\mathcal{H}_R$ и состояния $S \in \mathfrak{S}(\mathcal{H}_A \otimes \mathcal{H}_R)$, состояние $(\Phi \otimes \mathrm{Id}_R)[S] \in \mathfrak{S}(\mathcal{H}_B \otimes \mathcal{H}_R)$, где Id_R тождественный канал в $\mathfrak{T}(\mathcal{H}_R)$, является разделимым.*

Для описания структуры таких каналов нам потребуется обобщение понятия наблюдаемой (определение 2.2.1) на случай произвольного множества исходов.

Определение 10.7.2 *Пусть $\mathcal{X}$ - измеримое пространство с σ-алгеброй измеримых подмножеств $\mathfrak{B}$. Наблюдаемой со значениями в $\mathcal{X}$ называется вероятностная операторно-значная мера на $\mathcal{X}$, то есть семейство $M = \{M(B), B \in \mathfrak{B}\}$ эрмитовых операторов в $\mathcal{H}$, удовлетворяющее условиям*

- $M(B) \geq 0; \quad B \in \mathfrak{B}$;
- $M(\mathcal{X}) = I$;
- *для любого счетного разбиения $B = \cup B_j, \quad (B_i \cap B_j = \emptyset, i \neq j)$ выполняется $M(B) = \sum_j M(B_j)$ в смысле слабой сходимости операторов.*

Вероятностная операторно-значная мера E называется спектральной мерой, если

$$E(B_1 \cap B_2) = E(B_1)E(B_2) \quad B_1, B_2 \in \mathfrak{B}.$$

В этом случае все операторы $E(B)$ являются проекторами, причем проекторы, отвечающие непересекающимся подмножествам, взаимно ортогональны. Соответствующая наблюдаемая называется четкой.

Распределение вероятностей наблюдаемой M в состоянии S задается вероятностной мерой

$$\mu_S^M(B) = \operatorname{Tr} SM(B), \quad B \in \mathfrak{B}. \tag{10.45}$$

Отметим, что линейное продолжение аффинного отображения $S \to \mu_S^M$ можно рассматривать как обобщение понятия квантово-классического (c-q) канала (см. раздел 6.4 в гл. 6).

Задача 10.7.1 *Докажите, что формула (10.45) задает вероятностную меру на $\mathfrak{B}$.*

Теорема 10.7.2 *Канал Φ разрушает сцепленность тогда и только тогда, когда найдутся полное сепарабельное метрическое пространство $\mathcal{X}$, борелевская $\mathfrak{S}(\mathcal{H}_B)$-значная функция $S_B(x)$ и наблюдаемая M в $\mathcal{H}_A$ с множеством исходов $\mathcal{X}$, задаваемая вероятностной операторно-значной мерой $M(dx)$, такие что*

$$\Phi[S] = \int\limits_{\mathcal{X}} S_B(x)\mu_S^M(dx). \tag{10.46}$$

Соотношение (10.46) является континуальной версией представления (6.27), а утверждение можно рассматривать как бесконечномерное обобщение предложения 6.4.1. Таким образом, и в бесконечномерном случае квантовый канал, разрушающий сцепленность, является композицией q-c канала (измерение наблюдаемой M) и c-q канала (приготовляющего состояние $S_B(x)$ в зависимости от исхода измерения).

Набросок доказательства. Предположим, что канал имеет вид (10.46). Рассмотрим состояние $\omega \in \mathfrak{S}(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{R}})$. Тогда

$$(\Phi \otimes \mathrm{Id}_R(\omega) = \int\limits_{\mathcal{X}} S_B(x) \otimes m_\omega(dx), \tag{10.47}$$

где

$$m_\omega(B) = \mathrm{Tr}_A\, \omega(M(B) \otimes I_R), \quad B \subseteq \mathcal{X}.$$

Любой матричный элемент операторно-значной меры m_ω (в фиксированном базисе) является комплексной мерой на $\mathcal{X}$, абсолютно непрерывной относительно вероятностной меры $\mu_\omega(B) = \mathrm{Tr}\, m_\omega(B),\ B \subseteq \mathcal{X}$. Из теоремы Радона-Никодима вытекает представление

$$m_\omega(B) = \int\limits_B \sigma_\omega(x)\mu_\omega(dx),$$

где $\sigma_\omega(x)$ – функция на $\mathcal{X}$ со значениями в $\mathfrak{S}(\mathcal{K})$. Используя это представление, мы можем переписать (10.47) в виде

$$(\Phi \otimes \mathrm{Id}_{\mathcal{K}})(\omega) = \int\limits_{\mathcal{X}} S'(x) \otimes \sigma_\omega(x)\mu_\omega(dx), \tag{10.48}$$

что является разделимым состоянием согласно предложению 10.7.1.

Обратно, пусть Φ – канал, разрушающий сцепленность. Фиксируем невырожденное состояние S_A в $\mathfrak{S}(\mathcal{H}_A)$ и пусть $\{|e_j\rangle; j = 1, \dots\}$ – базис собственных векторов оператора S_A с соответствующими (положительными) собственными значениями $\{\lambda_j\}$. Рассмотрим единичный вектор

$$|\Omega\rangle = \sum_{j=1}^{+\infty} \sqrt{\lambda_j}|e_j\rangle \otimes |e_j\rangle$$

в пространстве $\mathcal{H}_A \otimes \mathcal{H}_A$, тогда $|\Omega\rangle\langle\Omega|$ является очищением состояния S_A. Поскольку Φ разрушает сцепленность, состояние

$$S_{AB} = (\mathrm{Id}_A \otimes \Phi)\,[|\Omega\rangle\langle\Omega|] \tag{10.49}$$

в $\mathfrak{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ является разделимым. В силу представления (10.41) найдется вероятностная мера π на $\mathfrak{S}(\mathcal{H}_A) \times \mathfrak{S}(\mathcal{H}_B)$, такая что

$$(\mathrm{Id}_A \otimes \Phi)(|\Omega\rangle\langle\Omega|) = \int\limits_{\mathfrak{S}(\mathcal{H}_A)} \int\limits_{\mathfrak{S}(\mathcal{H}_B)} S_A \otimes S_B\, \pi(dS_A dS_B). \tag{10.50}$$

Отсюда вытекает

$$\begin{aligned}\sigma &= \mathrm{Tr}_B(\mathrm{Id}_{\mathcal{H}_A}\otimes\Phi)(|\Omega\rangle\langle\Omega|)\\ &= \int\limits_{\mathfrak{S}(\mathcal{H}_A)}\int\limits_{\mathfrak{S}(\mathcal{H}_B)} S_A\,\pi(dS_A dS_B)\\ &= \int\limits_{\mathfrak{S}(\mathcal{H}_A)}\int\limits_{\mathfrak{S}(\mathcal{H}_B)} \bar{S}_A\,\pi(dS_A dS_B),\end{aligned} \tag{10.51}$$

где черта обозначает комплексное сопряжение в базисе $\{|e_i\rangle\}$. В силу этого равенства, для любого борелевского $B\subseteq\mathfrak{S}(\mathcal{H}_B)$, оператор

$$M(B)=\sigma^{-1/2}\left[\int\limits_{\mathfrak{S}(\mathcal{H}_A)}\int\limits_{B}\bar{S}_A\,\pi(dS_A dS_B)\right]\sigma^{-1/2} \tag{10.52}$$

может быть корректно определен как ограниченный положительный оператор в $\mathcal{H}_A$, такой что $M(B)\le M(\mathcal{X})=I_A$. Нетрудно убедиться, что M является наблюдаемой с исходами в $\mathcal{X}=\mathfrak{S}(\mathcal{H}_B)$.

Теперь рассмотрим канал, разрушающий сцепленность

$$\hat{\Phi}(S)=\int\limits_{\mathfrak{S}(\mathcal{H}_B)} S_B\mu_S^M(dS_B),$$

и покажем, что $\Phi(S)=\hat{\Phi}(S)$. Для этого достаточно проверить, что

$$\hat{\Phi}(|e_i\rangle\langle e_j|)=\Phi(|e_i\rangle\langle e_j|)$$

для всех i,j. Однако

$$\begin{aligned}\hat{\Phi}(|e_i\rangle\langle e_j|) &= \int\limits_{\mathfrak{S}(\mathcal{H}_B)} S_B\langle e_j|M(dS_B)|e_i\rangle\\ = \lambda_i^{-1/2}\lambda_j^{-1/2}\int\limits_{\mathfrak{S}(\mathcal{H}_A)}\int\limits_{\mathfrak{S}(\mathcal{H}_B)}&\langle e_i|S_A|e_j\rangle S_B\pi(dS_A dS_B)=\Phi(e_{ij}),\end{aligned}$$

где

$$e_{ij}=\lambda_i^{-1/2}\lambda_j^{-1/2}\,\mathrm{Tr}_A(|e_j\rangle\langle e_i|\otimes I_A)|\Omega\rangle\langle\Omega|=|e_i\rangle\langle e_j|.$$

□

Как было показано в разделе 8.3.3, в конечномерном случае каналы, разрушающие сцепленность, образуют обширный класс, для которого выполняется гипотеза аддитивности. Этот факт обобщается и на бесконечномерный случай.

Предложение 10.7.3 *Пусть имеется два канала Φ_1, Φ_2 с соответствующими ограничениями F_1, F_2, удовлетворяющие условиям следствия 10.5.1, причем Φ_1 – канал, разрушающий сцепленность. Тогда выполняются все свойства аддитивности (8.35), (8.33) и (8.39). В частности, классическая пропускная способность канала Φ, разрушающего сцепленность, равна*

$$C(\Phi, F, E) = C_\chi(\Phi, F, E).$$

Доказательство этого предложения получается обобщением рассуждений в конечномерном случае (предложение 8.3.5), с заменой сумм на интегралы и ансамблей на обобщенные ансамбли, и с учетом того обстоятельства, что условия следствия 10.5.1 гарантируют конечность всех входящих в рассмотрение энтропий. Приведем также без доказательства обобщение следствия 9.3.1, касающегося квантовой пропускной способности:

Предложение 10.7.4 *Всякий канал в сепарабельном гильбертовом пространстве, разрушающий сцепленность, является анти-деградируемым и следовательно, имеет нулевую квантовую пропускную способность.*

10.8 Приложение: спектральное разложение

Пусть $\mathcal{H}$ – сепарабельное гильбертово пространство, и пусть $\{E(B); B \in \mathcal{B}(\mathbb{R})\}$ спектральная мера на $\mathbb{R}$, где $\mathcal{B}(\mathbb{R})$ – σ-алгебра борелевских подмножеств. Согласно определению 10.7.2, спектральная мера E задает четкую вещественную наблюдаемую. Как объясняется ниже, такие наблюдаемые находятся во взаимно однозначном соответствии с самосопряженными операторами в $\mathcal{H}$.

Для любого единичного вектора $\psi \in \mathcal{H}$ соотношение

$$\mu_\psi(B) = \langle\psi|E(B)|\psi\rangle$$

определяет борелевскую вероятностную меру на $\mathbb{R}$. Если $f(x)$ – борелевская функция на $\mathbb{R}$, то интеграл $X_f = \int_{-\infty}^{\infty} f(x)E(dx)$ сходится сильно на плотной области

$$\mathcal{D}(X_f) = \left\{\psi : \int_{-\infty}^{\infty} |f(x)|^2 \mu_\psi(dx) < \infty\right\}$$

и таким образом, однозначно определяет линейный оператор с областью определения $\mathcal{D}(X_f)$. (Напомним, что сильная сходимость последовательности операторов $\{X_n\}$ к оператору X на области $\mathcal{D}$ означает, что $\lim_{n\to\infty} \|X_n\psi - X\psi\| = 0$ для $\psi \in \mathcal{D}$, а если $\mathcal{D} = \mathcal{H}$, то это называется сильной сходимостью операторов).

В самом деле, если $f(x) = \sum_{i=1}^{\infty} f_i 1_{B_i}(x)$ – функция с конечным или счетным множеством значений (здесь $\{B_i\}$ – измеримое разбиение $\mathbb{R}$), то полагаем

$$X_f = \sum_{i=1}^{\infty} f_i E(B_i),$$

так что

$$\|X_f\psi\|^2 = \sum_{i=1}^{\infty} |f_i|^2 \mu_\psi(B_i).$$

Это определение можно расширить на произвольную борелевскую функцию $f(x)$, равномерно приближая ее функциями со счетным множеством значений. В частности, для функции $f(x) = x$ получаем оператор

$$X = \int_{-\infty}^{\infty} xE(dx) \tag{10.53}$$

с областью определения

$$\mathcal{D}(X) = \left\{\psi : \int_{-\infty}^{\infty} |x|^2 \mu_\psi(dx) < \infty\right\},$$

для которого

$$\langle\psi|X|\psi\rangle = \int_{-\infty}^{\infty} x\mu_\psi(dx); \quad \|X\psi\|^2 = \int_{-\infty}^{\infty} |x|^2 \mu_\psi(dx), \quad \psi \in \mathcal{D}(X). \tag{10.54}$$

Для любого плотно определенного оператора X существует и единствен сопряженный оператор X^*, удовлетворяющий соотношению

$$\langle X^*\varphi|\psi\rangle = \langle\varphi|X\psi\rangle, \quad \psi \in \mathcal{D}(X),\ \varphi \in \mathcal{D}(X^*), \tag{10.55}$$

где $\mathcal{D}(X^*)$ – (плотное) подпространство всех векторов φ, таких что правая часть является ограниченным линейным функционалом от ψ. Оператор X называется *эрмитовым* (симметричным), если

$$\langle X\varphi|\psi\rangle = \langle\varphi|X\psi\rangle, \tag{10.56}$$

для всех $\varphi, \psi \in \mathcal{D}(X)$ и *самосопряженным*, если $X = X^*$ в том смысле, что $\mathcal{D}(X^*) = \mathcal{D}(X)$ и выполняется соотношение (10.56). Часто приходится иметь дело с *существенно самосопряженными* операторами, которые заданы не на максимальной области определения, но однозначно продолжаются до самосопряженных. **Спектральная теорема** утверждает, что для любого самосопряженного оператора X существует единственная спектральная мера E на $\mathbb{R}$, для которой имеет место (10.53), (10.54). Таким образом, самосопряженные операторы – это в точности те, которые имеют спектральное разложение со спектром в $\mathbb{R}$ и поэтому естественно

ассоциируются с (четкими) вещественными наблюдаемыми в бесконечномерном случае. Мера μ_ψ дает распределение вероятностей наблюдаемой X в чистом состоянии $|\psi\rangle\langle\psi|$, а формулы (10.54) – первый и второй моменты наблюдаемой X.

В определенном смысле, который мы не будем здесь подробно раскрывать, имеют место соотношения $X_f = f(X)$ и $X_{fg} = X_f X_g$. Рассмотрим случай экспоненты:

$$V_t = \int_{-\infty}^{\infty} \exp(itx) E(dx) = \exp itX.$$

Тогда семейство $\{V_t; t \in \mathbb{R}\}$ является (сильно непрерывной) группой унитарных операторов с инфинитезимальным генератором X:

$$\lim_{t\to 0} t^{-1}(V_t - I)|\psi\rangle = iX|\psi\rangle; \quad \psi \in \mathcal{D}(X).$$

Если оператор X ограничен, то $\exp itX$ определяется соответствующим степенным рядом, сходящимся по норме к V_t.

Обратно, всякая сильно непрерывная группа унитарных операторов $\{V_t; t \in \mathbb{R}\}$ имеет вид $V_t = \exp itX$, где X – однозначно определенный самосопряженный оператор. Это утверждение составляет содержание **теоремы Стоуна**.

Задача 10.8.1 *1. Пусть $\mathcal{H} = L^2(\mathbb{R})$ – гильбертово пространство квадратично интегрируемых функций $\psi(\xi)$, $\xi \in \mathbb{R}$, и $E = \{E(B); B \in \mathcal{B}(\mathbb{R})\}$ определено соотношением*

$$(E(B)\psi)(\xi) = 1_B(\xi)\psi(\xi); \quad \psi \in \mathcal{H}.$$

Докажите, что E – спектральная мера оператора X умножения на ξ, V_t – оператор умножения на $\exp it\xi$, и вообще, X_f – оператор умножения на $f(\xi)$.

2. Рассмотрим унитарный оператор в $\mathcal{H}$, определяемый преобразованием Фурье

$$\left(\tilde{F}\psi\right)(\lambda) = \frac{1}{\sqrt{2\pi}} \int \exp(-i\lambda\xi)\psi(\xi)\, d\xi.$$

Тогда $\tilde{E}(B) = \tilde{F}^* E(B)\tilde{F}$ – спектральная мера на $\mathbb{R}$, которой отвечает самосопряженный оператор $\tilde{X} = \frac{1}{i}\frac{d}{d\xi}$, порождающий группу унитарных операторов

$$\left(\tilde{V}_t\psi\right)(\xi) = \psi(\xi + t).$$

Самосопряженные операторы A_j; $j = 1, \ldots, s$ называются *коммутирующими*, если коммутируют их спектральные меры. Имеет место следующее многомерное обобщение теоремы Стоуна

Теорема 10.8.1 *Пусть V_x; $x = [x_1, \ldots, x_s] \in \mathbf{R}^s$ – сильно непрерывная группа унитарных операторов в сепарабельном гильбертовом пространстве $\mathcal{H}$, тогда существует семейство A_j; $j = 1, \ldots, s$ коммутирующих самосопряженных операторов в $\mathcal{H}$, такое что*

$$V_x = \exp\left(i \sum_{j=1}^{s} x_j A_j\right). \tag{10.57}$$

Обратно, для всякого семейства A_j; $j = 1, \ldots, s$ коммутирующих самосопряженных операторов в $\mathcal{H}$, формула (10.57) определяет сильно непрерывную унитарную группу в $\mathcal{H}$.

10.9 Комментарии

1. Первая попытка математически строгого рассмотрения основных понятий квантовой механики в сепарабельном гильбертовом пространстве была сделана в классической монографии фон Неймана [21]. В частности, было подчеркнуто различие между эрмитовыми (симметрическими) и самосопряженными операторами, на которое не обращалось внимания в предшествующих работах физиков, и отмечена ключевая роль самосопряженности для спектрального разложения операторов. Другой важный круг вопросов связан с понятием следа и ядерных операторов. Современное рассмотрение этих вопросов в связи с приложениями к квантовой теории имеется в книгах [24], [75], [37].

Лемма 10.1.1 получена Дель'Антонио [76], см. также приложение в книге Дэвиса [75]. Критерий компактности для подмножеств квантовых состояний представляет собой модификацию результата, полученного Сарымсаковым [26] и названного им "некоммутативной теоремой Прохорова". Последняя дает критерий слабой компактности семейств вероятностных мер на метрических пространствах, см. [132], и используется в разделе 10.4. По поводу компактности единичного шар в $\mathfrak{B}(\mathcal{H})$ в слабой операторной топологии см., например, [24].

2. Хорошо известно, что свойства энтропии в бесконечномерном случае существенно отличаются от конечномерного: в последнем случае энтропия – ограниченная непрерывная функция на $\mathfrak{S}(\mathcal{H})$, тогда как в бесконечномерном она разрывна всюду и бесконечна "почти всюду" в том смысле, что множество состояний с конечной энтропией является подмножеством первой категории в $\mathfrak{S}(\mathcal{H})$, см. обзор Верля [152], где получен ряд полезных свойств энтропии и относительной энтропии, включая теорему 10.2.1 и лемму 10.2.2. В этом обзоре можно также найти ссылки на оригинальные работы Либа, Рускаи, Саймона и других авторов, которые внесли значительный вклад в изучение свойств квантовой энтропии. Новые результаты о свойствах квантовой энтропии, в частности, удобные условия непрерывности получены в работе Широкова [47].

3. Важность рассмотрения входных ограничений для квантовых каналов была понятна с момента появления идеи квантовых коммуникаций; подробное физическое обсуждение вопроса об определении пропускной способности квантовых оптических каналов с ограничением на среднюю энергию входного сигнала см. в обзоре [70]. Исследование классической пропускной способности c-q канала с ограничением основано на работах Холево [35], [42]. В них используется модификация случайного кодирования (10.20), которая была предложена для соответствующей классической задачи. Другие подходы рассматривались Винтером [156] и Хайаши [92].

4. Каналы с непрерывным алфавитом рассматривались в работах Холево [42], Холево и Широкова [44], где систематически использовалось понятие обобщенного ансамбля. Доказательства необходимых фактов из теории вероятностных мер на метрических пространствах, таких как задача 10.4.1 и лемма 10.4.1, можно найти в монографии Партасарати [132].

5. Доказательство ограниченности отображения Φ в задаче 10.5.1 см. в книге [75], лемма 2.2.1. Рассмотрение χ- пропускной способности квантовых каналов с ограничениями основано на работах [42], [44]. Обстоятельное исследование свойств $\chi-$пропускной способности и других энтропийных характеристик бесконечномерных каналов без предположений типа конечности выходной энтропии было проведено Широковым [140], [48].

6. Этот раздел основан на результатах работы [42].

7. Концепция квантовой наблюдаемой как вероятностной операторнозначной меры подробно излагается в книгах [37], [36]. Теорема 10.7.2 доказана в работе Холево, Широкова и Вернера [39]. По поводу доказательства предложений 10.7.3, 10.7.4, см. статью [43].

8. Спектральная теорема для самосопряженных операторов в сепарабельном гильбертовом пространстве, теорема Стоуна и связанные с этим вопросы подробно рассматриваются в книге Рида и Саймона [24].

11. Гауссовские системы

Фундаментальным физическим носителем информации является электромагнитное поле, в частности, свет и радиоволны. С математической точки зрения излучение эквивалентно ансамблю осцилляторов. В квантовой оптике рассматриваются квантованные поля и, следовательно, квантовые осцилляторы. Это – типичная бозонная квантовая система с "непрерывными переменными", в которой основные наблюдаемые (амплитуды осцилляторов) удовлетворяют каноническим коммутационным соотношениям (ККС). Многие из современных экспериментов по обработке квантовой информации реализованы именно в таких системах.

Имеется особенно важный класс состояний бозонной системы, которые естественно соответствуют классическим гауссовским распределениям. С физической точки зрения этот класс содержит состояния теплового равновесия, а также когерентные и сжатые состояния, реализуемые в лазерах и "нелинейных" квантовых оптических приборах. С математической точки зрения эти состояния полностью характеризуются, аналогично классическим гауссовским распределениям, средним значением и матрицей ковариаций. При рассмотрении конечного числа осцилляторных мод, что является обычным приближением в квантовой оптике, возможно описание таких состояний с помощью методов конечномерной линейной алгебры.

В связи с бозонными системами и ККС неизбежно возникают аналитические сложности, связанные с бесконечной размерностью и неограниченностью некоторых операторов. Мы будем уделять основное внимание тем вопросам, которые имеют первостепенное значение с точки зрения квантовой теории информации, иногда обращаясь с неограниченными операторами достаточно бегло. Более детальное математическое рассмотрение таких вопросов можно при желании найти в соответствующей литературе (см. Комментарии).

11.1 Операторы, ассоциированные с коммутационным соотношением Гейзенберга

В гильбертовом пространстве $\mathcal{H} = L^2(\mathbb{R})$ рассмотрим операторы

$$(q\psi)(\xi) = \xi\psi(\xi); \quad (p\psi)(\xi) = \frac{\hbar}{i}\frac{d}{d\xi}\psi(\xi),$$

заданные на общей области определения $\mathcal{D}$, в качестве которой можно взять пространство $\mathcal{S}(\mathbb{R})$ бесконечно дифференцируемых функций, все производные которых убывают на бесконечности быстрее, чем любая степень $|\xi|$. Эти операторы являются существенно самосопряженными, следовательно, представляют (четкие) вещественные наблюдаемые (см. раздел 10.8).

На указанной области определения операторы q и p удовлетворяют коммутационному соотношению Гейзенберга

$$[q, p] = i\hbar I. \tag{11.1}$$

Рассмотрим чистое состояние $|\psi\rangle\langle\psi|$, $\psi \in \mathcal{D}$, и введем обозначения

$$x = \langle\psi|q|\psi\rangle, \quad y = \langle\psi|p|\psi\rangle,$$

$$\mathsf{D}_\psi(q) = \|(q-x)\,\psi\|^2, \mathsf{D}_\psi(p) = \|(p-y)\,\psi\|^2$$

для средних значений и дисперсий наблюдаемых q, p. Для любого вещественного числа ω

$$\omega^2\mathsf{D}_\psi(q) - \omega\hbar + \mathsf{D}_\psi(p) = \omega^2\mathsf{D}_\psi(q) + i\omega\langle\psi|[q,p]|\psi\rangle + \mathsf{D}_\psi(p)$$

$$= \|[\omega\,(q-x) + i\,(p-y)]\,\psi\|^2 \geq 0,$$

следовательно,

$$\mathsf{D}_\psi(q)\mathsf{D}_\psi(p) \geq \frac{\hbar^2}{4}, \tag{11.2}$$

причем равенство достигается тогда и только тогда, когда найдется ω такое, что ψ является решением дифференциального уравнения

$$\left[\omega\,(\xi - x) + \left(\hbar\frac{d}{d\xi} - iy\right)\right]\psi\,(\xi) = 0, \tag{11.3}$$

нормированное решение которого существует при $\omega > 0$ и имеет вид

$$\psi\,(\xi) = \frac{k}{\sqrt[4]{\pi\hbar/\omega}}\exp\left[\frac{iy\xi}{\hbar} - \frac{\omega\,(\xi-x)^2}{2\hbar}\right], \tag{11.4}$$

где $|k| = 1$.

Неравенство (11.2) есть *соотношение неопределенностей Гейзенберга*, а формула (11.4) описывает векторы состояний *минимальной неопределенности*, для которых $\mathsf{D}_\psi(q) = \frac{\hbar}{2\omega}, \mathsf{D}_\psi(p) = \frac{\omega\hbar}{2}$. Далее полагаем $k = \exp\left(-\frac{ixy}{2\hbar}\right)$, что приводит к

$$\psi\,(\xi) = \sqrt[4]{\frac{\omega}{\pi\hbar}}\exp\left[\frac{iy}{\hbar}\left(\xi - \frac{x}{2}\right) - \frac{\omega\,(\xi-x)^2}{2\hbar}\right]. \tag{11.5}$$

Вводя комплексные комбинации

$$a = \frac{1}{\sqrt{2\hbar\omega}}\left(\omega q + ip\right); \quad \zeta = \frac{1}{\sqrt{2\hbar\omega}}\left(\omega x + iy\right)$$

и обозначая соответствующий вектор состояния (11.5) через $|\zeta\rangle$, можно переписать определяющее уравнение (11.3) в виде

$$a|\zeta\rangle = \zeta|\zeta\rangle; \quad \zeta \in \mathbb{C}. \tag{11.6}$$

В частности,

$$a|0\rangle = 0, \tag{11.7}$$

где $|0\rangle$ – вектор, соответствующий функции

$$\psi\left(\xi\right) = \sqrt[4]{\frac{\omega}{\pi\hbar}} \exp\left[-\frac{\omega\xi^2}{2\hbar}\right].$$

Оператор $a = \frac{1}{\sqrt{2\hbar\omega}}\left(\omega q + ip\right)$ имеет сопряженный $a^\dagger = \frac{1}{\sqrt{2\hbar\omega}}\left(\omega q - ip\right)$, а коммутационное соотношение (11.1) принимает вид

$$\left[a, a^\dagger\right] = I.$$

Рассмотрим оператор

$$\mathcal{N} = a^\dagger a = aa^\dagger - I, \tag{11.8}$$

который является симметричным и существенно самосопряженным на $\mathcal{D}$. В силу (11.7) имеем

$$\mathcal{N}|0\rangle = 0.$$

Задача 11.1.1 *Последовательно применяя равенство (11.8), покажите, что векторы*

$$|n\rangle = \frac{\left(a^\dagger\right)^n}{\sqrt{n!}}|0\rangle \tag{11.9}$$

образуют ортонормированную систему собственных векторов оператора $\mathcal{N}$:

$$\mathcal{N}|n\rangle = n|n\rangle; \quad n = 0, 1, \ldots$$

Соответствующие собственные функции в $L^2\left(\mathbb{R}\right)$ с точностью до нормировки имеют вид

$$\left[\omega\xi - \hbar\frac{d}{d\xi}\right]^n \exp\left[-\frac{\omega\xi^2}{2\hbar}\right] \simeq H_n\left(\sqrt{\frac{\omega}{\hbar}}\xi\right) \exp\left[-\frac{\omega\xi^2}{2\hbar}\right],$$

где $H_n; \quad n = 0, 1, \ldots$ – полиномы Эрмита, которые образуют полную систему в $L^2\left(\mathbb{R}\right)$. Следовательно, $\{|n\rangle; n = 0, 1, \ldots\}$ – ортонормированный базис в $\mathcal{H}$,

$$\sum_{n=0}^{\infty} |n\rangle\langle n| = I,$$

и $\mathcal{N}$ имеет спектральное разложение

$$\mathcal{N} = \sum_{n=0}^{\infty} n|n\rangle\langle n|. \tag{11.10}$$

Задача 11.1.2 *Оператор $\mathcal{N}$ – самосопряженный оператор типа $\mathfrak{F}$ с областью определения*

$$\mathcal{D}(\mathcal{N}) = \left\{ \psi : \sum_{n=0}^{\infty} n^2 \left|\langle n|\psi\rangle\right|^2 < \infty \right\}$$

и (11.10) выполнено в смысле сильной сходимости на $\mathcal{D}(\mathcal{N})$.

Имеют место соотношения

$$a|n\rangle = \sqrt{n}|n-1\rangle; \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.$$

Вводя группу унитарных операторов $\left\{e^{i\mathcal{N}\omega t}; t \in \mathbb{R}\right\}$, имеем

$$e^{i\mathcal{N}\omega t} a e^{-i\mathcal{N}\omega t} = a e^{-i\omega t}; \quad e^{i\mathcal{N}\omega t} a^\dagger e^{-i\mathcal{N}\omega t} = a^\dagger e^{i\omega t}. \tag{11.11}$$

Возвращаясь к операторам q, p, видим, что

$$\hbar\omega \left(\mathcal{N} + \frac{I}{2}\right) = \frac{1}{2}\left(\omega^2 q^2 + p^2\right) \equiv \hbar H \tag{11.12}$$

– *оператор энергии* квантового гармонического осциллятора с частотой ω и уравнения (11.11) описывают динамику осциллятора с гамильтонианом H. Оператор $\mathcal{N}$ называется *оператором числа квантов*, $|n\rangle$ – вектором состояния с n квантами, $|0\rangle$ – вектором *вакуумного* состояния, оператор a (соответственно $a^\dagger$) – оператором *уничтожения* (соответственно *рождения* квантов). В квантовой оптике операторы $a, a^\dagger$ описывают одну гармонику (*моду*) поля, соответствующую определенной частоте ω (а также определенной поляризации).

Состояния $|\zeta\rangle\langle\zeta|; \zeta \in \mathbb{C}$, называются в квантовой оптике *когерентными*. Из (11.9), (11.6) получаем

$$\langle n|\zeta\rangle = \frac{1}{\sqrt{n!}}\langle 0|a^n\zeta\rangle = \frac{\zeta^n}{\sqrt{n!}}\langle 0|\zeta\rangle = \frac{\zeta^n}{\sqrt{n!}} \exp\left(-\frac{|\zeta|^2}{2}\right),$$

где также была использована формула

$$\langle\zeta_1|\zeta_2\rangle = \exp\left[-\frac{1}{2}\left(|\zeta_1|^2 + |\zeta_2|^2 - 2\bar{\zeta_1}\zeta_2\right)\right]. \tag{11.13}$$

Задача 11.1.3 *Докажите (11.13), используя вещественную параметризацию (11.5) векторов* $|\zeta\rangle$.

Задача 11.1.4 *Система векторов* $\{|\zeta\rangle; \zeta \in \mathbb{C}\}$ *является переполненной в том смысле, что*

$$\frac{1}{\pi}\int |\zeta\rangle\langle\zeta| d^2\zeta = I,$$

где $d^2\zeta = \frac{1}{2\hbar} dxdy$. *Указание: матричные элементы интеграла в базисе* $\{|n\rangle\}$ *имеют вид*

$$\frac{1}{\pi}\int \frac{\zeta^n}{\sqrt{n!}}\frac{\bar{\zeta}^m}{\sqrt{m!}} \exp\left(-|\zeta|^2\right) d^2\zeta = \delta_{nm}.$$

Пусть $p(\zeta)$ – плотность распределения вероятностей на $\mathbb{C}$. Тогда

$$S = \int |\zeta\rangle\langle\zeta| p(\zeta) d^2\zeta \tag{11.14}$$

– оператор плотности в $\mathcal{H}$, который соответствует *квазиклассическому* состоянию. В частности, для комплексного гауссовского распределения с нулевым средним и дисперсией N получаем оператор плотности

$$S_0 = \frac{1}{\pi N}\int |\zeta\rangle\langle\zeta| \exp\left(-\frac{|\zeta|^2}{N}\right) d^2\zeta \tag{11.15}$$

с матричными элементами

$$\langle n|S_0|m\rangle = \frac{1}{\pi N}\int \frac{\zeta^n}{\sqrt{n!}}\frac{\bar{\zeta}^m}{\sqrt{m!}} \exp\left(-\frac{(N+1)|\zeta|^2}{N}\right) d^2\zeta = \delta_{nm}\frac{1}{N+1}\left(\frac{N}{N+1}\right)^n.$$

Поэтому S_0 имеет спектральное разложение

$$S_0 = \frac{1}{N+1}\sum_{n=0}^{\infty}\left(\frac{N}{N+1}\right)^n |n\rangle\langle n|. \tag{11.16}$$

Это выражение имеет смысл и при $N = 0$, когда S_0 определяется вырожденным распределением, сосредоточенным в точке $\zeta = 0$.

Учитывая спектральное разложение (11.10) и (11.12), последнее соотношение можно переписать в виде

$$S_0 = \frac{\exp\left[-\theta H\right]}{\mathrm{Tr}\exp\left[-\theta H\right]}, \tag{11.17}$$

что является оператором плотности гиббсовского состояния квантового гармонического осциллятора , т. е. равновесного состояния при обратной температуре $\theta/\hbar = \frac{1}{\hbar\omega}\ln\frac{N+1}{N}$. Величина $N = \mathrm{Tr}S_0\mathcal{N} = \mathrm{Tr}S_0 a^\dagger a$ интерпретируется как среднее число квантов энергии осциллятора.

Другой важный оператор плотности соответствует смещенной плотности распределения

$$S_\mu = \frac{1}{\pi N}\int |\zeta\rangle\langle\zeta| \exp\left(-\frac{|\zeta-\mu|^2}{N}\right) d^2\zeta, \tag{11.18}$$

где $\mu = \frac{1}{\sqrt{2\hbar\omega}}(\omega m_q + i m_p)$. Введем унитарные операторы

$$D(\zeta)\psi(\xi) = \exp\left[\frac{iy}{\hbar}\left(\xi - \frac{x}{2}\right)\right]\psi(\xi - x). \tag{11.19}$$

Очевидно, что

$$|\zeta\rangle = D(\zeta)|0\rangle;$$
$$D(\zeta_2)|\zeta_1\rangle = \exp(i\Im\bar{\zeta_1}\zeta_2)|\zeta_1+\zeta_2\rangle,$$

т. е. действие этих операторов приводит к смещению средних значений наблюдаемых q, p. В физическом эксперименте такое смещение происходит при внешнем воздействии, осуществляемом идеальным лазером, которое преобразует вакуум $|0\rangle\langle 0|$ в когерентное состояние $|\zeta\rangle\langle\zeta|$. Квазиклассическое состояние (11.14) представляет собой статистическую смесь когерентных состояний, тогда как

$$S_\mu = D(\mu)S_0 D(\mu)^* \tag{11.20}$$

есть преобразование равновесного состояния S_0 при внешнем воздействии. Заметим, что состояние S_μ является чистым тогда и только тогда, когда $N = 0$, причем в этом случае оно совпадает с когерентным состоянием $|\mu\rangle\langle\mu|$.

Задача 11.1.5 *Докажите, что операторы смещения удовлетворяют коммутационному соотношению*

$$D(\zeta_2)D(\zeta_1) = \exp(i\Im\bar{\zeta_1}\zeta_2)D(\zeta_1+\zeta_2), \tag{11.21}$$

из которого следует

$$D(\mu)^* D(\zeta)D(\mu) = \exp(2i\Im\bar{\mu}\zeta)D(\zeta). \tag{11.22}$$

В дальнейшем нам понадобится некоммутативный аналог характеристической функции для состояния S_μ:

$$\mathrm{Tr}S_\mu D(\zeta) = \exp\left[2i\Im\bar{\mu}\zeta - \left(N+\frac{1}{2}\right)|\zeta|^2\right]. \tag{11.23}$$

Доказательство формулы (11.23). В силу (11.20), (11.22) имеем

$$\mathrm{Tr}S_\mu D(\zeta) = \exp\left(2i\Im\bar{\mu}\zeta\right)\mathrm{Tr}S_0 D(\zeta)$$

и достаточно доказать (11.23) для $\mu = 0$. Имеем

$$\begin{aligned}\operatorname{Tr} S_0 D(\zeta) &= \int \langle \zeta_1 | D(\zeta) | \zeta_1 \rangle \frac{1}{\pi N} \exp\left(-\frac{|\zeta_1|^2}{N}\right) d^2\zeta_1 \qquad (11.24)\\ &= \int \langle 0 | D(\zeta_1)^* D(\zeta) D(\zeta_1) | 0 \rangle \frac{1}{\pi N} \exp\left(-\frac{|\zeta_1|^2}{N}\right) d^2\zeta_1 \\ &= \int \exp(i\Im \bar{\zeta}_1 \zeta) \langle 0 | D(\zeta) | 0 \rangle \frac{1}{\pi N} \exp\left(-\frac{|\zeta_1|^2}{N}\right) d^2\zeta_1\end{aligned}$$

Используя (11.13), получаем

$$\langle 0 | D(\zeta) | 0 \rangle = \langle 0 | \zeta \rangle = \exp\left(-\frac{1}{2}|\zeta|^2\right).$$

Подставляя это выражение в (11.24) и используя выражение для характеристической функции гауссовского распределения, получаем

$$\operatorname{Tr} S_0 D(\zeta) = \exp\left(-\left(N + \frac{1}{2}\right)|\zeta|^2\right).$$

□

11.2 Канонические коммутационные соотношения

В квантовой механике канонические коммутационные соотношения (ККС) возникают при квантовании классической системы с конечным либо бесконечным числом степеней свободы, т. е. поля. В квантовой оптике обычно имеют дело только с конечным набором существенных частот, что сводит проблему к рассмотрению системы с конечным числом s степеней свободы.

Рассмотрим гильбертово пространство $\mathcal{H} = L^2(\mathbb{R}^s)$ комплексных квадратично интегрируемых функций вещественных переменных $\xi_j; j = 1, \ldots, s$, где $\mathbb{R}^s$ является координатным пространством рассматриваемой классической системы. В пространстве $\mathcal{H}$ зададим две группы унитарных операторов

$$V_x \psi(\xi) = \exp(i\xi^\top x)\psi(\xi); \quad U_y \psi(\xi) = \psi(\xi + \hbar y), \qquad (11.25)$$

где $\xi, x, y \in \mathbf{R}^s$ рассматриваются как векторы-столбцы, а $^\top$ обозначает транспонирование. Эти группы удовлетворяют каноническим коммутационным соотношениям (ККС) Вейля

$$U_y V_x = \exp(i\hbar y^\top x) V_x U_y. \qquad (11.26)$$

Отметим аналогию с дискретными операторами Вейля, определенными соотношениями (6.47). Операторы U_y, V_x описывают изменение квантового состояния при смещении координаты, соответственно импульса, а

соотношения Вейля задают кинематику квантово-механической системы; на самом деле, эти соотношения могут быть выведены из принципа галилеевой относительности.

Чтобы использовать очевидную симметрию между x, y, введем вектор-столбец вещественных параметров $z = [x_1, y_1, \dots, x_s, y_s]^\top$ и унитарные операторы

$$W(z) = \exp(\frac{i}{2}\hbar y^\top x) V_x U_y. \tag{11.27}$$

Из (11.27) и (11.26) следует, что операторы $W(z)$ удовлетворяют *ККС Вейля-Сигала*

$$W(z)W(z') = \exp[\frac{i}{2}\Delta(z, z')]W(z + z'), \tag{11.28}$$

где

$$\Delta(z, z') = \hbar \sum_{j=1}^{s} (x'_j y_j - x_j y'_j) \tag{11.29}$$

– каноническая *симплектическая форма*. Из (11.28) и унитарности операторов $W(z)$ в частности следует

$$W(-z) = W(z)^* \quad z \in Z.$$

Из (11.28) следует

$$W(z')^* W(z) W(z') = \exp[i\Delta(z, z')]W(z). \tag{11.30}$$

Сходство коммутационных соотношений (11.21) и (11.28) неслучайно: на самом деле операторы $W(z)$ с точностью до параметризации представляют собой многомодовое обобщение унитарных операторов смещения.

Задача 11.2.1 *Покажите, что в случае одной степени свободы*

$$D(\mu) = W(m_p/\hbar, -m_q/\hbar), \tag{11.31}$$

так что

$$D(\mu)^* W(z) D(\mu) = \exp i\,(m_q x + m_p y)\, W(z). \tag{11.32}$$

Поскольку $\Delta(z, z) \equiv 0$, из (11.28) вытекает, что для любого фиксированного z однопараметрическое семейство $\{W(tz); t \in \mathbf{R}\}$ является унитарной группой. Согласно теореме Стоуна (см. раздел 10.8), существует самосопряженный оператор $R(z)$, такой что

$$W(z) = \exp i\, R(z). \tag{11.33}$$

Задача 11.2.2 *Используйте (11.30), чтобы доказать соотношение*

$$W(z')^* R(z) W(z') = R(z) + \Delta(z, z')I. \tag{11.34}$$

Задача 11.2.3 *Проверьте, что соответствие $z \to R(z)$ является линейным, и что операторы $R(z)$ удовлетворяют коммутационным соотношениям*

$$[R(z), R(z')] = -i\Delta(z, z')I. \tag{11.35}$$

Указание: в соотношении (11.28) замените z, z' на tz, sz' и рассмотрите первые и вторую смешанную производные по s, t при $s = t = 0$.

Пространство $Z = \mathbb{R}^{2s}$, снабженное невырожденной кососимметричной формой $\Delta(z, z')$, является *симплектическим пространством*. Симплектическое пространство (Z, Δ) описывает фазовое пространство классической механической системы, квантование которой дается семейством унитарных операторов $W(z)$ в гильбертовом пространстве $\mathcal{H}$. Такое квантование в существенном единственно: всякое неприводимое семейство унитарных операторов $W(z)$ в каком-либо гильбертовом пространстве, удовлетворяющих ККС Вейля-Сигала, унитарно эквивалентно описанному выше представлению в $L^2(\mathbb{R}^s)$, называемому представлением Шредингера (теорема единственности Стоуна-фон Неймана, см. Комментарии). В дальнейшем $W(z); z \in Z$, – некоторое неприводимое представление ККС.

Базис $\{e_j, h_j; j = 1, \ldots, s\}$ в Z называется *симплектическим*, если

$$\Delta(e_j, h_k) = -\hbar\delta_{jk},\ \Delta(e_j, e_k) = \Delta(h_j, h_k) = 0; \quad j, k = 1, \ldots, s. \tag{11.36}$$

В таком базисе симплектическая форма $\Delta(z, z')$ имеет стандартный вид (11.29) и

$$R(z) = Rz = \sum_{j=1}^{s} (x_j q_j + y_j p_j),$$

где

$$R = [q_1, p_1, \ldots, q_s, p_s]$$

– вектор-строка самосопряженных операторов – *канонических наблюдаемых*, удовлетворяющих ККС Гейзенберга

$$[q_j, p_k] = i\delta_{jk}\hbar I, \ \ [q_j, q_k] = 0, \ \ [p_j, p_k] = 0. \tag{11.37}$$

Представление Шредингера ККС (11.37) дается самосопряженными операторами

$$q_j = \xi_j, \quad p_j = \frac{\hbar}{i}\frac{\partial}{\partial \xi_j}$$

в гильбертовом пространстве $L^2(\mathbb{R}^s)$. Операторы q_j, p_j не ограничены, поэтому соотношения (11.37) следует рассматривать только на общей плотной области определения, в качестве которой удобно выбрать пространство $\mathcal{S}(\mathbb{R})$. Мы уже видели, что подход, основанный на использовании ККС Вейля, является более фундаментальным с физической точки

зрения. С другой стороны, он также позволяет избежать ряда не относящихся к делу математических патологий, связанных с областями определения неограниченных операторов.

В дальнейшем в пространстве $Z = \mathbb{R}^{2s}$ будут рассматриваться различные билинейные формы $\Delta, \alpha, \ldots$, при этом матрицы форм будут обозначаться тем же символом, т. е. $\alpha(z, z') = z^\top \alpha z'$, и т. д. В частности, $\Delta(z, z') = z^\top \Delta z'$, где

$$\Delta = \begin{bmatrix} 0 & -\hbar & & & \\ \hbar & 0 & & & \\ & & \ddots & & \\ & & & 0 & -\hbar \\ & & & \hbar & 0 \end{bmatrix} \equiv \operatorname{diag} \begin{bmatrix} 0 & -\hbar \\ \hbar & 0 \end{bmatrix}. \tag{11.38}$$

Лемма 11.2.1 *Пусть $\alpha(z, z') = z^\top \alpha z'$ – произвольное скалярное произведение в симплектическом пространстве (Z, Δ). Существует симплектический базис $\{e_j, h_j; j = 1, \ldots, s\}$ в Z, в котором форма α имеет диагональный вид с матрицей*

$$\tilde{\alpha} = \operatorname{diag} \begin{bmatrix} \alpha_j & 0 \\ 0 & \alpha_j \end{bmatrix},$$

причем $\alpha_j > 0$.

Доказательство. Рассмотрим матрицу $\hat{\alpha} = \Delta^{-1}\alpha$, которая является матрицей оператора (обозначаемого тем же символом), связывающего две формы:

$$\alpha(z, z') = \Delta(z, \hat{\alpha} z').$$

Оператор $\hat{\alpha}$ является кососимметричным в евклидовом пространстве (Z, α) : $\hat{\alpha}^* = -\hat{\alpha}$. По известной теореме линейной алгебры существует ортогональный базис $\{e_j, h_j\}$ в (Z, α) и положительные числа $\{\alpha_j\}$ такие, что

$$\hat{\alpha} e_j = -\hbar^{-1} \alpha_j h_j; \quad \hat{\alpha} h_j = \hbar^{-1} \alpha_j e_j.$$

Выбирая нормировку $\alpha(e_j, e_j) = \alpha(h_j, h_j) = \alpha_j$, получаем симплектический базис в (Z, Δ) с требуемыми свойствами. □

Обозначая через T матрицу перехода от исходного симплектического базиса в Z к новому симплектическому базису, т. е. матрицу со столбцами $\{e_j, h_j; j = 1, \ldots, s\}$, имеем

$$\Delta = T^\top \Delta T; \quad \tilde{\alpha} = T^\top \alpha T.$$

Канонические наблюдаемые в новом базисе даются соотношениями

$$\tilde{q}_j = R(e_j), \tilde{p}_j = R(h_j); j = 1, \ldots, s,$$

так что $RT = \tilde{R}$, где $\tilde{R} = [\tilde{q}_1, \tilde{p}_1, \ldots, \tilde{q}_s, \tilde{p}_s]^\top$. Матрица T задает *симплектическое преобразование* в (Z, Δ), которое характеризуется свойством

$$\Delta(Tz, Tz') = \Delta(z, z'); \quad z, z' \in Z.$$

11.3 Динамика, квадратичные операторы и комплексные структуры

Для любого симплектического преобразования T в Z существует *каноническое унитарное преобразование* U_T в $\mathcal{H}$ такое, что

$$U_T^* W(z) U_T = W(Tz).$$

Это следует из теоремы единственности, поскольку операторы $W(Tz); z \in Z$, также образуют (неприводимое) представление ККС (11.28) в пространстве $\mathcal{H}$. В терминах канонических наблюдаемых

$$U_T^* R U_T = RT. \tag{11.39}$$

Для получения явного выражения для операторов U_T рассмотрим однопараметрическую группу

$$T_t = \mathrm{e}^{tD}; \quad t \in \mathbb{R}$$

симплектических преобразований в Z, которая описывает линейную динамику в классическом фазовом пространстве. Здесь D – генератор, который удовлетворяет условию

$$\Delta(T_t z, T_t z') \equiv \Delta(z, z').$$

Будем предполагать, что оператор D является строго $\Delta-$*положительным*, т. е. $\Delta(z, Dz')$ является скалярным произведением в Z. Отсюда получаем, в матричных обозначениях

$$\Delta D = -D^\top \Delta \geq 0. \tag{11.40}$$

Квантование линейной динамики T_t дается следующей теоремой.

Теорема 11.3.1 *Рассмотрим оператор $H = \frac{1}{\hbar} R\epsilon R^\top$ в $\mathcal{H}$, квадратичный относительно канонических наблюдаемых R, где $\epsilon = -\frac{\hbar}{2} D\Delta^{-1}$. Тогда H – положительный самосопряженный оператор в $\mathcal{H}$, порождающий унитарную группу $\{\mathrm{e}^{itH}\}$ в $\mathcal{H}$, такую что*

$$W(T_t z) = \mathrm{e}^{itH} W(z) \mathrm{e}^{-itH}. \tag{11.41}$$

Доказательство (набросок). Рассмотрим билинейную форму

$$\epsilon(z,z') = -z^\top \Delta\epsilon\Delta z' = \frac{\hbar}{2}\Delta(z,Dz'),$$

которая является скалярным произведением в силу предположения об операторе D. Согласно лемме 11.2.1 существует симплектический базис $\{e_j,h_j\}$, в котором матрица этой формы имеет диагональный вид

$$-\Delta\tilde{\epsilon}\Delta = \hbar^2\mathrm{diag}\begin{bmatrix}\omega_j/2 & 0\\ 0 & \omega_j/2\end{bmatrix},$$

где $\omega_j > 0$, следовательно

$$\tilde{\epsilon} = \mathrm{diag}\begin{bmatrix}\omega_j/2 & 0\\ 0 & \omega_j/2\end{bmatrix}.$$

Вводя канонические наблюдаемые $\tilde{q}_j = R(e_j), \tilde{p}_j = R(h_j); j = 1,\dots,s$, имеем

$$H = \sum_{j=1}^{s}\frac{\omega_j}{2\hbar}\left(\tilde{q}_j{}^2 + \tilde{p}_j{}^2\right), \tag{11.42}$$

откуда видно, что H – положительный самосопряженный оператор, как сумма операторов типа (11.12), относящихся к разным модам. Определяя *операторы рождения-уничтожения*

$$\tilde{a}_j^\dagger = \frac{1}{\sqrt{2\hbar}}\left(\tilde{q}_j - i\tilde{p}_j\right), \tilde{a}_j = \frac{1}{\sqrt{2\hbar}}\left(\tilde{q}_j + i\tilde{p}_j\right),$$

удовлетворяющие коммутационным соотношениям

$$[\tilde{a}_j,\tilde{a}_k^\dagger] = \delta_{jk}I, \tag{11.43}$$

имеем

$$H = \sum_{j=1}^{s}\omega_j\left(\tilde{a}_j^\dagger\tilde{a}_j + I/2\right). \tag{11.44}$$

В терминах генераторов канонических наблюдаемых соотношение (11.41) сводится к

$$R\mathrm{e}^{tD} = \mathrm{e}^{-it\frac{1}{2}RD\Delta^{-1}R^\top}R\mathrm{e}^{it\frac{1}{2}RD\Delta^{-1}R^\top}, \tag{11.45}$$

где $D = -2\epsilon\Delta/\hbar$. В базисе $\{e_j,h_j\}$ матрица оператора D имеет вид

$$\tilde{D} = -2\hbar^{-1}\tilde{\epsilon}\Delta = \mathrm{diag}\begin{bmatrix}0 & \omega_j\\ -\omega_j & 0\end{bmatrix},$$

следовательно, (11.45) сводится к уравнениям типа (11.11)

$$\tilde{a}_j \mathrm{e}^{-i\omega_j t} = \mathrm{e}^{i\omega_j t \tilde{a}_j^\dagger \tilde{a}_j} \tilde{a}_j \mathrm{e}^{-i\omega_j t \tilde{a}_j^\dagger \tilde{a}_j}; \quad j = 1, \ldots, s, \tag{11.46}$$

описывающим динамику квантовых гармонических осцилляторов с частотами ω_j. □

Подход, использующий представление энергетической матрицы ϵ в диагональной форме, называется в квантовой оптике разложением на нормальные моды.

Оператор J в (Z, Δ) называется *оператором комплексной структуры*, если он $\Delta-$положителен и

$$J^2 = -1, \tag{11.47}$$

где 1 – единичный оператор в Z. В частности,

$$\Delta J = -J^\top \Delta \geq 0. \tag{11.48}$$

Задача 11.3.1 *Оператор J порождает в Z структуру комплексного унитарного пространства (размерности s), если положить по определению $iz = Jz$ и ввести скалярное произведение*

$$h(z, z') = j(z, z') + i\Delta(z, z') = \Delta(z, Jz') + i\Delta(z, z').$$

С любой комплексной структурой можно связать циклическую однопараметрическую группу $\left\{\mathrm{e}^{\varphi J}\right\}$ симплектических преобразований, которую будем называть *калибровочной группой*. Согласно доказанной теореме, калибровочная группа в Z порождает унитарную группу *калибровочных преобразований* в $\mathcal{H}$ по формуле

$$W(\mathrm{e}^{\varphi J} z) = \mathrm{e}^{-i\varphi G} W(z) \mathrm{e}^{i\varphi G}, \tag{11.49}$$

где $G = -\frac{1}{2} R J \Delta^{-1} R^\top$ – самосопряженный положительный оператор в $\mathcal{H}$. В терминах канонических наблюдаемых, соотношение (11.49) сводится к линейному преобразованию

$$R\mathrm{e}^{\varphi J} = \mathrm{e}^{-i\varphi G} R \mathrm{e}^{i\varphi G}, \tag{11.50}$$

которое является частным случаем (11.45).

Из доказательства теоремы в случае $D = J$, с использованием дополнительного свойства (11.47), вытекает существование симплектического базиса $\{e_j, h_j; j = 1, \ldots, s\}$, ассоциированного со скалярным произведением $j(z, z') = \Delta(z, Jz')$, в котором оператор J имеет матрицу

$$\tilde{J} = \mathrm{diag} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{11.51}$$

В этом базисе

$$G = \sum_{j=1}^{s} \frac{1}{2\hbar} \left({\tilde{q}_j}^2 + {\tilde{p}_j}^2 \right) = \sum_{j=1}^{s} \left(\tilde{a}_j^\dagger \tilde{a}_j + I/2 \right), \tag{11.52}$$

что с точностью до несущественного слагаемого совпадает с *оператором полного числа квантов*

$$\mathcal{N} = \sum_{j=1}^{s} \tilde{a}_j^\dagger \tilde{a}_j.$$

Соотношение (11.50) сводится к (11.46) с заменой $\omega_j t$ на φ.

Оператор A в $\mathcal{H}$ называется *калибровочно инвариантным*, если

$$\mathrm{e}^{-i\varphi G} A \mathrm{e}^{i\varphi G} = A$$

для всех вещественных φ. Используя (11.50), получаем, что квадратичный оператор $A = R\epsilon R^\top$, где ϵ – симметричная положительная матрица, является калибровочно инвариантным, если $\mathrm{e}^{\varphi J} \epsilon \mathrm{e}^{\varphi J^\top} = \epsilon$, или, эквивалентно, $J\epsilon + \epsilon J^\top = 0$, что с учетом (11.48) равносильно условию

$$[J, \epsilon \Delta] = 0.$$

Для любой энергетической матрицы ϵ существует оператор комплексной структуры, удовлетворяющий этому условию: это оператор $J = T\tilde{J}T^{-1}$, который имеет матрицу (11.51) в симплектическом базисе, ассоциированном со скалярным произведением $\epsilon(z, z') = \Delta(z, \epsilon \Delta z')$.

Пример Пусть $\omega_j; j = 1, \ldots, s$ – положительные числа, и

$$H = \sum_{j=1}^{s} \frac{1}{2\hbar} \left(\omega_j^2 {q_j}^2 + {p_j}^2 \right)$$

– гамильтониан ансамбля осцилляторов с частотами ω_j, тогда $\tilde{q}_j = \sqrt{\omega_j} q_j, \tilde{p}_j = \sqrt{\omega_j}^{-1} p_j$, так что

$$\tilde{a}_j^\dagger = \frac{1}{\sqrt{2\hbar\omega_j}} \left(\omega_j q_j - i p_j \right), \quad \tilde{a}_j = \frac{1}{\sqrt{2\hbar\omega_j}} \left(\omega_j q_j + i p_j \right)$$

и

$$H = \sum_{j=1}^{s} \frac{\omega_j}{2\hbar} \left({\tilde{q}_j}^2 + {\tilde{p}_j}^2 \right) = \sum_{j=1}^{s} \omega_j \left(\tilde{a}_j^\dagger \tilde{a}_j + I/2 \right).$$

Соответствующий оператор комплексной структуры имеет вид

$$J = \mathrm{diag} \begin{bmatrix} 0 & \omega_j \\ -\omega_j^{-1} & 0 \end{bmatrix},$$

что соответствует умножению на i операторов уничтожения $\tilde{a}_j$, при этом калибровочный оператор в пространстве $\mathcal{H}$ дается формулой (11.52), т. е. равен $\mathcal{N} + \frac{s}{2} I$.

11.4 Гауссовские состояния

11.4.1 Характеристическая функция

Характеристическая функция квантового состояния S определяется соотношением

$$\phi(z) = \mathrm{Tr} SW(z); \quad z \in Z.$$

Для произвольного ядерного оператора S это выражение можно рассматривать как некоммутативный аналог преобразования Фурье, однозначно определяющий оператор S.

Положительность оператора S влечет неотрицательную определенность $n \times n-$матриц с элементами

$$\phi(z_r - z_s) \exp\left[-\frac{i}{2}\Delta(z_r, z_s)\right] = \mathrm{Tr} W(z_s)^* S W(z_r). \tag{11.53}$$

для всех $n = 1, 2, \ldots$ и всевозможных наборов $\{z_1, \ldots, z_n\} \in Z$.

Оператор плотности S имеет *конечные вторые моменты*, если $\mathrm{Tr} S q_j^2 < \infty, \mathrm{Tr} S p_j^2 < \infty$ для всех j, где след определяется как в (10.6). В этом случае определены *вектор средних значений* $m = \mathrm{Tr} SR$ и *матрица ковариаций* $\mathsf{B}_S(R) = \alpha$ (ср. раздел 2.2.3). В силу ККС (11.35) коммутационная матрица $\mathsf{C}_S(R) = \Delta$, так что

$$\alpha - \frac{i}{2}\Delta = \mathrm{Tr}(R - m)^\top S(R - m). \tag{11.54}$$

Отсюда вытекает неравенство

$$\alpha \geq \frac{i}{2}\Delta, \tag{11.55}$$

которое есть не что иное, как соотношение неопределенностей (2.21) для канонических наблюдаемых R. Далее будет показано (теорема 11.4.1), что условие (11.55) не только необходимо, но и достаточно для того, чтобы α была матрицей ковариаций некоторого состояния. В дальнейшем будем обозначать $\Sigma(m, \alpha)$ множество всех состояний с фиксированным вектором средних значений m и матрицей ковариаций α.

Как и в классическом случае, компоненты m, α (как и моменты более высокого порядка) можно выразить через производные характеристической функции.

Задача 11.4.1 *Предполагая существование моментов соответствующего порядка, доказать, что*

$$\mathrm{Tr} SR(z)^n = i^{-n} \left.\frac{d^n}{dt^n}\phi(tz)\right|_{t=0}.$$

11.4.2 Определение и свойства гауссовских состояний

Состояние S называется *гауссовским*, если его характеристическая функция $\phi(z) = \mathrm{Tr}SW(z)$ имеет вид

$$\phi(z) = \exp\left(i\, m(z) - \frac{1}{2}\alpha(z,z)\right), \tag{11.56}$$

где $m(z) = mz$ –линейная форма, а $\alpha(z,z) = z^\top \alpha z$ – билинейная форма на (Z, Δ). Здесь m – вещественный $(2s)$-вектор-строка, а α – вещественная симметричная $(2s) \times (2s)$-матрица.

Задача 11.4.2 *Найдите первые и вторые производные функции (11.56) при $z = 0$, и покажите, используя результат задачи 11.4.1, что m – вектор средних значений, а α – матрица ковариаций состояния S.*

Теорема 11.4.1 *Для того, чтобы соотношение (11.56) определяло квантовое состояние, необходимо и достаточно, чтобы матрица α удовлетворяла условию (11.55). При выполнении этого условия формула (11.56) определяет единственное гауссовское состояние в $\Sigma(m, \alpha)$.*

Доказательство. Как уже было отмечено, необходимость вытекает из соотношения (11.54). Альтернативное доказательство вытекает из следующей леммы, которая понадобится и в дальнейшем.

Лемма 11.4.2 *Пусть α – скалярное произведение, Δ – кососимметричная форма в Z, такие что для всех $n = 1, 2, \ldots$ и всевозможных наборов $\{z_1, \ldots, z_n\} \in Z$, эрмитовы матрицы с элементами*

$$\exp\{\alpha(z_r, z_s) - \frac{i}{2}\Delta(z_r, z_s)\} \tag{11.57}$$

неотрицательно определены. Тогда матрицы форм α, Δ удовлетворяют условию (11.55).

Доказательство. Пусть $t > 0$. Записывая условие неотрицательной определенности для набора $\{0, \sqrt{t}z_1, \ldots, \sqrt{t}z_n\}$ с переменными $\{c_0, c_1, \ldots, c_n\} \in \mathbb{C}$ такими, что $c_0 = -\sum_{j=1}^n c_j$, имеем

$$\sum_{r,s=1}^n \bar{c}_r c_s \left(\exp t\left[\alpha(z_r, z_s) - \frac{i}{2}\Delta(z_r, z_s)\right] - 1\right) \geq 0.$$

Деля на t и устремляя t к 0, получаем, что матрицы с элементами $\alpha(z_r, z_s) - \frac{i}{2}\Delta(z_r, z_s)$ неотрицательно определены для всевозможных наборов $z_1, \ldots, z_n$, что равносильно условию (11.55). □

Подставляя в (11.53) гауссовское выражение (11.56), можно видеть, что

$$\phi(z_r - z_s)\,\exp\left[-\frac{i}{2}\Delta(z_r, z_s)\right] = \phi(z_r)\overline{\phi(z_s)}\exp\{\alpha(z_r, z_s) - \frac{i}{2}\Delta(z_r, z_s)\},$$

таким образом, условие леммы выполнено, откуда следует неравенство (11.55).

Пусть теперь матрица α удовлетворяет неравенству (11.55). Из леммы 11.2.1 следует, что существует симплектическое преобразование T, такое что

$$\tilde{\alpha} = T^\top \alpha T = \operatorname{diag}\begin{bmatrix}\alpha_j & 0\\ 0 & \alpha_j\end{bmatrix}, \tag{11.58}$$

где $\alpha_j > 0$, причем неравенство (11.55) равносильно тому, что $\alpha_j \geq \frac{\hbar}{2}, \quad j = 1, \ldots, s$. Полагая $\alpha_j = \hbar\left(N_j + \frac{1}{2}\right)$, получаем $N_j \geq 0$.

Из (11.58) следует, что

$$\phi(Tz) = \exp(i\tilde{m}^\top z - \frac{1}{2} z^\top \tilde{\alpha} z)$$

$$= \exp\sum_{j=1}^{s}[i(\tilde{m}_{jq}x_j + \tilde{m}_{jp}y_j) - \frac{1}{2}\alpha_j(x_j^2 + y_j^2)],$$

где $\tilde{m} = T^\top m$. Однако в силу соотношений (11.23), (11.31)

$$\phi_j(z) = \exp\left[i(\tilde{m}_{jq}x + \tilde{m}_{jp}y) - \frac{\alpha_j}{2}(x^2 + y^2)\right] \tag{11.59}$$

является характеристической функцией оператора плотности типа (11.18) для одной степени свободы (при $\omega = 1$), в канонических наблюдаемых $\tilde{R} = RT$. Обозначая $S^{(j)}$ соответствующее состояние и полагая

$$S = \bigotimes_{j=1}^{s} S^{(j)}, \tag{11.60}$$

получаем

$$\phi(Tz) = \operatorname{Tr} S\,\exp i\,\tilde{R}z = \operatorname{Tr} S\,W(Tz),$$

откуда следует, что состояние S имеет характеристическую функцию $\phi(z)$. □

В квантовой оптике представление гауссовского состояния в виде (11.60) обычно связано с разложением на нормальные моды.

Пусть J – оператор комплексной структуры в Z и $\left\{\mathrm{e}^{i\varphi G}\right\}$ – соответствующая калибровочная группа в $\mathcal{H}$. Из (11.49) следует, что гауссовский оператор плотности S является калибровочно инвариантным, $\mathrm{e}^{-i\varphi G} S\,\mathrm{e}^{i\varphi G} = S, \varphi \in \mathbf{R}$, тогда и только тогда, когда его характеристическая функция удовлетворяет условию $\phi(\mathrm{e}^{\varphi J} z) = \phi(z)$, что равносильно

условиям $m = 0$ и $J^\top \alpha + \alpha J = 0$. Используя (11.48), последнее равенство можно переписать в виде

$$[J, \Delta^{-1}\alpha] = 0. \tag{11.61}$$

Напомним, что $\hat{\alpha} = \Delta^{-1}\alpha$ является матрицей кососимметричного оператора. Для любой матрицы ковариаций α найдется по крайней мере один оператор комплексной структуры в Z, коммутирующий с оператором $\Delta^{-1}\alpha$, а именно, ортогональный оператор J из полярного разложения

$$\hat{\alpha} = \left|\Delta^{-1}\alpha\right| J = J \left|\Delta^{-1}\alpha\right| \tag{11.62}$$

в евклидовом пространстве (Z, α). Пусть T – симплектическое преобразование, диагонализующее α, тогда $J = T\tilde{J}T^{-1}$, где $\tilde{J}$ определено соотношением (11.51). Более того, $\left|\Delta^{-1}\alpha\right| = T\tilde{A}T^{-1}$, где

$$\tilde{A} = \operatorname{diag}\begin{bmatrix} \alpha_j/\hbar & 0 \\ 0 & \alpha_j/\hbar \end{bmatrix} = \operatorname{diag}\begin{bmatrix} N_j + \frac{1}{2} & 0 \\ 0 & N_j + \frac{1}{2} \end{bmatrix}.$$

Задача 11.4.3 *Гауссовское состояние S является чистым тогда и только тогда, когда выполняется одно из равносильных условий:*

i. $N_j = 0; j = 1, \ldots, s;$
ii. $|\det(2\Delta^{-1}\alpha)| = 1;$
iii. $\alpha = \frac{1}{2}\Delta J$, *где* J – *оператор комплексной структуры.*

При выполнении этих условий множество $\Sigma(m, \alpha)$ состоит из одного чистого гауссовского состояния.

Пусть S – гауссовское состояние с параметрами $(0, \alpha)$, J – оператор комплексной структуры из полярного разложения (11.62), и S_0 – чистое гауссовское состояние с параметрами $(0, \frac{1}{2}\Delta J)$. Таким образом, оба этих состояния калибровочно-инвариантны.

Задача 11.4.4 *Докажите следующее многомерное обобщение формулы (11.15)*

$$S = \int W(z) S_0 W(z)^* P(d^{2s}z) = \int W(Jz) S_0 W(Jz)^* P(d^{2s}z), \tag{11.63}$$

где P – гауссовское распределение вероятностей с симплектической характеристической функцией

$$\int \mathfrak{e}^{i\Delta(w,z)} P(d^{2s}z) = \exp\left[-w^\top \left(\alpha - \frac{1}{2}\Delta J\right) w\right].$$

Это распределение калибровочно-инвариантно, т. е. инвариантно относительно действия оператора комплексной структуры J_A в Z_A. Заметим также, что состояния $W(Jz)S_0W(Jz)^*$ являются когерентными относительно рассматриваемой комплексной структуры.

11.4.3 Оператор плотности гауссовского состояния

Заметим, что $\Delta^{-1}\alpha = T(\Delta^{-1}\tilde{\alpha})T^{-1}$, где

$$\Delta^{-1}\tilde{\alpha} = \operatorname{diag}\begin{bmatrix} 0 & \alpha_j/\hbar \\ -\alpha_j/\hbar & 0 \end{bmatrix}. \tag{11.64}$$

Следовательно, числа $\pm i\alpha_j/\hbar$; $j = 1, \ldots, s$ являются собственными значениями оператора $\Delta^{-1}\alpha$.

Задача 11.4.5 *Оператор $-(\Delta^{-1}\alpha)^2 - \frac{1}{4}I$ является положительным в евклидовом пространстве (Z, α), причем состояние S является чистым тогда и только тогда, когда этот оператор равен 0, т. е.*

$$(\Delta^{-1}\alpha)^2 = -\frac{1}{4}I. \tag{11.65}$$

Спектральное разложение произвольного гауссовского состояния (11.60) можно получить как тензорное произведение спектральных разложений одномодовых состояний $S^{(j)}$. Следующая теорема обобщает представление (11.17) на случай произвольных гауссовских состояний, разложение которых не содержит "чистой" моды.

Теорема 11.4.3 *Предположим, что оператор $-(\Delta^{-1}\alpha)^2 - \frac{1}{4}I$ невырожден, тогда гауссовский оператор плотности с нулевым средним и матрицей ковариаций α представляется в виде*

$$S_0 = c\exp\left(-\frac{1}{\hbar}R\epsilon R^\top\right), \tag{11.66}$$

где

$$c = \left[\det\left(-(\Delta^{-1}\alpha)^2 - \frac{1}{4}I\right)\right]^{-\frac{1}{4}} = \left[\det\left(-4\sin^2\frac{\epsilon\Delta}{\hbar}\right)\right]^{\frac{1}{4}}, \tag{11.67}$$

а ϵ находится из соотношения

$$2\Delta^{-1}\alpha = \cot\frac{\epsilon\Delta}{\hbar}. \tag{11.68}$$

Доказательство. Переписывая представление (11.17) (при $\omega = 1$) в виде

$$\left(\mathrm{e}^{\theta/2} - \mathrm{e}^{-\theta/2}\right)\exp\left(-\frac{\theta}{2\hbar}\left(q^2 + p^2\right)\right),$$

и используя разложение (11.58), диагонализующее матрицу α, получаем

$$S_0 = c\exp\left(-\frac{1}{\hbar}\tilde{R}\tilde{\epsilon}\tilde{R}^\top\right),$$

где

$$c = \prod_{j=1}^{s} \left(\mathrm{e}^{\theta_j/2} - \mathrm{e}^{-\theta_j/2} \right), \quad \tilde{\epsilon} = \operatorname{diag} \begin{bmatrix} \theta_j/2 & 0 \\ 0 & \theta_j/2 \end{bmatrix},$$

и $\theta_j = \ln \left(\frac{\frac{\alpha_j}{\hbar} + \frac{1}{2}}{\frac{\alpha_j}{\hbar} - \frac{1}{2}} \right)$, так что

$$\mathrm{e}^{\theta_j/2} - \mathrm{e}^{-\theta_j/2} = \left[\left(\frac{\alpha_j}{\hbar} \right)^2 - \frac{1}{4} \right]^{-\frac{1}{2}}. \tag{11.69}$$

Оператор $\Delta^{-1}\alpha$ имеет собственные значения $\pm i \frac{\alpha_j}{\hbar}$, откуда следует первая из формул (11.67). Переходя к исходным каноническим наблюдаемым R, получаем (11.66), где $\epsilon = T\tilde{\epsilon}T^{\top}$.

Обращая соотношение (11.69), получаем

$$\frac{\alpha_j}{\hbar} = \frac{1}{2} \coth \frac{\theta_j}{2}. \tag{11.70}$$

Имеем

$$\frac{\tilde{\epsilon}\Delta}{\hbar} = \operatorname{diag} \begin{bmatrix} 0 & -\theta_j/2 \\ \theta_j/2 & 0 \end{bmatrix}.$$

Заметим, что $\epsilon\Delta = T\tilde{\epsilon}\Delta T^{-1}$ и $\Delta^{-1}\alpha = T\Delta^{-1}\tilde{\alpha}T^{-1}$ – матрицы операторов, причем $\Delta^{-1}\tilde{\alpha}$ дается соотношением (11.64). Операторы $-\frac{\tilde{\epsilon}\Delta}{\hbar}$ и $\tilde{\Delta}^{-1}\tilde{\alpha}$ имеют одинаковые собственные векторы с собственными значениями $\pm i\theta_j/2$ и $\pm i\alpha_j/\hbar$ соответственно, поэтому из (11.70) следует (11.68).

□

11.4.4 Энтропия гауссовского состояния

Для вычисления энтропии произвольного гауссовского состояния используем разложение на нормальные моды. Для одной моды ($s = 1$) оператор плотности $S^{(j)}$ унитарно эквивалентен оператору (11.16). Поэтому энтропия фон Неймана равна

$$H(S_0) = \frac{1}{N+1} \sum_{n=0}^{\infty} \left(\frac{N}{N+1} \right)^n [(n+1)\log(N+1) - n \log N] = g(N),$$

где введено обозначение

$$g(x) = (x+1)\log(x+1) - x \log x, \quad x > 0;$$
$$g(0) = 0.$$

Для гауссовского состояния общего вида (11.60) получаем суммированием по нормальным модам

$$H(S) = \sum_{j=1}^{s} g\left(N_j\right), \tag{11.71}$$

где $N_j = \alpha_j/\hbar - 1/2$.

Чтобы получить общее выражение, не зависящее от выбора базиса, заметим, что оператор $\Delta^{-1}\alpha$ имеет, в силу (11.64), собственные значения $\pm i\frac{\alpha_j}{\hbar}$, и следовательно, диагонализуем (в поле комплексных чисел). Для любой диагонализуемой матрицы $M = U\ \mathrm{diag}(m_j)U^{-1}$ по аналогии с другими непрерывными функциями на комплексной плоскости введем функцию $\mathrm{abs}(M) = U\mathrm{diag}(|m_j|)U^{-1}$. Тогда выражение (11.71) можно переписать в виде

$$H(S) = \frac{1}{2}\mathrm{Sp}\ g\left(\mathrm{abs}(\Delta^{-1}\alpha) - \frac{I}{2}\right), \tag{11.72}$$

где Sp – след матрицы (который следует отличать от следа Tr оператора в гильбертовом пространстве).

Другое выражение для энтропии гауссовского состояния можно получить из представления (11.66). По определению энтропии имеем

$$H(S) = -\log c + \frac{1}{\hbar}\mathrm{Sp}\epsilon\alpha,$$

что в силу (11.67) и (11.68) равно

$$\begin{aligned} &\frac{1}{4}\mathrm{Sp}\log\left[-(\Delta^{-1}\alpha)^2 - \frac{1}{4}I\right] + \mathrm{Sp}(\Delta^{-1}\alpha)\mathrm{arc\,cot}\left(2\Delta^{-1}\alpha\right) \qquad (11.73)\\ &= -\frac{1}{4}\mathrm{Sp}\log\left(-\sin^2\frac{\epsilon\Delta}{\hbar}\right) + \frac{1}{2\hbar}\mathrm{Sp}\epsilon\Delta\cot\frac{\epsilon\Delta}{\hbar}. \end{aligned}$$

Задача 11.4.6 *Используя соотношение*

$$g\left(\alpha - \frac{1}{2}\right) = \frac{1}{2}\log\left(\alpha^2 - \frac{1}{4}\right) + \alpha\log\frac{\alpha + \frac{1}{2}}{\alpha - \frac{1}{2}},$$

доказать совпадение выражений (11.72) и (11.73).

Пример В случае одномодового гауссовского оператора плотности общего вида имеем $\alpha = \begin{bmatrix} \alpha^{qq} & \alpha^{qp} \\ \alpha^{qp} & \alpha^{pp} \end{bmatrix}$, причем $\alpha^{qq}\alpha^{pp} - (\alpha^{qp})^2 \geq \frac{\hbar^2}{4}$ (это неравенство равносильно условию (11.55)). Тогда

$$-(\Delta^{-1}\alpha)^2 = \frac{\alpha^{qq}\alpha^{pp} - (\alpha^{qp})^2}{\hbar^2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

следовательно,

$$\left|\Delta^{-1}\alpha\right| = \frac{\sqrt{\alpha^{qq}\alpha^{pp} - (\alpha^{qp})^2}}{\hbar}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

и

$$H(S) = g\left(\sqrt{\frac{\alpha^{qq}\alpha^{pp} - (\alpha^{qp})^2}{\hbar^2}} - \frac{1}{2}\right).$$

Геометрически, величина $\alpha^{qq}\alpha^{pp} - (\alpha^{qp})^2$ равна квадрату площади эллипсоида рассеяния

$$\pi\hbar z^\top \alpha^{-1} z = 1, \quad z = [x, y]^\top,$$

для двумерного гауссовского распределения с матрицей ковариаций α.

Заметим, что в соответствии с (11.65) гауссовское состояние является чистым тогда и только тогда, когда оно имеет минимальную неопределенность $\alpha^{qq}\alpha^{pp} - (\alpha^{qp})^2 = \frac{\hbar^2}{4}$. Это соотношение имеет место как для когерентных состояний с

$$\alpha^{qq} = \alpha^{pp} = \frac{\hbar}{2}, \quad \alpha^{qp} = 0$$

так и для *сжатых* состояний, для которых последнее условие не выполняется (в отличие от квантовой оптики, мы используем здесь вещественную, а не комплексную параметризацию). Заметим, что состояние, которое является "сжатым" относительно данной комплексной структуры, всегда можно сделать "когерентным", используя комплексную структуру, ассоциированную с матрицей ковариаций. В многомодовом случае используется разложение на нормальные моды. В квантовой оптике выделяется та комплексная структура, для которой гамильтониан осциллятора имеет канонический вид (11.42) и когерентные/сжатые состояния определяются относительно этой комплексной структуры.

Тогда как чистые гауссовские состояния являются состояниями минимальной неопределенности канонических наблюдаемых, произвольные гауссовские состояния характеризуются следующим свойством максимальной энтропийной неопределенности при фиксированных моментах канонических наблюдаемых.

Лемма 11.4.4 *Гауссовское состояние имеет максимальную энтропию среди всех состояний с фиксированными средним m и матрицей ковариаций α. Поэтому для любого множества состояний, заданного ограничениями на первые и вторые моменты, при нахождении максимальной энтропии достаточно ограничиться рассмотрением гауссовских состояний.*

Доказательство. Пусть $S \in \Sigma(m, \alpha)$ и пусть $\tilde{S}$ – единственное гауссовское состояние в $\Sigma(m, \alpha)$. Без ограничения общности можно считать, что состояние $\tilde{S}$ невырождено. Общий случай можно свести к этому посредством отделения чистых компонент в разложении $\tilde{S}$ в тензорное произведение. Действительно, рассматривая разложение на нормальные моды, можно считать, что $Z = Z_1 + Z_2$ при $\Delta = \Delta_1 + \Delta_2, \alpha = \alpha_1 + \alpha_2$, где

α_1 удовлетворяет аналогу условия (11.65) и поэтому соответствует единственному, обязательно гауссовскому, состоянию в Z_1, в то время как α_2 соответствует невырожденному оператору плотности. Имеем

$$H(\tilde{S}) - H(S) = H\left(S;\tilde{S}\right) + \mathrm{Tr}\left(S-\tilde{S}\right)\log\tilde{S},$$

где последнее слагаемое равно нулю, поскольку для невырожденного гауссовского состояния $\tilde{S}$ оператор $\log\tilde{S}$ является многочленом второго порядка от канонических наблюдаемых в силу теоремы 11.4.3, а первые и вторые моменты состояний $S,\tilde{S}$ совпадают. Поэтому $H(\tilde{S}) - H(S) = H\left(S;\tilde{S}\right) \geq 0$.

□

Аналогичный результат имеет место для условной квантовой энтропии.

Лемма 11.4.5 *Гауссовское состояние имеет максимальную условную энтропию среди всех состояний составной бозонной системы AB с данными средним m и матрицей ковариаций α.*

Доказательство. Пусть $S_{AB} \in \Sigma\left(m_{AB},\alpha_{AB}\right)$ и пусть $\tilde{S}_{AB}$ – единственное гауссовское состояние в $\Sigma\left(m_{AB},\alpha_{AB}\right)$. Тогда $S_A \in \Sigma\left(m_A,\alpha_A\right)$ и $\tilde{S}_A$ – гауссовское состояние в $\Sigma\left(m_A,\alpha_A\right)$. Мы вновь можем считать, что состояние $\tilde{S}_{AB}$ невырождено. Обозначая

$$H(A|B) = H(S_{AB}) - H(S_B); \quad H(\tilde{A}|\tilde{B}) = H(\tilde{S}_{AB}) - H(\tilde{S}_B),$$

имеем

$$\begin{aligned} H(\tilde{A}|\tilde{B}) - H(A|B) &= H(S_{AB};\tilde{S}_{AB}) - H(S_A;\tilde{S}_A) \\ &\quad + \mathrm{Tr}\left(S_{AB}-\tilde{S}_{AB}\right)\log\tilde{S}_{AB} - \mathrm{Tr}\left(S_A-\tilde{S}_A\right)\log\tilde{S}_A \\ &= H(S_{AB};\tilde{S}_{AB}) - H(S_A;\tilde{S}_A), \end{aligned}$$

поскольку $\log\tilde{S}_{AB}, \log\tilde{S}_A$ – многочлены второго порядка от канонических переменных. В силу монотонности относительной энтропии, $H(S_{AB};\tilde{S}_{AB}) - H(S_A;\tilde{S}_A) \geq 0$. □

11.4.5 Очищение гауссовского состояния

Рассмотрим две бозонные системы, описываемые ККС с симплектическими пространствами $(Z_1,\Delta_1),(Z_2,\Delta_2)$. Симплектическим пространством составной системы является прямая сумма $Z_{12} = Z_1 \oplus Z_2$, состоящая из

пар (z_1, z_2) с компонентами $z_i \in Z_i$, причем $\Delta_{12} = \Delta_1 + \Delta_2$. По определению, симплектическая матрица Δ_{12} является блочно-диагональной по отношению к разложению $Z_{12} = Z_1 \oplus Z_2$. Операторы Вейля для составной системы определяются как $W_{12}(z_1, z_2) = W_1(z_1) \otimes W_2(z_2)$.

Пусть S_{12} – гауссовское состояние составной системы с матрицей ковариаций α_{12}. Сужение гауссовского состояния S_{12} на первую подсистему определяется математическим ожиданием операторов Вейля $W_1(z_1) \otimes I_2 = W_{12}(z_1, 0)$, а значит, в соответствии с (11.56), имеет матрицу ковариаций α_1, которая является первым диагональным блоком в блочно-матричном разложении

$$\alpha_{12} = \begin{bmatrix} \alpha_1 & \beta \\ \beta^\top & \alpha_2 \end{bmatrix} \quad ; \quad \Delta_{12} = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}. \tag{11.74}$$

Корреляционная матрица α_{12} состояния S_{12} составной системы блочно-диагональна тогда и только тогда, когда это состояние является состоянием-произведением.

Пусть S_1 – гауссовское состояние с матрицей ковариаций $\alpha_1 = \alpha$ на симплектическом пространстве (Z_1 с формой $\Delta_1 = \Delta$. Рассмотрим $Z_2 = Z_1$ с формой $\Delta_2 = -\Delta$, так что

$$\Delta_{12} = \begin{bmatrix} \Delta & 0 \\ 0 & -\Delta \end{bmatrix}. \tag{11.75}$$

Задача 11.4.7 *Рассмотрим симплектическое пространство $Z_{12} = Z_1 \oplus Z_2$ и гауссовское состояние S_{12} с матрицей ковариаций*

$$\alpha_{12} = \begin{bmatrix} \alpha & \Delta\sqrt{-(\Delta^{-1}\alpha)^2 - I/4} \\ -\Delta\sqrt{-(\Delta^{-1}\alpha)^2 - I/4} & \alpha \end{bmatrix}. \tag{11.76}$$

Тогда S_{12} является очищением состояния S_1.

Указание: Проверить, что матрица

$$\Delta_{12}^{-1}\alpha_{12} = \begin{bmatrix} \Delta^{-1}\alpha & \sqrt{-(\Delta^{-1}\alpha)^2 - I/4} \\ \sqrt{-(\Delta^{-1}\alpha)^2 - I/4} & -\Delta^{-1}\alpha \end{bmatrix} \tag{11.77}$$

удовлетворяет условию (11.65) для чистого гауссовского состояния.

В случае элементарных одномодовых гауссовских состояний (11.16) эта конструкция дает

$$\Delta_{12} = \hbar \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix};$$

$$\alpha_{12} = \begin{bmatrix} N+1/2 & 0 & \sqrt{N^2+N} & 0 \\ 0 & N+1/2 & 0 & -\sqrt{N^2+N} \\ \sqrt{N^2+N} & 0 & N+1/2 & 0 \\ 0 & -\sqrt{N^2+N} & 0 & N+1/2 \end{bmatrix}$$

и

$$\Delta_{12}^{-1}\alpha_{12} = \begin{bmatrix} 0 & -(N+1/2) & 0 & \sqrt{N^2+N} \\ N+1/2 & 0 & \sqrt{N^2+N} & 0 \\ 0 & -\sqrt{N^2+N} & 0 & N+1/2 \\ -\sqrt{N^2+N} & 0 & -(N+1/2) & 0 \end{bmatrix}. \quad (11.78)$$

11.5 Гауссовские каналы

11.5.1 Классически-квантовый гауссовский канал

Как отмечалось в конце раздела 11.1, состояние S_μ, данное соотношением (11.18), описывает в квантовой оптике когерентный сигнал лазера в одномодовом равновесном электромагнитном поле. Отображение $\mu \to S_\mu$ можно рассматривать как классически-квантовый (c-q) канал, осуществляющий передачу классического сигнала μ на фоне аддитивного квантового гауссовского шума. Вводя вещественную параметризацию $m = [m_q, m_p]$ вместо комплексной $\mu = \frac{1}{\sqrt{2\hbar}}(m_q + im_p)$, получаем канал $m \to S_m$ с непрерывным алфавитом $\mathcal{X} = \mathbf{R}^2$, для которого

$$\mathrm{Tr} S_m q = m_q, \quad \mathrm{Tr} S_m p = m_p, \quad \mathrm{Tr} S_m a^\dagger a = N + \frac{1}{2\hbar}\|m\|^2, \quad (11.79)$$

где $\|m\|^2 = m_q^2 + m_p^2$. Для нахождения его пропускной способности используем общий подход раздела 10.4.

Предполагая, что слова $w = (m_1, \ldots, m_n)$ передаются посредством независимой передачи букв, введем аддитивное энергетическое ограничение типа (10.15):

$$\frac{1}{2\hbar}\left(\|m_1\|^2 + \cdots + \|m_n\|^2\right) \le nE, \quad (11.80)$$

которое соответствует выбору функции ограничения $f(m) = \frac{1}{2\hbar}\|m\|^2$. Тогда условия теоремы 10.4.3 выполнены при выборе оператора $F = a^\dagger a$, и классическая пропускная способность канала $m \to S_m$ с энергетическим ограничением (11.80) на входе равна

$$C = \max_{\pi \in \mathcal{P}_E} \left[H(\bar{S}_\pi) - \int H(S_m)\, \pi(d^2 m) \right], \quad (11.81)$$

где $\mathcal{P}_E$ определяется как множество входных распределений $\pi(d^2 m)$, удовлетворяющих условию

$$\frac{1}{2\hbar}\int \|m\|^2\,\pi(d^2m) \le E, \tag{11.82}$$

а

$$\bar{S}_\pi = \int S_m\,\pi(d^2m).$$

Заметим, что состояния S_m унитарно эквивалентны в силу (11.20) и поэтому имеют одинаковую энтропию $H(S_m) = H(S_0) = g(N)$. Следовательно, для любого входного распределения $\pi(d^2m)$

$$\chi(\pi) = H(\bar{S}_\pi) - H(S_0),$$

и задача сводится к максимизации энтропии $H(\bar{S}_\pi)$ при ограничении (11.82). В силу (11.79) из (11.82) следует, что

$$\operatorname{Tr}\bar{S}_\pi\, a^\dagger a \le N + E. \tag{11.83}$$

Это неравенство является ограничением на второй момент состояния $\bar{S}_\pi$; в соответствии с леммой 11.4.4, максимум энтропии, равный

$$H(\bar{S}_\pi) = g(N+E), \tag{11.84}$$

достигается на гауссовском операторе плотности

$$\bar{S}_\pi = \frac{1}{N+E+1}\sum_{n=0}^{\infty}\left(\frac{N+E}{N+E+1}\right)^n |n\rangle\langle n|, \tag{11.85}$$

для которого имеет место равенство в (11.83). Такое состояние соответствует оптимальному распределению

$$\pi(d^2m) = \frac{1}{2\pi\hbar E}\exp\left(-\frac{\|m\|^2}{2\hbar E}\right)d^2m. \tag{11.86}$$

Окончательно, пропускная способность гауссовского c-q канала дается выражением

$$C = C_\chi = g(N+E) - g(N) \tag{11.87}$$
$$= \log\left(1+\frac{E}{N+1}\right) + (N+E)\log\left(1+\frac{1}{N+E}\right) - N\log\left(1+\frac{1}{N}\right),$$

которое при $N \to \infty, E/N \to \text{const}$ асимптотически ведет себя как $\log\left(1+\frac{E}{N+1}\right)$. Таким образом, соотношение (11.87) можно рассматривать как квантовое обобщение знаменитой формулы Шеннона (4.48) с гауссовским шумом мощности N. Отсутствие множителя $1/2$ в формуле (11.87) объясняется наличием двух независимых вещественных амплитуд m_q, m_p для одной квантовой моды.

11.5.2 Открытые бозонные системы

В этом разделе рассматриваются каналы, которые естественно порождаются взаимодействием бозонной системы с бозонным окружением (шумом). Пусть Z_A, Z_B – симплектические пространства, описывающие вход и выход канала, а Z_D, Z_E – соответствующие окружения, так что

$$Z_A \oplus Z_D = Z_B \oplus Z_E = Z, \tag{11.88}$$

и пусть $W_A(z_A)\ldots$ – операторы Вейля в гильбертовых пространствах $\mathcal{H}_A\ldots$ соответствующих бозонных систем. Предположим, что составная система, находящаяся в исходном состоянии $S_A \otimes S_D$, эволюционирует согласно каноническому унитарному преобразованию U_T, которое отвечает симплектическому преобразованию T в Z. В соответствии с разложениями (11.88) в прямые суммы, T можно записать в блочно-матричной форме

$$T = \begin{bmatrix} K & L \\ K_D & L_D \end{bmatrix}, \tag{11.89}$$

где $K : Z_B \to Z_A, L : Z_E \to Z_A, K_D : Z_B \to Z_D, L_D : Z_E \to Z_D$.

Характеристическая функция состояния системы после взаимодействия, описываемого унитарным преобразованием U_T, имеет вид

$$\begin{aligned}
\varphi_B(z_B) &= \operatorname{Tr} U_T \left(S_A \otimes S_D\right) U_T^* \left(W_B(z_B) \otimes I_E\right) \\
&= \operatorname{Tr} \left(S_A \otimes S_D\right) U_T^* \left(W_B(z_B) \otimes W_E(0)\right) U_T \\
&= \operatorname{Tr} \left(S_A \otimes S_D\right) \left(W_A(K z_B) \otimes W_D(K_D z_B)\right) \\
&= \varphi_A(K z_B) \varphi_D \left(K_D z_B\right),
\end{aligned}$$

где $\varphi_D\left(z_D\right) = \operatorname{Tr} S_D W_D(z_B)$ – характеристическая функция исходного состояния окружения. Это преобразование можно переписать в виде

$$\varphi_B(z_B) = \varphi_A(K z_B) f(z_B), \tag{11.90}$$

где

$$f(z_B) = \varphi_D\left(K_D z_B\right). \tag{11.91}$$

Если исходное состояние окружения является гауссовским с параметрами (m_D, α_D), то f имеет вид гауссовской характеристической функции

$$f(z_B) = \exp\left[i l(z_B) - \frac{1}{2}\alpha(z_B, z_B)\right], \tag{11.92}$$

где $l(z_B) = m_D(K_D z_B)$ и

$$\alpha(z_B, z_B') = \alpha_D(K_D z_B, K_D z_B'). \tag{11.93}$$

Задача 11.5.1 *Преобразование состояний согласно формуле (11.90) задает квантовый канал $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$. Указание: преобразование представляет собой композицию унитарной эволюции и частичного следа, которые являются каналами в смысле определения 10.5.1.*

Определение 11.5.1 *Канал, преобразующий состояния согласно формуле (11.90), называется линейным бозонным каналом. Если, дополнительно, f – гауссовская характеристическая функция (11.92), где l – линейная форма, а α – скалярное произведение на Z_B, то линейный бозонный канал называется гауссовским с параметрами (K, l, α).*

Пусть $R_A, \ldots$ – векторы канонических наблюдаемых в $\mathcal{H}_A, \ldots$ с коммутационными матрицами $\Delta_A, \ldots$. Согласно (11.39), действие оператора U_T описывается линейным преобразованием

$$[R'_B\, R'_E] \equiv U_T^*[R_B\, R_E]U_T = [R_A\, R_D]T, \tag{11.94}$$

где T – блочная матрица (11.89), и для упрощения обозначений мы пишем $R_A, \ldots$ вместо $R_A \otimes I_D, \ldots$. В частности,

$$R'_B = R_A K + R_D K_D, . \tag{11.95}$$

Коммутационная матрица и матрица ковариаций канонических наблюдаемых R_B на выходе канала даются выражением

$$\alpha_B - \frac{i}{2}\Delta_B = \mathrm{Tr}(R'_B)^\top S R'_B,$$

где $S = S_A \otimes S_D$, и S_A, S_D – операторы плотности в $\mathcal{H}_A$ и $\mathcal{H}_D$ с матрицами ковариаций α_A и α_D, соответственно. Используя (11.95), получаем

$$\Delta_B = K^\top \Delta_A K + K_D^\top \Delta_D K_D \tag{11.96}$$

$$\alpha_B = K^\top \alpha_A K + K_D^\top \alpha_D K_D. \tag{11.97}$$

Заметим, что в гауссовском случае $\alpha = K_D^\top \alpha_D K_D$ является матрицей соответствующей формы, входящей в выражение для функции (11.92), определяющей канал. Поскольку $\alpha_D \geq \frac{i}{2}\Delta_D$, из (11.96) вытекает неравенство

$$\alpha \geq \frac{i}{2}\left[\Delta_B - K^\top \Delta_A K\right], \tag{11.98}$$

которое в дальнейшем будет играть ключевую роль.

Конечное состояние окружения определяется аналогичным уравнением для канонических наблюдаемых R_E после взаимодействия:

$$R'_E = R_A L + R_D L_D.$$

Слабо комплементарный канал $\Phi_E : S_A \to S_E$ (см. раздел 6.6) определяется соотношением

$$\varphi_E(z_E) = \varphi_A\left(L z_E\right) \cdot \varphi_D\left(L_D z_E\right),$$

и, следовательно, также является линейным бозонным каналом, причем гауссовским, если S_D гауссовское. Если исходное состояние окружения чистое, $S_D = |\psi_D\rangle\langle\psi_D|$, то канал Φ_E является комплементарным к каналу Φ, причем изометрия Стайнспринга дается формулой $V = U_T|\psi_D\rangle$.

Теорема 11.5.1 *Неравенство (11.98) является необходимым и достаточным условием для того, чтобы набор (K, l, α) определял гауссовский канал.*

Доказательство. Необходимость условия (11.98) была установлена выше. Покажем, что при выполнении этого условия можно построить симплектическое преобразование T, определяющее динамику открытой бозонной системы. Прежде всего, заметим, что для данных операторов K, K_D, удовлетворяющих соотношению (11.96), всегда возможно расширить преобразование (11.95) до симплектического преобразования T. Соотношение (11.96) означает, что преобразование $z_B \to \begin{bmatrix} Kz_B \\ K_D z_B \end{bmatrix}$ является симплектическим вложением Z_B в $Z_A \oplus Z_D = Z$, т. е. столбцы матрицы $\begin{bmatrix} K \\ K_D \end{bmatrix}$ образуют систему, которая может быть расширена до симплектического базиса в Z, формирующего матрицу T. Таким образом, проблема сводится к следующей: для данных K, α, удовлетворяющих неравенству (11.98), найти симплектическое пространство (Z_D, Δ_D), оператор $K_D : Z_B \to Z_D$ и скалярное произведение в Z_D, задаваемое симметричной матрицей $\alpha_D \geq \frac{i}{2}\Delta_D$, так что

$$K_D^\top \Delta_D K_D = \Delta_B - K^\top \Delta_A K \equiv \tilde{\Delta}_D; \tag{11.99}$$

$$K_D^\top \alpha_D K_D = \alpha. \tag{11.100}$$

Сначала предположим, что α невырождена. Рассмотрим евклидово пространство (Z_B, α) с кососимметричной формой $\tilde{\Delta}_D$. Пусть $2n = \dim Z_B$. По условию, $\alpha \geq \frac{i}{2}\tilde{\Delta}_D$. Рассмотрим кососимметричный оператор S, определяемый соотношением

$$\tilde{\Delta}_D(z_B, z'_B) = \alpha_D(z_B, Sz'_B); \quad z_B, z'_B \in Z_B.$$

Согласно теореме из линейной алгебры, найдется ортонормированный базис $\tilde{e}_j; j = 1, \ldots, 2n-l; \tilde{h}_j; j = 1, \ldots, l$ в (Z_B, α), такой что

$$S\tilde{e}_j = s_j \tilde{h}_j, S\tilde{h}_j = -s_j \tilde{e}_j; \quad j = 1, \ldots, l; \quad S\tilde{e}_j = 0; \quad j = l+1, \ldots, 2n-l.$$

Условие $\alpha \geq \frac{i}{2}\tilde{\Delta}_D$ влечет $I - \frac{i}{2}S \geq 0$, поэтому $0 < s_j \leq 2$.

Пусть (Z_D, Δ_D) – стандартное симплектическое пространство размерности $2l+2(2n-2l) = 2(2n-l)$ с базисом $e_j, h_j; j = 1, \ldots, 2n-l$. Определим α_D как форму с диагональной матрицей

$$\alpha_D e_j = s_j^{-1} e_j, \quad \alpha_D h_j = s_j^{-1} h_j; \quad j = 1, \ldots, l;$$

$$\alpha_D e_j = e_j, \quad \alpha_D h_j = \frac{1}{4} h_j; \quad j = l+1, \ldots, 2n-l,$$

тогда $\alpha_D \geq \frac{i}{2}\Delta_D$. Определим оператор $K_D : Z_B \to Z_D$ формулой

$$K_D\tilde{e}_j = \sqrt{s_j}e_j, \quad K_D\tilde{h}_j = \sqrt{s_j}h_j; \quad j = 1, \ldots, l;$$
$$K_D\tilde{e}_j = e_j, \qquad j = l+1, \ldots, 2n-l. \tag{11.101}$$

Тогда соотношения (11.99),(11.100) вытекают из рассмотрения значений соответствующих квадратичных форм на базисных векторах $\tilde{e}_j; j = 1, \ldots, 2n-l; \tilde{h}_j; j = 1, \ldots, l$ в Z_B. Отметим, что по построению, α_D является матрицей ковариаций чистого состояния тогда и только тогда, когда $s_j = 2, j = 1, \ldots, l$.

Если же α обращается в нуль на нетривиальном подпространстве $Z_0 \subset Z_B$, то $\tilde{\Delta}_D$ также равна нулю на Z_0 в силу условия $\alpha \geq \frac{i}{2}\tilde{\Delta}_D$. Поэтому векторы $\tilde{e}_j; \quad j = l+1, \ldots, 2n-l$ можно выбрать таким образом, что часть из них будет образовывать базис в Z_0. Тогда можно модифицировать определение (11.101), потребовав $K_D\tilde{e}_j = 0$ для $\tilde{e}_j \in Z_0$, при этом будут выполнены соотношения (11.99), (11.100). □

11.5.3 Основные свойства гауссовских каналов

Формула (11.90) равносильна следующему соотношению для двойственного канала

$$\Phi^*[W_B(z_B)] = W(Kz_B)f(z_B), \tag{11.102}$$

которое показывает, что операторы Вейля с точностью до множителя переходят в операторы Вейля. В случае гауссовского канала

$$\Phi^*[W_B(z_B)] = W(K_Bz_B)\exp\left[il(z_B) - \frac{1}{2}\alpha(z_B, z_B)\right]. \tag{11.103}$$

Следующее утверждение показывает, что гауссовские каналы могут быть определены соотношениями (11.90), (11.103) безотносительно к картине взаимодействия с окружением, рассмотренной в предыдущем разделе.

Предложение 11.5.2 *Для того, чтобы отображение (11.103) было вполне положительно, необходимо и достаточно выполнения условия (11.98).*

Доказательство. Достаточность была установлена в теореме 11.5.1. Для доказательства необходимости заметим, что полная положительность отображения (11.102) влечет неотрицательную определенность матриц с операторными элементами

$$W(Kz_s)\Phi^*[W(z_s)^*W(z_r)]W(Kz_r)^*$$
$$= f(z_r - z_s)\,\exp\frac{i}{2}\big[\Delta(z_r, z_s) - \Delta(K^\top z_r, K^\top z_s)\big], \tag{11.104}$$

где $z_1, \ldots, z_n$ – произвольные конечные наборы элементов из Z. В гауссовском случае (11.103) это равносильно неотрицательной определенности эрмитовых матриц с элементами

$$\exp\{\alpha(z_r, z_s) - \frac{i}{2}\Delta(z_r, z_s) + \frac{i}{2}\Delta(Kz_r, Kz_s)\}, \tag{11.105}$$

и условие (11.98) вытекает из леммы 11.4.2. □

Отметим следующие свойства:

i. Гауссовский канал преобразует гауссовские состояния в гауссовские состояния.

ii. Двойственный к линейному бозонному каналу преобразует любой многочлен от канонических наблюдаемых R_B в многочлен от R_A того же порядка, при условии, что функция f имеет производные достаточно высокого порядка. Это получается дифференцированием соотношения (11.102) в точке $z_B = 0$.

iii. Композиция гауссовских каналов является гауссовским каналом. Действительно, пусть $\Phi_j; j = 1, 2,$ – два гауссовских канала с параметрами K_j, l_j, α_j, тогда, используя определение (11.103), получаем гауссовский канал $\Phi_2 \circ \Phi_1$ с параметрами

$$\begin{aligned} K &= K_1 K_2, \\ l &= K_2^\top l_1 + l_2, \\ \alpha &= K_2^\top \alpha_1 K_2 + \alpha_2. \end{aligned} \tag{11.106}$$

iv. Линейный бозонный, в частности, гауссовский канал ковариантен в следующем смысле

$$\Phi[W(z)^* S W(z)] = W(K'z)^* \Phi[S] W(K'z), \tag{11.107}$$

где $K' = \Delta^{-1} K^\top \Delta$ – симплектически сопряженное преобразование.

11.5.4 Гауссовские наблюдаемые

Пусть заданы два симплектических пространства Z_A, Z_B с соответствующими системами Вейля в гильбертовых пространствах $\mathcal{H}_A, \mathcal{H}_B$. Пусть M наблюдаемая в $\mathcal{H}_A$ с множеством исходов Z_B, которая задается вероятностной операторно-значной мерой $M(d^{2n}z)$. Далее в этом разделе будем опускать индекс $_B$, так что $z = z_B$ и т. д. Наблюдаемая полностью определяется *операторной характеристической функцией*

$$\phi_M(w) = \int \mathrm{e}^{i\Delta(w,z)} M(d^{2n}z).$$

Отметим следующее очевидное свойство характеристической функции наблюдаемой: для любого конечного подмножества $\{w_j\} \subset Z_B$ блочная матрица с операторными элементами $\phi(w_j - w_k)$ является неотрицательно определенной.

Наблюдаемая M называется *гауссовской*, если ее операторная характеристическая функция имеет вид

$$\phi_M(w) = W_A(Kw) \exp\left(-\frac{1}{2}\alpha(w,w)\right) = \exp\left(iR_A^\top Kw - \frac{1}{2}\alpha(w,w)\right), \tag{11.108}$$

где $K : Z_B \to Z_A$ – линейный оператор, а α – билинейная форма на Z_B.

Теорема 11.5.3 *Для того, чтобы соотношение (11.108) определяло наблюдаемую, необходимо и достаточно выполнения матричного неравенства*

$$\alpha \geq \frac{i}{2} K^\top \Delta_A K. \tag{11.109}$$

Доказательство. Для операторной функции, заданной соотношением (11.108),

$$\begin{aligned}\phi(w_j - w_k) &= W(Kw_k)^* W(Kw_j) \exp\left[-\frac{i}{2}\Delta(Kw_j, Kw_k) - \frac{1}{2}\alpha(w_j - w_k, w_j - w_k)\right] \\ &= C_k^* C_j \exp\left[\alpha(w_j, w_k) - \frac{i}{2}\Delta(Kw_j, Kw_k)\right],\end{aligned}$$

где $C_j = W(Kw_j) \exp\left[-\frac{1}{2}\alpha(w_j, w_j)\right]$, и неотрицательная определенность матриц со скалярными коэффициентами $\exp\left[\alpha(w_j, w_k) - \frac{i}{2}\Delta(Kw_j, Kw_k)\right]$ влечет неравенство (11.109) согласно лемме 11.4.2.

Достаточность условия (11.109) будет доказана прямым построением расширения Наймарка наблюдаемой M в духе следствия 3.1.1.

Предложение 11.5.4 *Предположим, что выполняется условие (11.109), тогда найдутся бозонная система в пространстве $\mathcal{H}_C$ с каноническими наблюдаемыми R_C, для которой $\mathcal{H}_B \subseteq \mathcal{H}_A \otimes \mathcal{H}_C$ и гауссовское состояние $S_C \in \mathfrak{S}(\mathcal{H}_C)$, такие что*

$$M(U) = \mathrm{Tr}_C \left(I_A \otimes S_C\right) E_{AC}(S), \quad U \subseteq Z_B, \tag{11.110}$$

где E_{AC} – четкая наблюдаемая в пространстве $\mathcal{H}_A \otimes \mathcal{H}_C$, задаваемая спектральной мерой коммутирующих самосопряженных операторов

$$X_B = \Delta_B^{-1} \left(R_A K + R_C K_C\right)^\top, \tag{11.111}$$

где $K_C : Z_B \to Z_C$ – оператор, такой что

$$K_C^\top \Delta_C K_C = -K^\top \Delta_A K. \tag{11.112}$$

Доказательство. Условие (11.112) означает, что $K_C^\top \Delta_C K_C + K^\top \Delta_A K = 0$, то есть коммутативность операторов (11.111). Модифицируя доказательство предложения 11.5.1, построим симплектическое пространство (Z_C, Δ_C), оператор $K_C : Z_B \to Z_C$ и скалярное произведение в Z_C, задаваемое симметричной матрицей $\alpha_C \geq \frac{i}{2}\Delta_C$, так что выполняется (11.112) вместе с

$$K_C^\top \alpha_C K_C = \alpha.$$

Тогда характеристическая функция наблюдаемой E_{AC} равна

$$\begin{aligned}\phi_{E_{AC}}(w_B) &= \int \exp[i\Delta(w_B, z_B)] E_{AC}\left(d^{2n} z_B\right) \\ &= \exp i w_B^\top \Delta_B X_B = \exp i \left(R_A K + R_C K_C\right) w_B \\ &= W_A(K w_B) W_C(K_C w_B),\end{aligned}$$

откуда следует

$$\begin{aligned}\mathrm{Tr}_C \left(I_A \otimes S_C\right) \phi_{E_{AC}}(w_B) &= W_A(K w_B) \exp\left(-\frac{1}{2}\alpha_C(K_C w_B, K_C w_B)\right) \\ &= W_A(K w_B) \exp\left(-\frac{1}{2}\alpha(w_B, w_B)\right) = \phi_M(w_B),\end{aligned}$$

и соотношение (11.110). □

Наблюдаемая M является четкой тогда и только тогда, когда $\alpha = 0$, при этом она является спектральной мерой коммутирующих самосопряженных операторов $R_A K$.

11.5.5 Гауссовские каналы, разрушающие сцепленность

Теорема 11.5.5 *Гауссовский канал Φ с параметрами $(K, 0, \alpha)$ является каналом, разрушающим сцепленность тогда и только тогда, когда α допускает разложение*

$$\alpha = \alpha_A + \alpha_B, \qquad \alpha_A \geq \frac{i}{2} K^\top \Delta_A K, \quad \alpha_B \geq \frac{i}{2}\Delta_B. \tag{11.113}$$

В этом случае Φ имеет представление

$$\Phi[S] = \int_{Z_B} W_B(z) \sigma_B W_B(z)^* \mu_S(d^{2n} z), \tag{11.114}$$

где σ_B – гауссовское состояние с параметрами $(0, \alpha_B)$, а $\mu_S(U) = \mathrm{Tr} S M_A(U), U \subseteq Z_B$, – распределение вероятностей гауссовской наблюдаемой M_A с характеристической функцией

$$\phi_{M_A}(w) = W_A(K w) \exp\left(-\frac{1}{2}\alpha_A(w, w)\right). \tag{11.115}$$

Доказательство. Предположим сначала, что α допускает указанное разложение и рассмотрим канал, определяемый соотношением (11.114); надо показать, что

$$\Phi^*[W_B(w)] = W_A(Kw)\exp\left[-\frac{1}{2}\alpha(w,w)\right]. \tag{11.116}$$

В самом деле, для произвольного состояния S

$$\begin{aligned}
\mathrm{Tr}S\Phi^*[W_B(w)] &= \mathrm{Tr}\Phi[S]W_B(w) = \int_{Z_B}\mathrm{Tr}W_B(z)\sigma_B W_B(z)^* W_B(w)\mu_S(d^{2n}z)\\
&= \int_{Z_B}\mathrm{Tr}\sigma_B W_B(z)^* W_B(w)W_B(z)\mu_S(d^{2n}z)\\
&= \mathrm{Tr}\sigma_B W_B(w)\int_{Z_B}\exp\left[i\Delta(w,z)\right]\mu_S(d^{2n}z)\\
&= \exp\left[-\frac{1}{2}\alpha_B(w,w)\right]\mathrm{Tr}S\phi_{M_A}(w)\\
&= \mathrm{Tr}SW_A(Kw)\exp\left[-\frac{1}{2}\alpha_B(w,w)-\frac{1}{2}\alpha_A(w,w)\right],
\end{aligned} \tag{11.117}$$

откуда и следует соотношение (11.116).

Обратно, пусть канал Φ гауссовский и разрушает сцепленность. Мы используем гауссовскую версию доказательства теоремы 10.7.2. Фиксируем гауссовское невырожденное состояние S_A в $\mathfrak{S}(\mathcal{H}_A)$ и пусть $\{|e_j\rangle\}_{j=1}^{+\infty}$ – базис собственных векторов оператора S_A с соответствующими (положительными) собственными значениями $\{\lambda_j\}_{j=1}^{+\infty}$. Рассмотрим единичный вектор

$$|\Omega\rangle = \sum_{j=1}^{+\infty}\sqrt{\lambda_j}|e_j\rangle\otimes|e_j\rangle$$

в пространстве $\mathcal{H}_A\otimes\mathcal{H}_A$, тогда $|\Omega\rangle\langle\Omega|$ является гауссовским очищением состояния S_A. Поскольку Φ разрушает сцепленность, гауссовское состояние

$$S_{AB} = (\mathrm{Id}_A\otimes\Phi)\left[|\Omega\rangle\langle\Omega|\right] \tag{11.118}$$

в $\mathfrak{S}(\mathcal{H}_A\otimes\mathcal{H}_B)$ является разделимым. Отсюда вытекает представление

$$S_{AB} = \int\limits_{Z_A}\int\limits_{Z_B} W_A(z_A)\sigma_A W_A(z_A)^*\otimes W_B(z_B)\sigma_B W_B(z_B)^* P(d^{2m}z_A d^{2n}z_B), \tag{11.119}$$

где σ_A, σ_B – гауссовские состояния, а P – гауссовское распределение вероятностей. В самом деле, из представления (10.41) для разделимого состояния следует, что матрица ковариаций α_{AB} состояния S_{AB} удовлетворяет неравенству

$$\alpha_{AB} \geq \begin{bmatrix} \alpha_A & 0 \\ 0 & \alpha_B \end{bmatrix} \equiv \hat{\alpha}_{AB},$$

где α_A, α_B – матрицы ковариаций некоторых квантовых состояний в системах A, B. Обозначая σ_A, σ_B гауссовские состояния с такими матрицами ковариаций и P – гауссовское распределение вероятностей с симплектическим преобразованием Фурье, соответствующим матрице ковариаций $\alpha_{AB} - \hat{\alpha}_{AB}$, получаем (11.119).

Беря частичный след по системе B, получаем

$$\begin{aligned} S_A &= \mathrm{Tr}_B(\mathrm{Id}_A \otimes \Phi)\left[|\Omega\rangle\langle\Omega|\right] \\ &= \int\limits_{Z_A}\int\limits_{Z_B} W_A(z_A)\sigma_A W_A(z_A)^* P(d^{2m}z_A d^{2n}z_B) \\ &= \int\limits_{Z_A}\int\limits_{Z_B} \overline{W_A(z_A)\sigma_A W_A(z_A)^*} P(d^{2m}z_A d^{2n}z_B), \end{aligned}$$

где черта означает комплексное сопряжение в базисе из собственных значений оператора S_A. Тогда аналогично (10.52) получаем, что соотношение

$$\begin{aligned} &M_A(U) \\ &= S_A^{-1/2}\left[\int\limits_{Z_A}\int\limits_{U} \overline{W_A(z_A)\sigma_A W_A(z_A)^*}\, P(d^{2m}z_A d^{2n}z_B)\right] S_A^{-1/2}, \end{aligned}$$

задает ограниченный положительный оператор, такой что $M_A(U) \leq M_A(Z_B) = I_A$. Нетрудно показать, что $M_A(U)$ является вероятностной операторно-значной мерой на борелевских подмножествах $U \subseteq Z_B$.

Покажем, что имеет место представление (11.114) для канала Φ с так определенными M_A и σ_B. Рассмотрим канал, разрушающий сцепленность

$$\hat{\Phi}[S] = \int_{Z_B} W_B(z)\sigma_B W_B(z)^* \mu_S(d^{2n}z).$$

где $\mu_S(U) = \mathrm{Tr}SM_A(U); U \subseteq Z_B$. Чтобы доказать, что $\Phi = \hat{\Phi}$, достаточно показать $\hat{\Phi}[|e_j\rangle\langle e_k|] = \Phi[|e_j\rangle\langle e_k|]$ для всех j, k. Однако

$$\begin{aligned} &\hat{\Phi}[|e_j\rangle\langle e_k|] = \textstyle\int_{Z_B} W_B(z)\sigma_B W_B(z)^* \langle e_k|M_A(d^{2n}z)|e_j\rangle \\ &= \lambda_j^{-1/2}\lambda_k^{-1/2} \int\limits_{Z_A}\int\limits_{Z_B} \langle e_j|W_A(z_A)\sigma_A W_A(z_A)^*|e_k\rangle W_B(z_B)\sigma_B W_B(z_B)^* P(d^{2m}z_A d^{2n}z_B) \\ &= \lambda_j^{-1/2}\lambda_k^{-1/2}\mathrm{Tr}_A(|e_k\rangle\langle e_j| \otimes I_B)S_{AB} = \Phi[|e_j\rangle\langle e_k|], \end{aligned}$$

в силу (11.118), (11.119).

Остается доказать, что M_A – гауссовская наблюдаемая с характеристической функцией (11.115), где $\alpha_A = \alpha - \alpha_B$, а α_B – корреляционная

функция состояния σ_B; не ограничивая общности, мы можем предположить, что среднее равно нулю, так как его всегда можно устранить действием операторов Вейля. Однако из (11.117) следует

$$\Phi^*[W_B(w)] = \exp\left[-\frac{1}{2}\alpha_B(w,w)\right]\phi_{M_A}(w)$$

для любого канала Φ с представлением (11.114), откуда, принимая во внимание (11.116), в самом деле получаем (11.115) с $\alpha_A = \alpha - \alpha_B$. □

Условие теоремы автоматически выполнено в частном случае, когда

$$K^\top \Delta_A K = 0.$$

В этом случае компоненты векторного оператора $R_A K$ коммутируют, поэтому M_A – четкая наблюдаемая, задаваемая их спектральной мерой, и распределение вероятностей $\mu_S(d^{2n}z)$ может быть сколь угодно сконцентрировано в любой точке z при надлежащем выборе состояния S. Таким образом, в этом случае естественно отождествить Φ с классически-квантовым (c-q) каналом, определяемым семейством состояний $z \to W(z)\sigma_B W(z)^*$.

11.6 Пропускные способности гауссовских каналов

11.6.1 Максимизация квантовой взаимной информации

Мы начнем с рассмотрения классической пропускной способности с использованием сцепленного состояния, поскольку она находится наиболее эффективно в гауссовском случае. Если состояние S и канал Φ являются гауссовскими, то величины $H(S)$, $H(\Phi[S]), H(S,\Phi)$ и $I(S,\Phi)$ могут быть вычислены с использованием формул (11.72), (11.97), (11.77). В частности, энтропия $H(\Phi[S])$ дается формулой (11.72), с заменой α на матрицу α', вычисленную по формуле (11.97). Производя очищение входного состояния и используя формулу (7.40), имеем

$$H(S,\Phi) = \frac{1}{2}\mathrm{Sp}\, g\left(\mathrm{abs}(\Delta_{12}^{-1}\alpha'_{12}) - \frac{I}{2}\right),$$

где матрица

$$\alpha'_{12} = \begin{bmatrix} \alpha' & K\beta \\ \beta^\top K^\top & \alpha \end{bmatrix} \tag{11.120}$$

с $\beta = \Delta\sqrt{-(\Delta^{-1}\alpha)^2 - I/4}$, получается подстановкой уравнения (11.95) в соотношение

$$\alpha'_{12} - \frac{i}{2}\Delta'_{12} = \mathrm{Tr}\,[R'\,R_2]^\top\, S\,[R'\,R_2]\,,$$

в котором R_2 обозначает неизменяемые канонические наблюдаемые эталонной системы. С другой стороны, если имеется явное описание комплементарного канала $\tilde{\Phi}$, то обменную энтропию можно найти как выходную энтропию $H(\tilde{\Phi}[S])$.

Следующий результат существенно упрощает вычисление пропускной способности гауссовских каналов для передачи классической информации с помощью сцепленного состояния, сводя вычисления к случаю гауссовских входных состояний.

Теорема 11.6.1 *Пусть Φ – гауссовский канал. Максимум взаимной информации $I(S,\Phi)$ на множестве состояний $\Sigma(m,\alpha)$ с фиксированными первыми и вторыми моментами достигается на гауссовском состоянии.*

Доказательство. Используя представление (7.45) для квантовой взаимной информации, можно записать

$$I(S,\Phi) = H(B|E) + H(B),$$

где B – выход канала и E – окружение. Для гауссовского канала первые и вторые моменты преобразуются одинаково для всех состояний из $\Sigma(m,\alpha)$. Применяя лемму 11.4.4, получаем

$$H(B) = H(\Phi[S]) \leq H(\widetilde{\Phi[S]}) = H(\Phi[\tilde{S}]) = H(\tilde{B}),$$

где $\widetilde{\Phi[S]}$ (соответственно, $\tilde{S}$) – гауссовское состояние с теми же первыми и вторыми моментами, что и $\Phi[S]$ (соответственно, S.) Канал $S \to S_{BE} = VSV^*$, где V – изометрия Стайнспринга для Φ, также является гауссовским, поскольку он осуществляется каноническим унитарным преобразованием с гауссовским состоянием окружения, именно,

$$\operatorname{Tr} S_{BE} W(z_B) \otimes W(z_E) = \phi_A(Kz_B + Lz_E)\phi_D(K_D z_B + L_D z_E). \quad (11.121)$$

Поэтому аналогичное рассуждение, основанное на лемме 11.4.5, показывает, что $H(B|E) \leq H(\tilde{B}|\tilde{E})$. Таким образом, $I(S,\Phi) \leq I(\tilde{S},\Phi)$. □

Из этого результата вытекает, что если максимум информации $I(S,\Phi)$ на некотором множестве состояний, задаваемом ограничениями на первые и вторые моменты, вообще достигается, то он достигается на гауссовском состоянии. Рассмотрим энергетическое ограничение

$$\operatorname{Tr} SF \leq E, \quad (11.122)$$

где $F = R\epsilon R^\top$ – квадратичный оператор с положительной невырожденной матрицей энергии ϵ. Заметим, что

$$\mathrm{Tr}SF = \mathrm{Sp}(\epsilon\alpha_A) + m_A\epsilon m_A^\top,$$

где m_A – вектор-строка средних значений и α_A – матрица ковариаций состояния S. Предложение 10.6.2 дает выражение для классической пропускной способности передачи с использованием сцепленного состояния $C_{ea}(\Phi, F, E)$ в виде максимума квантовой взаимной информация $I(S, \Phi)$ на множестве всех состояний S, удовлетворяющих ограничению (11.122). Покажем, что оператор $F = R\epsilon R^\top$ удовлетворяет условиям этого предложения. Действительно, в силу теоремы 11.4.3, оператор F удовлетворяет условию (10.9). Выбирая $\tilde{F} = c[RR^\top - (\mathrm{Sp}\alpha_E K_E^\top K_E)I]$, имеем $\Phi^*[\tilde{F}] = cRK^\top KR^\top$ и всегда можно подобрать положительное c такое, что $\Phi^*[\tilde{F}] \leq F$. Более того, $\tilde{F}$ удовлетворяет условию (10.9). Поэтому максимум величины $I(S, \Phi)$ достигается и имеет место формула (10.39). Поскольку энергетическое ограничение выражается через первые и вторые моменты m_A, α_A, то значение $C_{ea}(\Phi, F, E)$ достигается на гауссовском состоянии.

Дальнейшее упрощение максимизации получается в калибровочно-инвариантном случае.

11.6.2 Калибровочно-ковариантные каналы

Рассмотрим канал $\Phi : \mathfrak{T}(\mathcal{H}_A) \to \mathfrak{T}(\mathcal{H}_B)$. Предположим, что в Z_A, Z_B фиксированы некоторые операторы комплексной структуры J_A, J_B, и пусть G_A, G_B – операторы, порождающие унитарные группы калибровочных преобразований в $\mathcal{H}_A, \mathcal{H}_B$ согласно формуле (11.49). Канал называется *калибровочно-ковариантным*, если

$$\Phi[\mathrm{e}^{i\varphi G_A} S \mathrm{e}^{-i\varphi G_A}] = \mathrm{e}^{i\varphi G_B}\Phi[S]\mathrm{e}^{-i\varphi G_B} \tag{11.123}$$

для всех входных состояний S и вещественных чисел φ. В случае линейного бозонного канала Φ, в силу (11.49), (11.90), это равносильно следующему

$$\phi(\mathrm{e}^{\varphi J_A} K z_B) f(z_B) = \phi(K\mathrm{e}^{\varphi J_B} z_B) f(\mathrm{e}^{\varphi J_B} z_B),$$

где $\phi(z_A)$ – характеристическая функция состояния S. Для гауссовского канала с параметрами $(K, 0, \alpha)$ это сводится к соотношениям

$$KJ_B - J_A K = 0, \quad [\Delta_B^{-1}\alpha, J] = 0.$$

Таким образом, естественный выбор комплексной структуры в Z_B дается любым оператором J_B, коммутирующим с оператором $\Delta_B^{-1}\alpha$. Существование такой комплексной структуры доказывается аналогично (11.61) с той разницей, что матрица α может быть вырожденной.

В задачах максимизации с ограниченной средней энергией естественная комплексная структура в Z_A определяется оператором энергии $F = R\epsilon R^\top$, а именно, J_A – оператор комплексной структуры в Z_A, коммутирующий с оператором $\epsilon\Delta_A$, так что

$$J_A\epsilon + \epsilon J_A^\top = 0. \tag{11.124}$$

В случае обычного гамильтониана системы осцилляторов $H = \frac{1}{\hbar}F$ действие оператора J_A сводится к умножению на i в соответствующей комплексификации.

Пусть теперь гауссовский канал Φ ковариантен по отношению к указанным естественным комплексным структурам. Тогда

$$I(\mathrm{e}^{i\varphi G_A} S \mathrm{e}^{-i\varphi G_A}, \Phi) \equiv I(S, \Phi),$$

что вытекает из аналогичного свойства всех трех слагаемых, составляющих $I(S,\Phi)$. Для $H(S)$ это просто следствие унитарной инвариантности энтропии, для $H(\Phi[S])$ дополнительно используется ковариантность канала Φ, а для $H(S,\Phi) = H(\tilde{\Phi}[S])$ это следует из ковариантности комплементарного канала $\tilde{\Phi}$ (задача 6.7.2). Определяя усредненное G_A- инвариантное состояние

$$\bar{S} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{i\varphi G_A} S \mathrm{e}^{-i\varphi G_A} d\varphi,$$

имеем

$$\mathrm{Tr}SF = \mathrm{Sp}(\epsilon\alpha_A) + m_A\epsilon m_A^\top \geq \mathrm{Sp}(\epsilon\alpha_A) = \mathrm{Tr}\bar{S}F,$$

где m_A, α_A – первые и вторые моменты состояния S и последнее равенство вытекает из (11.124):

$$\mathrm{Tr}\bar{S}F = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Tr}\mathrm{e}^{i\varphi G_A} S \mathrm{e}^{-i\varphi G_A} \mathrm{Sp}\left(R\epsilon R^\top\right) d\varphi$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Sp}\left(\mathrm{e}^{\varphi J_A}\epsilon \mathrm{e}^{\varphi J_A^T}\alpha_A\right) d\varphi = \mathrm{Sp}(\epsilon\alpha_A).$$

Теперь, используя вогнутость взаимной информации $I(S,\Phi)$, мы заключаем, что $I(S,\Phi) \leq I(\bar{S},\Phi)$. Таким образом, максимум квантовой взаимной информации при энергетическом ограничении $\mathrm{Tr}SF \leq E$ достигается на калибровочно-инвариантном (G_A-инвариантном) состоянии. Первые и вторые моменты такого состояния с необходимостью удовлетворяют соотношениям $m_A = 0, [J_A, \Delta^{-1}\alpha_A] = 0$. Рассматривая гауссовское состояние с этими же первыми и вторыми моментами, в итоге получаем

Следствие 11.6.1 *Пусть Φ – гауссовский канал, калибровочно ковариантный относительно естественных комплексных структур J_A, J_B (где J_A ассоциирована с оператором энергии). Тогда максимум квантовой взаимной информации на множестве состояний с ограниченной средней энергией достигается на калибровочно-инвариантном (G_A-инвариантном) гауссовском состоянии.*

11.6.3 Максимизация когерентной информации

Рассмотрим когерентную информацию

$$I_c\left(S,\Phi\right)=H(\Phi[S])-H\left(S,\Phi\right).$$

Предложение 11.6.2 *Пусть Φ – гауссовский деградируемый канал, так что*

$$\tilde{\Phi}=T\circ\Phi \tag{11.125}$$

где T – также гауссовский канал, тогда квантовая пропускная способность канала Φ равна

$$Q(\Phi)=\sup_{\tilde{S}}I_c\left(\tilde{S},\Phi\right),$$

где супремум берется по множеству гауссовских состояний $\tilde{S}$. Если, дополнительно, все каналы калибровочно-ковариантны относительно некоторых комплексных структур, то супремум достаточно брать только по множеству калибровочно-инвариантных (G_A-инвариантных) гауссовских состояний.

Доказательство. Если канал Φ является деградируемым, т. е. имеет место (11.125), где $\tilde{\Phi}$ – комплементарный канал и T – некоторый канал, то (см. раздел 9.3.3)

$$I_c\left(S,\Phi\right)=H\left(E'|E\right)=H\left(S_{E'E}\right)-H\left(S_E\right), \tag{11.126}$$

где $S_{E'E}=V'S_BV'^*$ и $V':\mathcal{H}_B\to\mathcal{H}_E\otimes\mathcal{H}_{E'}$ – изометрия Стайнспринга для канала T. Если же канал T можно выбрать гауссовским, то канал $S_B\to S_{E'E}$ также гауссовский и значит применимо рассуждение с использованием леммы 11.4.5. Поэтому $I_c\left(S,\Phi\right)=H\left(E'|E\right)\leq H\left(\tilde{E}'|\tilde{E}\right)=I_c\left(\tilde{S},\Phi\right)$. Следовательно, квантовая пропускная способность канала Φ равна

$$Q(\Phi)=\sup_{S}I_c\left(S,\Phi\right)=\sup_{\tilde{S}}I_c\left(\tilde{S},\Phi\right),$$

где супремум берется по множеству гауссовских состояний $\tilde{S}$.

Если, дополнительно, канал Φ калибровочно-ковариантен, то из вогнутости условной квантовой энтропии (11.126) следует, что этот супремум можно брать только по множеству калибровочно-инвариантных гауссовских состояний, аналогично рассуждениям из предыдущего раздела. □

11.6.4 Классическая пропускная способность

Классическую пропускную способность гауссовского канала Φ естественно рассматривать при аддитивном входном ограничении (10.28), соответствующем оператору

$$F^{(n)} = F \otimes I \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes I \otimes F,$$

где $F = R^\top \epsilon R$ – квадратичный оператор энергии с положительно определенной матрицей ϵ. Однако нахождение значения пропускной способности $C(\Phi, F, E)$ зависит от решения до сих пор открытой проблемы аддитивности для гауссовских каналов. Естественной оценкой для $C(\Phi, F, E)$ является величина $C_\chi(\Phi, F, E)$, определяемая выражением (10.31), которая совпадает с $C(\Phi, F, E)$ в случае аддитивности, например, для каналов, разрушающих сцепленность. Во всяком случае, эта величина дает нижнюю границу для $C(\Phi, F, E)$.

Однако даже вычисление величины $C_\chi(\Phi, F, E)$ для гауссовских каналов остается в общем случае открытым вопросом (за исключением c-q каналов и специального случая, который будет рассмотрен в предложении 11.7.2). Отметим, что по крайней мере в рассматриваемой ситуации оптимальный обобщенный ансамбль всегда существует, поскольку набор условий в следствии 10.5.1 выполнен для $F = R^\top \epsilon R$, как было показано в разделе 11.6.1. Таким образом, максимум величины $\chi_\Phi(\pi)$ достигается, и для $C_\chi(\Phi, F, E)$ имеет место формула (10.36).

Выдвигается следующая **гипотеза о гауссовских оптимальных ансамблях**. *Для гауссовского канала Φ с квадратичным энергетическим ограничением максимум величины $\chi_\Phi(\pi)$ в выражении (10.36) для $C_\chi(\Phi, F, E)$ достигается на гауссовском обобщенном ансамбле π, который состоит из обобщенных когерентных состояний $W(z)S_0W(z)^*$, где S_0 – чистое гауссовское состояние, с гауссовским распределением $P(d^{2n}z)$.*

Для ансамбля описанного типа в силу свойства ковариантности (11.107) гауссовских каналов $H(\Phi[W(z)S_0W(z)^*]) = H(\Phi[S_0])$, поэтому

$$\chi_\Phi(\pi) = H(\Phi[\bar{S}_\pi]) - H(\Phi[S_0]), \tag{11.127}$$

что приводит к **гипотезе о гауссовской минимальной энтропии**: *Для гауссовского канала Φ минимум выходной энтропии достигается на (чистом) гауссовском состоянии S_0.*

Рассмотрим калибровочные преобразования в $\mathcal{H}_A$, ассоциированные с оператором комплексной структуры. Определим действие калибровочных преобразований на обобщенные ансамбли состояний соотношением

$$\pi_\varphi(U) = \pi(\{S : \mathrm{e}^{i\varphi G_A} S \mathrm{e}^{-i\varphi G_A} \in U\}), \quad \varphi \in [0, 2\pi],$$

для борелевских подмножеств $U \in \mathfrak{S}(\mathcal{H}_A)$. Обобщенный ансамбль *калибровочно-инвариантен*, если $\pi_\varphi \equiv \pi$. Используя вогнутость функционала $\chi_\Phi(\pi)$

(задача 10.5.2) и рассуждая как в разделе 11.6.1, нетрудно доказать, что если гауссовский канал Φ калибровочно-ковариантен относительно комплексной структуры, ассоциированной с матрицей энергии ϵ, то максимум в соотношении (10.36) достигается на калибровочно-инвариантном обобщенном ансамбле. Отсюда также следует, что среднее состояние $\bar{S}_\pi$ такого оптимального ансамбля является калибровочно-инвариантным.

Однако в том же предположении можно доказать следующее условное утверждение, связывающее две гипотезы, сформулированные выше.

Предложение 11.6.3 *Пусть гауссовский канал Φ калибровочно-ковариантен относительно естественных комплексных структур J_A, J_B. Предположим также, что минимум выходной энтропии достигается на G_A-инвариантном гауссовском состоянии S_0. Тогда выполняется гипотеза о гауссовских оптимальных ансамблях, причем оптимальный ансамбль π может быть выбран так, что выходное состояние $\bar{S}_B = \Phi[\bar{S}_\pi]$ является G_B-инвариантным гауссовским состоянием.*

Доказательство. Выше было отмечено, что оптимальный ансамбль π существует. Обозначим $\bar{S}_B = \Phi[\bar{S}_\pi]$, тогда $\mathrm{Tr}\bar{S}_\pi F \leq E$ и

$$C_\chi(\Phi, F, E) = \chi_\Phi(\pi) = H(\bar{S}_B) - \hat{H}_\Phi(\bar{S}_B)$$

$$\leq H(\bar{S}_B) - \check{H}(\Phi) = H(\bar{S}_B) - H(\Phi[S_0]), \tag{11.128}$$

в силу общего неравенства $\hat{H}_\Phi(\bar{S}_B) \geq \check{H}(\Phi)$ и предположения $H(\Phi[S_0]) = \check{H}(\Phi)$.

Рассмотрим G_A-инвариантное состояние

$$\bar{S}_A = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{i\varphi G_A}\bar{S}_\pi \mathrm{e}^{-i\varphi G_A} d\varphi,$$

тогда $\mathrm{Tr}\bar{S}_A F = \mathrm{Tr}\bar{S}_\pi F$. Пусть теперь $\tilde{S}_A$ – гауссовское состояние с теми же первыми и вторыми моментами, что и $\bar{S}_A$, тогда вновь

$$\mathrm{Tr}\tilde{S}_A F = \mathrm{Tr}\bar{S}_A F = \mathrm{Tr}\bar{S}_\pi F. \tag{11.129}$$

Более того, $\tilde{S}_A$ является G_A-инвариантным гауссовским состоянием, а $\tilde{S}_B = \Phi[\tilde{S}_A]$ – G_B-инвариантным гауссовским состоянием, причем

$$H(\tilde{S}_B) \geq H(\Phi[\bar{S}_A]) \geq H(\bar{S}_B). \tag{11.130}$$

Здесь первое неравенство следует из леммы 11.4.4, а второе – из вогнутости энтропии. Поскольку S_0 чистое G_A-инвариантное гауссовское состояние, то согласно формуле (11.63), имеет место разложение

$$\tilde{S}_A = \int W(z) S_0 W(z)^* P(d^{2s}z).$$

Обозначим $\tilde{\pi}$ ансамбль когерентных состояний $W(z)S_0W(z)^*$ с гауссовским распределением $P(d^{2s}z)$. Тогда

$$\tilde{S}_B = \Phi[\tilde{S}_A] = \int \Phi[W(z)S_0W(z)^*]P(d^{2s}z) = \int W(K'z)\Phi[S_0]W(K'z)^*P(d^{2s}z)$$

в силу свойства ковариантности (11.107) гауссовских каналов, поэтому

$$\hat{H}_\Phi(\tilde{S}_B) \leq H(\Phi[S_0]) = \check{H}(\Phi).$$

В силу равенства (11.129), ансамбль $\tilde{\pi}$ удовлетворяет энергетическому ограничению. Более того,

$$\chi_\Phi(\tilde{\pi}) = H(\tilde{S}_B) - \hat{H}_\Phi(\tilde{S}_B) \geq H(\tilde{S}_B) - H(\Phi[S_0]). \tag{11.131}$$

Сопоставляя неравенства (11.128), (11.128), (11.131), получаем $\chi_\Phi(\tilde{\pi}) \geq \chi_\Phi(\pi) = C_\chi(\Phi, F, E)$, поэтому $\tilde{\pi}$ является оптимальным ансамблем с требуемыми свойствами. □

11.7 Случай одной моды

11.7.1 Классификация гауссовских каналов

В этом разделе нас будет интересовать задача о приведении произвольного одномодового квантового гауссовского канала к простейшему виду путем подходящих канонических унитарных преобразований на входе и выходе канала:

$$\Phi^{*\prime}\left[V(z)\right] = U_{T_1}^*\Phi^*\left[U_{T_2}^*V(z)U_{T_2}\right]U_{T_1}$$

то есть

$$\Phi^{*\prime}\left[V(z)\right] = V(T_1KT_2z)f(T_2z).$$

Здесь мы дадим полное решение этой задачи в случае $s = 1$.

Пусть Z – двумерное симплектическое пространство, т. е. линейное пространство векторов $z = [x, y]^\top$ с симплектической формой

$$\Delta(z, z') = x'y - xy'. \tag{11.132}$$

(Для упрощения обозначений в этом разделе полагаем $\hbar = 1$). Базис e, h в Z является симплектическим, если $\Delta(e, h) = 1$, т. е. если площадь ориентированного параллелограмма со сторонами e, h равна 1. Линейное преобразование T в Z является симплектическим, если оно отображает симплектический базис в симплектический базис.

В соответствии с формулой (11.103) гауссовский канал характеризуется параметрами K, l, α, удовлетворяющими условию

$$\alpha \geq \frac{i}{2}\left[\Delta - K^{\top}\Delta K\right]. \tag{11.133}$$

Используя преобразование смещения (11.32), всегда можно получить $l = 0$, что и будем подразумевать далее. Поэтому

$$\Phi^*[W(z')] = W(Kz')\exp\left[-\frac{1}{2}\alpha(z', z')\right]. \tag{11.134}$$

Теорема 11.7.1 *Пусть e, h некоторый симплектический базис; в зависимости от значения*

$$A)\quad \Delta(Ke, Kh) = 0; \quad B)\quad \Delta(Ke, Kh) = 1;$$

$$C)\quad \Delta(Ke, Kh) = k^2 > 0, k \neq 1; \quad D)\quad \Delta(Ke, Kh) = -k^2 < 0$$

найдутся симплектические преобразования T_1, T_2 такие, что канал $\Phi^{*\prime}$ имеет вид (11.90) с

$$A_1)\quad K\,[x,y]^{\top} = [0,0]^{\top};$$
$$f(z) = \exp\left[-\frac{1}{2}\Big(N_0 + \frac{1}{2}\Big)\left(x^2 + y^2\right)\right]; \quad N_0 \geq 0;$$

$$A_2)\quad K\,[x,y]^{\top} = [x,0]^{\top};$$
$$f(z) = \exp\left[-\frac{1}{2}\Big(N_0 + \frac{1}{2}\Big)\left(x^2 + y^2\right)\right];$$

$$B_1)\quad K\,[x,y]^{\top} = [x,y]^{\top};$$
$$f(z) = \exp\left[-\frac{1}{4}y^2\right];$$

$$B_2)\quad K\,[x,y]^{\top} = [x,y]^{\top};$$
$$f(z) = \exp\left[-\frac{1}{2}\,N_c\,\left(x^2 + y^2\right)\right]; \quad N_c \geq 0;$$

$$C)\quad K\,[x,y]^{\top} = [kx,ky]^{\top}; \quad k > 0, k \neq 1;$$
$$f(z) = \exp\left[-\frac{1}{2}\Big(N_c + \frac{|k^2-1|}{2}\Big)\left(x^2 + y^2\right)\right];$$

$$D)\quad K\,[x,y]^{\top} = [kx,-ky]^{\top}; \quad k > 0;$$
$$f(z) = \exp\left[-\frac{1}{2}\Big(N_c + \frac{(k^2+1)}{2}\Big)\left(x^2 + y^2\right)\right].$$

Очевидно, что в случаях $A), B_2), C)$ канал является калибровочно инвариантным по отношению к естественной комплексной структуре, ассоциированной с умножением на i в комплексной плоскости x, y.

Доказательство. A) $\Delta(Ke, Kh) = 0$. В этом случае $\Delta(Kz, Kz') \equiv 0$, и либо $K = 0$, либо K имеет ранг 1. Тогда неравенство (11.133) есть просто

условие, характеризующее корреляционную матрицу квантового гауссовского состояния. Из леммы 11.2.1 следует, что найдется симплектическое преобразование T_2 такое, что

$$\alpha(T_2 z, T_2 z) = \left(N_0 + \frac{1}{2}\right)\left(x^2 + y^2\right). \tag{11.135}$$

Если $K = 0$, то мы получаем случай A_1).

Пусть теперь K имеет ранг 1, тогда и KT_2 имеет ранг 1, значит найдется вектор e' такой, что $KT_2\,[x,y]^\top = (k_1 x + k_2 y)\,e'$. Тогда найдется симплектическое преобразование T_1 такое, что $T_1 e' = [1,0]$ и следовательно, $T_1 K T_2\,[x,y]^\top = [k_1 x + k_2 y, 0]^\top$. Производя поворот T_2', который оставляет инвариантным α, мы можем преобразовать этот вектор к виду $[k_1' x, 0]^\top$ с $k_1' \neq 0$, и производя затем симплектическое масштабное преобразование T_1', мы приходим к случаю A_2).

B,C) $\Delta(Ke, Kh) = k^2 > 0$. Тогда $T = k^{-1}K$ – симплектическое преобразование и $\Delta(Kz, Kz') = k^2\Delta(z, z')$. Положим $T_1 = (TT_2)^{-1}$, где T_2 будет выбрано позднее, так что $T_1 K T_2 = kI$. Отображение T_2 выбирается следующим образом:

В случае B) $k = 1$, и условие (11.133) означает неотрицательную определенность матрицы α. Тогда имеем следующие подслучаи:

B_2) Если α невырождена, то согласно лемме 11.2.1 найдется симплектическое преобразование T_2 такое, что

$$\alpha(T_2 z, T_2 z) = \ N_c\,\left(x^2 + y^2\right),$$

где $N_c > 0$. Также если $\alpha = 0$, то получаем аналогичную формулу с $N_c = 0$.

B_1) С другой стороны, если форма α вырождена и имеет ранг 1, то $\alpha(z,z) = (k_1 x + k_2 y)^2$ при некоторых k_1, k_2, не равных нулю одновременно, тогда для произвольного $N_c > 0$ найдется симплектическое преобразование T_2, такое что

$$\alpha(T_2 z, T_2 z) = \ N_c\, y^2$$

в частности, можно взять $N_c = \frac{1}{2}$.

В случае C) $k \neq 1$, и условие (11.133) влечет, что $\alpha/|k^2 - 1|$ является корреляционной матрицей квантового гауссовского состояния, поэтому найдется симплектическое преобразование T_2, такое что

$$\alpha(T_2 z, T_2 z) = |k^2-1|\Big(\ N_0 + \frac{1}{2}\Big)\left(x^2 + y^2\right) = \Big(\ N_c + \frac{|k^2-1|}{2}\Big)\left(x^2 + y^2\right),$$

где $N_0 \geq 0, N_c = |k^2 - 1| N_0$.

D) $\Delta(Ke, Kh) = -k^2 < 0$. Тогда $T = k^{-1}K$ – антисимплектическое преобразование $(\Delta(Tz, Tz') = -\Delta(z, z'))$, и $\Delta(Kz, Kz') = -k^2\Delta(z, z')$. Аналогично случаям B,C), получаем

$$\alpha(T_2z,T_2z)=(k^2+1)\big(\ N_0+\frac{1}{2}\big)\ \big(x^2+y^2\big)=\big(\ N_c+\frac{(k^2+1)}{2}\big)\ \big(x^2+y^2\big)\,,$$

где $N_0\geq 0, N_c=(k^2+1)N_0$. Полагая $T_1=\Lambda\,(TT_2)^{-1}$, где $\Lambda\,[x,y]^\top=[x,-y]^\top$, мы получаем первое уравнение в D). Заметим, что T_1 – симплектическое преобразование, поскольку как T, так и Λ антисимплектические. □

Как показано в разделе 11.5.2, гауссовский канал Φ можно расширить до линейной динамики открытой бозонной системы, описываемой каноническими наблюдаемыми q,p и каноническими наблюдаемыми окружения $q_E,p_E,\ldots$ в гауссовском состоянии S_E; более того, эта линейная динамика также дает описание слабо комплементарного канала Φ_E, отображающего исходное состояние системы q,p в конечное состояние окружения $q'_E,p'_E,\ldots$; если состояние S_E чистое, то Φ_E – комплементарный канал $\tilde{\Phi}$, который определяется каналом Φ однозначно с точностью до унитарной эквивалентности. Заметим, что окружение может быть многомодовым, что отражается символом $\ldots$ в обозначении вспомогательных переменных.

Дадим описание в терминах открытой системы в каждом из случаев теоремы предыдущего раздела.

A_1) Это есть полностью деполяризующий канал

$$\begin{aligned} q'&=q_E\\ p'&=p_E \end{aligned}$$

где наблюдаемые q_E,p_E описывают окружение в равновесном состоянии S_E со средним числом квантов N_0. Слабо комплементарным является идеальный канал Id.

A_2) Линейное преобразование канонических переменных моды имеет вид

$$\begin{aligned} q'&=q+q_E\\ p'&=p_E, \end{aligned}\tag{11.136}$$

где мода q_E,p_E снова находится в равновесном состоянии S_E со средним числом квантов N_0. Слабо комплементарным к этому каналу является канал, который задается преобразованием

$$\begin{aligned} q'_E&=q\\ p'_E&=p-p_E, \end{aligned}\tag{11.137}$$

где p_E может рассматриваться как классическая случайная величина с дисперсией $N_0+\frac{1}{2}$.

B_1) Уравнение канала имеет вид (11.137), где p_E имеет дисперсию $\frac{1}{2}$, так что мода q_E, p_E находится в чистом (вакуумном) состоянии, при этом комплементарный канал задается преобразованием (11.136).

B_2) Канал с аддитивным комплексным гауссовским шумом интенсивности N_c

$$q' = q + \xi$$
$$p' = p + \eta.$$

Задача 11.7.1 *Доказать, что расширение Стайнспринга для этого канала дается окружением с двумя бозонными модами $q_j, p_j; j = 1, 2$ в чистом гауссовском состоянии, которое имеет нулевые средние, нулевые ковариации и дисперсии*

$$\mathsf{D}q_1 = \mathsf{D}q_2 = N_c; \quad \mathsf{D}p_1 = \mathsf{D}p_2 = \frac{1}{4N_c},$$

так что $\xi = q_1, \eta = q_2$, причем динамика составной системы описывается уравнениями

$$\begin{aligned} q' &= q + q_1, \\ p' &= p + q_2, \\ q'_1 &= q_1, \\ p'_1 &= p_1 - p - q_2/2, \\ q'_2 &= q_2, \\ p'_2 &= p_2 + q + q_1/2. \end{aligned}$$

C) Аттенюатор/усилитель с коэффициентом k и квантовым шумом со средним числом квантов N_0. В случае аттенюатора ($k < 1$) уравнение канала имеет вид

$$q' = kq + \sqrt{1-k^2}q_E$$
$$p' = kp + \sqrt{1-k^2}p_E,$$

где мода q_E, p_E находится в равновесном состоянии S_E со средним числом квантов N_0. Слабо комплементарный канал задается уравнениями

$$q'_E = \sqrt{1-k^2}q - kq_E \tag{11.138}$$
$$p'_E = \sqrt{1-k^2}p - kp_E, \tag{11.139}$$

и также является аттенюатором (с коэффициентом $k' = \sqrt{1-k^2}$).

В случае усилителя ($k > 1$) имеем

$$q' = kq + \sqrt{k^2-1}q_E$$
$$p' = kp - \sqrt{k^2-1}p_E,$$

со слабо комплементарным каналом

$$q'_E = \sqrt{k^2-1}q + kq_E$$
$$p'_E = -\sqrt{k^2-1}p + kp_E,$$

см. случай D).

D) Уравнение канала имеет вид

$$q' = kq + \sqrt{k^2+1}q_E$$
$$p' = -kp + \sqrt{k^2+1}p_E,$$

т. е. этот канал является слабо комплементарным к усиливающему каналу с коэффициентом $k' = \sqrt{k^2+1}$ и квантовым шумом со средним числом квантов N_0.

11.7.2 Каналы, разрушающие сцепленность

Применим теорему 11.5.5 к случаю одной бозонной моды $A = B$, где

$$\Delta_A(z,z') = \Delta_B(z,z') = \Delta(z,z') = x'y - xy'$$

Как было показано выше, произвольный одномодовый гауссовский канал приводится к одной из нормальных форм.

Нам остается только найти форму $K^\top \Delta_A K$ и проверить разложимость (11.113) в каждом из этих случаев. Мы опираемся на следующий простой факт:

Задача 11.7.2

$$\big(N + \frac{1}{2}\big) I \geq \frac{i}{2}\Delta$$

тогда и только тогда, когда $N \geq 0$.

A) $K^\top \Delta K = 0$, поэтому Φ является c-q каналом (фактически, классическим каналом);

B) $K^\top \Delta K = \Delta$, поэтому необходимое условие разложимости (11.113) влечет $\alpha \geq i\Delta$. Это не может быть выполнено в случае B_1) в силу вырожденности α, следовательно канал не является разрушающим сцепленность. С другой стороны, в случае B_2) условие (11.113) выполняется при $\alpha_B = \alpha_A = \alpha/2$ тогда и только тогда, когда $N_c \geq 1$, и Φ является каналом, разрушающим сцепленность;

C) $K^\top \Delta K = k^2\Delta$. Ясно, что в этом случае условие разложимости выполняется тогда и только тогда, когда $\alpha \geq \frac{i}{2}(1+k^2)\Delta$, что равносильно неравенству $N_c + \frac{|1-k^2|}{2} \geq \frac{(1+k^2)}{2}$ или

$$N_c \geq \min\left(1, k^2\right). \tag{11.140}$$

Это дает условие разрушения сцепленности (которое также формально включает случай B_2));

$D)$ $K^\top \Delta K = -k^2\Delta$. Условие разложимости выполняется тогда и только тогда, когда $\alpha \geq \frac{i}{2}(1+k^2)\Delta$, что выполнено всегда, поэтому канал является разрушающим сцепленность для всех $N_c \geq 0$.

Таким образом, свойство аддитивности (10.32) выполняется для одномодовых гауссовских каналов вида $A), D)$ с произвольными значениями параметров, и $B_2), C)$ с параметрами, удовлетворяющими (11.140). В общем случае каналы, разрушающие сцепленность, имеют нулевую квантовую пропускную способность, $Q(\Phi) = 0$, см. следствие 9.3.1. Однако в разделе 11.7.4 на основе анализа деградируемости будет найдена более широкая область нулевой квантовой пропускной способности.

11.7.3 Аттенюатор/усилитель

Рассмотрим подробно случай C). Введя параметр $N_c = |k^2 - 1|N_0 \geq 0$, можно представить действие канала Φ следующим образом

$$\begin{aligned} \mathrm{Tr}\ \Phi[S]\ W(z) = \mathrm{Tr}SW(kz) \times \\ \times \exp\left[-\frac{1}{2}\big(|k^2-1|/2 + N_c\big)\|z\|^2\right]. \end{aligned} \tag{11.141}$$

Как отмечалось, этот канал является калибровочно инвариантным по отношению к естественной комплексной структуре, ассоциированной с умножением на i в комплексной плоскости x, y. Поэтому далее, опираясь на результаты раздела 11.6, мы ограничимся рассмотрением калибровочно инвариантных состояний.

Пусть входное состояние $S = S(N)$ системы является гауссовским с характеристической функцией

$$\mathrm{Tr}S(N)W(z) = \exp\left[-\frac{1}{2}\big(N + \frac{1}{2}\big)\|z\|^2\right],$$

где параметр N представляет мощность (среднее число квантов) сигнала

$$\mathrm{Tr}\, S(N)a^\dagger a = N, \tag{11.142}$$

где $a = \frac{1}{\sqrt{2}}(q + ip)$. Энтропия состояния $S(N)$ равна

$$H(S(N)) = g(N). \tag{11.143}$$

Из (11.141) следует, что выходное состояние $\Phi[S]$ является гауссовским состоянием $S(N')$, где

$$N' = k^2 N + N_0', \tag{11.144}$$

а

$$N_0' = \max\{0, (k^2-1)\} + N_c$$

– среднее число фотонов на выходе, соответствующем входному вакуумному состоянию $S(0)$. Тогда

$$H(\Phi[S(N)]) = g(N'). \tag{11.145}$$

Найдем обменную энтропию $H(S(N), \Phi)$. Чистое входное состояние S_{12} расширенной системы $\mathcal{H} \otimes \mathcal{H}_0$ характеризуется $2 \times 2-$матрицей ковариаций (11.78). В соответствии с (11.120), действие расширенного канала $\Phi \otimes \mathrm{Id}$ преобразует эту матрицу в

$$\Delta_{12}^{-1}\alpha_{12}' = \begin{bmatrix} 0 & -(N'+1/2) & 0 & k\sqrt{N^2+N} \\ N'+1/2 & 0 & k\sqrt{N^2+N} & 0 \\ 0 & -k\sqrt{N^2+N} & 0 & N+1/2 \\ -k\sqrt{N^2+N} & 0 & -(N+1/2) & 0 \end{bmatrix}$$

Из формулы (11.72) получаем $H(S(N), \Phi) = g(|\lambda_1| - \frac{1}{2}) + g(|\lambda_2| - \frac{1}{2})$, где $\pm\lambda_1, \pm\lambda_2$ – собственные значения матрицы в правой части. Решая характеристическое уравнение, находим

$$\lambda_{1,2} = \frac{i}{2}\left((N'-N) \pm D\right), \tag{11.146}$$

где

$$D = \sqrt{(N+N'+1)^2 - 4k^2N(N+1)}.$$

Следовательно,

$$H(S(N), \Phi) = \tag{11.147}$$
$$= g\left(\frac{D+N'-N-1}{2}\right) + g\left(\frac{D-N'+N-1}{2}\right).$$

Зависимость энтропий $H(\Phi[S(N)]), H(S(N), \Phi)$ от k для случая $N_c = 0$ показана на рис. 11.1. Заметим, что для всех N когерентная информация $H(\Phi[S(N)]) - H(S(N), \Phi)$ оказывается положительной при $k > 1/\sqrt{2}$ и отрицательной при остальных значениях k. Она стремится к $-H(S(N))$ при $k \to 0$, равна $H(S(N))$ при $k = 1$, и быстро убывает к нулю при $k \to \infty$.

Рассмотрим классические пропускные способности канала Φ при аддитивных входных энергетических ограничениях, соответствующих оператору $F = a^\dagger a$. В соответствии с теоремой 11.6.1 и следствием из нее классическая пропускная способность с использованием сцепленного состояния равна

$$C_{ea}(\Phi, F, E) = \max_{S:\mathrm{Tr}Sa^\dagger a \le E} I(S, \Phi)$$

и достигается на калибровочно инвариантном гауссовском состоянии S. Поскольку условие (10.40) очевидно выполнено для данного канала, для

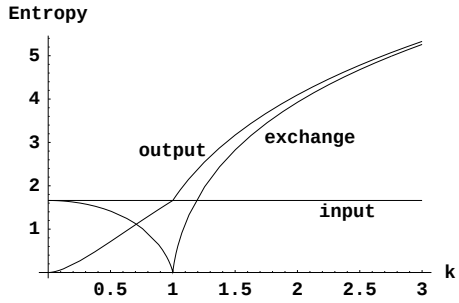


Рис. 11.1. *Энтропии: выходная энтропия (11.145), обменная энтропия (11.147) в случае* $N_c = 0$.

этого состояния $\mathrm{Tr} S a^\dagger a = E$. Поэтому $S = S(E)$, и для энтропий имеют место выражения (11.143), (11.145) и (11.147). Следовательно, указанный выше максимум равен

$$C_{ea}(\Phi, F, E) = I(S(E), \Phi) = H(S(E)) + H(\Phi[S(E)]) - H(S(E), \Phi).$$

Рассматривая величину $C_{ea}(\Phi, F, E)$ как функцию параметров E, k, N_c, интересно сравнить ее с величиной $C_\chi(\Phi, F, E)$, которая является нижней границей для классической пропускной способности $C(\Phi, F, E)$ (возможно, совпадающей с нею). Если верна гипотеза о гауссовских оптимальных ансамблях, то оптимальным является ансамбль, состоящий из когерентных состояний $S_\zeta = |\zeta\rangle\langle\zeta|; \zeta \in \mathbb{C}$, с гауссовской плотностью вероятности $p(\zeta) = (\pi E)^{-1} \exp\left(-|\zeta|^2/E\right)$, что приводит к значению

$$C_\chi(\Phi, F, E) = g\left(k^2 E + N_0'\right) - g\left(N_0'\right).$$

Отношение

$$G = \frac{C_{ea}(\Phi, F, E)}{C_\chi(\Phi, F, E)} \tag{11.148}$$

дает по крайней мере верхнюю границу для *выигрыша* за счет использования сцепленных состояний при передаче классических сообщений. В частности, если среднее число квантов E сигнала стремится к нулю, тогда как $N_0' > 0$, то

$$C_\chi(\Phi, F, E) \sim E k^2 \log\left(\frac{N_0' + 1}{N_0'}\right),$$
$$C_{ea}(\Phi, F, E) \sim -E \log E/(N_0' + 1),$$

и G стремится к бесконечности как $-\log E$. График отношения G как функции k при $N_c = 0$ показан на рис. 11.2.

Гипотеза о гауссовских оптимальных ансамблях верна в частном случае аттенюатора Φ с вакуумным шумом $N_c = 0, k < 1$. Более того, в этом случае выполняется и гипотеза аддитивности, так что величина $C_\chi(\Phi, F, E)$ дает классическую пропускную способность канала.

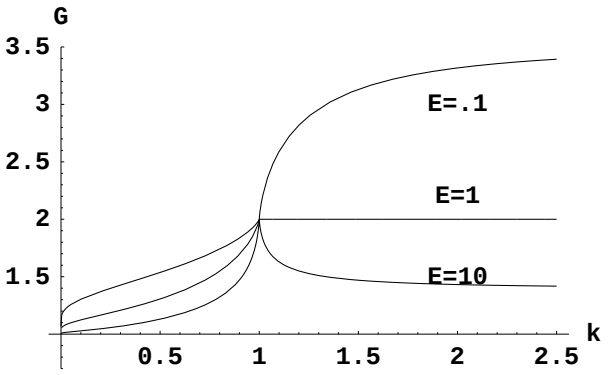


Рис. 11.2. *Выигрыш (11.148) за счет использования сцепленного состояния как функция k при $N_c = 0$. Параметр=входная мощность E.*

Предложение 11.7.2 *Для аттенюатора Φ с вакуумным шумом*

$$C(\Phi, F, E) = C_\chi(\Phi, F, E) = g(k^2 E). \tag{11.149}$$

Доказательство. Докажем свойство аддитивности

$$C_\chi^{(n)}(\Phi, H^{(n)}, nE) = nC_\chi(\Phi, F, E). \tag{11.150}$$

Действительно, для этого канала $N_0' = 0, N' = k^2 E$ и $H(\Phi[S(E)]) = g(k^2 E)$. Следовательно,

$$H(\Phi[S_z]) = 0 \tag{11.151}$$

для произвольного когерентного состояния S_z с $E = 0$ и поэтому минимальная выходная энтропия равна нулю

$$\min_S H(\Phi[S]) = 0.$$

Следовательно, $\min_{S^{(n)}} H(\Phi^{\otimes n}[S^{(n)}]) = 0$ и аддитивность выходной энтропии очевидна. Подставляя ансамбль когерентных состояний и используя (11.151), получаем

$$C_\chi(\Phi, F, E) \geq g(k^2 E) = \max_{S:\mathrm{Tr} S a^\dagger a \leq E} H(\Phi[S]).$$

С другой стороны, используя субаддитивность выходной энтропии, имеем

$$nC_\chi(\Phi, F, E) \leq C_\chi^{(n)}(\Phi, F, E) \leq n \max_{S:\mathrm{Tr} S a^\dagger a \leq E} H(\Phi[S]).$$

Сравнение этих неравенств дает (11.150) и (11.149). □

11.7.4 Квантовая пропускная способность

В этом разделе мы вычислим квантовую пропускную способность ряда одномодовых гауссовских каналов, опираясь на свойства (анти-)деградируемости (разделы 9.3.3, 11.6.3). При этом мы используем тот факт, что композиция одномодовых гауссовских каналов является одномодовым гауссовским каналом. В частности, из общих правил (11.106) вытекают следующие соотношения.

Обозначая $\Phi_{C,k}$ (соответственно $\Phi_{D,k}$) канал класса C (соответственно D) с коэффициентом k и фиксированным значением параметра N_c, имеем

$$\Phi_{C,k_2} \circ \Phi_{C,k_1} = \Phi_{C,k_1k_2}; \quad \text{если } k_1, k_2 < 1, \tag{11.152}$$

$$\Phi_{D,k_2} \circ \Phi_{C,k_1} = \Phi_{D,k_1k_2}; \quad \text{если } k_1 > 1. \tag{11.153}$$

Предложение 11.7.3 *Для аттенюатора с коэффициентом $k \leq \frac{1}{\sqrt{2}}$ квантовая пропускная способность $Q(\Phi_{C,k}) = 0$. Для аттенюатора/усилителя с $k \geq \frac{1}{\sqrt{2}}$ и $N_c = 0$*

$$Q(\Phi_{C,k}) = Q_G(\Phi_{C,k}) = \log \frac{k^2}{|k^2 - 1|}. \tag{11.154}$$

Здесь

$$Q_G(\Phi) = \sup_{N \geq 0} I_c(S(N), \Phi),$$

где супремум когерентной информации $I_c(S, \Phi) = H(\Phi[S]) - H(S, \Phi)$ берется по всем гауссовским калибровочно инвариантным входным состояниям вида $S(N)$.

Доказательство. Достаточно доказать, что аттенюатор с $k < \frac{1}{\sqrt{2}}$ – анти-деградируемый канал, тогда как аттенюатор/усилитель с $k \geq \frac{1}{\sqrt{2}}$ и $N_c = 0$ – деградируем.

Если $k < \frac{1}{\sqrt{2}}$, то $k_1 = \frac{k}{\sqrt{1-k^2}} < 1$. Принимая во внимание, что слабо комплементарным к каналу $\Phi_{C,k}$ является канал

$$\tilde{\Phi}^w_{C,k} = \Phi_{C,\sqrt{1-k^2}}, \tag{11.155}$$

см. (11.138), соотношение (11.152) можно записать в виде

$$\Phi_{C,k} = \Phi_{C,\frac{k}{\sqrt{1-k^2}}} \circ \tilde{\Phi}^w_{C,k}, \tag{11.156}$$

откуда следует, что $\Phi_{C,k}$ – анти-деградируемый канал, и следовательно, в силу теоремы 9.3.5, его квантовая пропускная способность равна нулю.

В случае $N_c = 0$, когда состояние окружения – чистое, комплементарный канал совпадает со слабо комплементарным: $\tilde{\Phi}_{C,k} = \tilde{\Phi}^w_{C,k} = \Phi_{C,k'}$, где $k' = \sqrt{1-k^2} > \frac{1}{\sqrt{2}}$, и соотношение (11.156) приобретает вид

$$\tilde{\Phi}_{C,k'} = \Phi_{C,\frac{k}{\sqrt{1-k^2}}} \circ \Phi_{C,k'},$$

откуда следует, что $\Phi_{C,k'}$ – деградируемый канал при $\frac{1}{\sqrt{2}} < k' < 1$. Поэтому из предложения 11.6.2 следует, что

$$Q(\Phi_{C,k}) = Q_G(\Phi_{C,k}). \tag{11.157}$$

Аналогично в случае усилителя ($k > 1, N_c = 0$) из соотношения (11.153) следует, что

$$\tilde{\Phi}_{C,k} = \Phi_{D,\frac{\sqrt{k^2-1}}{k}} \circ \Phi_{C,k},$$

поэтому $\Phi_{C,k}$ – деградируемый канал, и снова имеет место равенство (11.157). Остается найти величину $Q_G(\Phi_{C,k})$.

В случае $N_c \geq 0$ когерентная информация

$$I_c(S(N), \Phi) = g(N') - g\left(\frac{D + N' - N - 1}{2}\right) - g\left(\frac{D - N' + N - 1}{2}\right) \tag{11.158}$$

сложным образом зависит от входной мощности N. Имеем $I_c(S(0), \Phi) = 0$ и

$$\lim_{N\to\infty} I_c(S(N), \Phi) = \log k^2 - \log|k^2 - 1| - g\left(N_c/|k^2 - 1|\right), \quad k \neq 1.$$

Пусть теперь $N_c = 0$. В случае $k < 1$ находим $N' = k^2 N, D = (1 - k^2)N + 1$ и

$$I_c(S(N), \Phi) = g(k^2 N) - g\left((1 - k^2)N\right).$$

Задача 11.7.3 *Показать, что $I_c(S(N), \Phi)$ – выпуклая убывающая функция параметра N при $k^2 < \frac{1}{2}$ (соответственно вогнутая возрастающая функция при $k^2 > \frac{1}{2}$), следовательно,*

$$\sup_N I_c(S(N), \Phi) = \begin{cases} I_c(S(N), \Phi)|_{N=0} = 0, & k^2 < \frac{1}{2} \\ \lim_{N\to\infty} I_c(S(N), \Phi) = \log \frac{k^2}{1-k^2}, & k^2 > \frac{1}{2}. \end{cases}$$

В случае $k > 1$, используя (11.144), имеем $N' = k^2(N + 1) - 1, D = k^2(N + 1) - N$ и

$$I_c(S(N), \Phi) = g(k^2(N + 1) - 1) - g\left(\left(k^2 - 1\right)(N + 1)\right).$$

Задача 11.7.4 *Показать, что $I_c(S(N), \Phi)$ – вогнутая возрастающая функция N и, следовательно,*

$$\sup_N I_c(S(N), \Phi) = \lim_{N\to\infty} I_c(S(N), \Phi) = \log \frac{k^2}{k^2 - 1}.$$

Окончательно получаем общее выражение (11.154) для квантовой пропускной способности аттенюатора/усилителя при $N_c = 0$. □

Обширная область нулевой квантовой пропускной способности, перекрывающая область (11.140) каналов, разрушающих сцепленность, дается следующим предложением.

Предложение 11.7.4 *Пусть* $k \geq \frac{1}{\sqrt{2}}$, *тогда* $Q(\Phi_{C,k}) = 0$ *при*

$$N_c \geq \frac{1}{2}\left(k^2 - |k^2 - 1|\right) = \min\{k^2, 1\} - \frac{1}{2}. \tag{11.159}$$

Доказательство. Рассмотрим композицию $\Phi = \Phi_2 \circ \Phi_1$, где

$$\Phi_1^*[W(z)] = W(\sqrt{2}kz) \exp\left[-\frac{(2k^2-1)}{2}\left(N_1 + \frac{1}{2}\right)\right],$$
$$\Phi_2^*[W(z)] = W\left(\frac{z}{\sqrt{2}}\right) \exp\left[-\frac{1}{4}\left(N_2 + \frac{1}{2}\right),\right]$$

так что $Q(\Phi_2) = 0$ в силу предложения 11.7.3 и, следовательно, $Q(\Phi_2 \circ \Phi_1) = 0$ в силу (9.45). Тогда

$$\Phi_1^* \circ \Phi_2^*[W(z)] = W(kz) \exp\left[-\frac{1}{2}\left(\frac{|k^2-1|}{2} + N_c\right)\right],$$

где

$$\frac{|k^2-1|}{2} + N_c = \frac{1}{2}\left(N_2 + \frac{1}{2}\right) + \frac{(2k^2-1)}{2}\left(N_1 + \frac{1}{2}\right).$$

Варьируя числа $N_1, N_2 \geq 0$, можно получить все значения N_c, удовлетворяющие (11.159). □

Задача 11.7.5 *Доказать, что в случае A)*

$$\Phi = \Phi \circ \tilde{\Phi},$$

где $\tilde{\Phi}$ *– комплементарный канал, откуда следует, что канал* Φ *антидеградируемый и значит* $Q(\Phi) = 0$.

Задача 11.7.6 *Доказать, что в случае* B_1) $Q(\Phi) = \infty$, *рассматривая когерентную информацию* $I_c(S, \Phi) = H(\Phi[S]) - H(\tilde{\Phi}[S])$, *где* S *– гауссовское состояние с* $\sigma_q^2 = \mathsf{D}Q, \sigma_p^2 = \mathsf{D}P$ *и нулевыми средними и ковариацией. Показать, что*

$$H(\Phi[S]) = g\left(\sqrt{\sigma_Q^2\left(\sigma_P^2 + \frac{1}{2}\right)} - \frac{1}{2}\right),$$
$$H(\tilde{\Phi}[S]) = g\left(\sqrt{\left(\sigma_Q^2 + \frac{1}{2}\right)\frac{1}{2}} - \frac{1}{2}\right).$$

В частности, выбирая состояния с $\sigma_Q^2\sigma_P^2 \to \infty, \sigma_Q^2 \to 0$, *получаем* $I_c(S,\Phi) \to \infty$.

Задача 11.7.7 *Доказать, что в случае D) канал* Φ *является антидеградируемым и следовательно* $Q(\Phi) = 0$.

Поскольку $I_c(S,\Phi) = H(\Phi[S]) - H(\tilde{\Phi}[S])$, для пропускной способности $Q(\tilde{\Phi})$ имеет место выражение, аналогичное (9.40), в котором I_c заменено на $-I_c$. Следовательно,

$$\sup_S I_c(S,\Phi) \leq Q(\Phi),$$
$$\sup_S [-I_c(S,\Phi)] \leq Q(\tilde{\Phi}).$$

Таким образом, если когерентная информация $I_c(S,\Phi)$ принимает значения разных знаков, то обе пропускные способности положительны и следовательно, канал Φ не является ни деградируемым, ни антидеградируемым.

Задача 11.7.8 *Рассмотрим канал с аддитивным классическим гауссовским шумом интенсивности* N_c *(случай* B_2*),* $k = 1$*). Показать, что*

$$F(N) \equiv I_c(S(N),\Phi) = g(N+N_c) - g\left(\frac{D+N_c-1}{2}\right) - g\left(\frac{D-N_c-1}{2}\right),$$

где $D = \sqrt{(N_c+1)^2 + 4N_cN}$. *При этом* $F(0) = 0$,

$$\lim_{N\to\infty} F(N) = -\log eN_c,$$

где использована асимптотика $g(x) \simeq \log ex$, $x \to \infty$, *следовательно,* $\lim_{N\to\infty} F(N) = 0$, *если* $N_c = e^{-1}$. *Рассматривая функцию* $F(N)$ *при* $N_c = 0.99e^{-1}$, *показать, что она принимает значения разных знаков, и, следовательно, канал* Φ *не является ни деградируемым, ни антидеградируемым.*

11.8 Комментарии

1. Квантовый гармонический осциллятор, впервые исследованный Дираком [8], лежит в основе многих математических моделей квантовой оптики и квантовой электроники. В частности, полезное символическое исчисление, позволяющее получить алгебраические тождества, вытекающие из канонических коммутационных соотношений, такие как (11.46), развито в книге Люиселла [15]. Анализ, основанный на когерентных состояниях, разработан Глаубером [7]. Формула (11.14) – это знаменитое P-представление операторов плотности, широко используемое в квантовой оптике. Представление поля излучения в виде ансамбля квантовых

осцилляторов, также предложенное Дираком, в подходящей для приложений в квантовой оптике форме рассмотрено, например, в книгах Клаудера и Сударшана [12] и Хелстрома [30].

Используя переполненность системы когерентных векторов (задача 11.1.4)), можно определить нечеткую наблюдаемую со значениями в $\mathbb{C} \equiv \mathbb{R}^2$

$$M(B) = \frac{1}{\pi} \int_B |\zeta\rangle\langle\zeta| d^2\zeta. \tag{11.160}$$

которая играет ключевую роль для описания "приближенного совместного измерения" наблюдаемых q и p. Подробное обсуждение этого вопроса и дальнейшие ссылки см. в книгах Холево [37], Дэвиса [75].

2. На полезность представления ККС в симметричной форме (11.28) было указано Сигалом [27]. Доказательство следующей теоремы можно найти, например, в [37].

Теорема единственности Стоуна - фон Неймана. *Пусть V_x, U_y; $x, y \in \mathbf{R}^s$ – сильно непрерывные группы унитарных операторов в сепарабельном гильбертовом пространстве $\mathcal{H}$, удовлетворяющие ККС Вейля (11.26). Тогда V_x, U_y унитарно эквивалентны прямой сумме не более чем счетного числа копий представления Шредингера (11.25). В частности, любое неприводимое представление ККС унитарно эквивалентно представлению Шредингера.*

Работа Вильямсона [155] содержит общую классификацию классических квадратичных гамильтонианов (необязательно положительных). Доказательство леммы 11.2.1 см. также в главе V книги [37].

3. Многочисленные физические приложения линейной динамики систем с квадратичным гамильтонианом рассматриваются в книге Малкина и Манько [18].

4. Рассмотрение квантовых гауссовских состояний основано на главе V книги Холево [37], где можно найти доказательства утверждений задач 11.4.1, 11.4.2, 11.4.3, 11.4.5. Условие (11.53) является не только необходимым, но и достаточным для положительности оператора S. Подход к квантовым гауссовским состояниям, основанный на характеристической функции, опирается на очевидные аналогии с классической теорией вероятностей, и является, по-видимому, наиболее прямым и аналитически прозрачным.

Формулы для энтропии квантовых состояний общего вида были получены Холево, Сома и Хирота [102]. Характеризация гауссовского состояния как состояния с максимальной энтропией (лемма 11.4.4) тесна связана с принципом максимальной энтропии в статистической физике, см. например [15], раздел 6.6; аналогичное свойство для условной квантовой энтропии (лемма 11.4.5) было установлено Айсертом и Вольфом [81].

Очищение произвольного гауссовского состояния построено в работе Холево [32].

5. Формула для пропускной способности гауссовского c-q канала была предугадана Гордоном [89]. Приведенное доказательство следует работе Холево [35], где рассмотрен и случай многомодового гауссовского c-q канала, в частности, реалистичная модель временно́го классического сигнала на фоне окрашенного квантового шума. Пропускная способность для канала со сжатым гауссовским шумом вычислена в работе Холево, Сома и Хирота [102]. Более детальную информацию о скорости сходимости вероятности ошибки для гауссовских каналов с чистыми состояниями можно получить, модифицируя оценки функции надежности из раздела 5.7 на случай каналов с бесконечным алфавитом и входными ограничениями [103].

Общее описание бозонных гауссовских каналов было дано в работе Холево и Вернера [105] (см. также обзор [81]), следуя работе [31]. Используя общее описание бозонного гауссовского канала, можно показать, что в квантовой оптике такие каналы могут быть реализованы из блоков, соответствующих линейным многоканальным смесителям и основным одномодовым преобразователям, таким как параметрический усилитель [67].

С другой стороны, гауссовские каналы представляют собой вполне положительные отображения C^*-алгебры ККС; Демоэн, Ванеуверцвийн и Вербер [77] показали, что неотрицательная определенность матриц (11.104) необходима и достаточна для полной положительности отображения Φ.

Гауссовские наблюдаемые фактически соответствуют гауссовским q-c каналам; они рассматриваются в гл. VI книги [37]. В частности, там показано, что расширение Наймарка для нечеткой наблюдаемой (11.160) дается спектральной мерой коммутирующих операторов $q + q_C, p - p_C$, где мода C находится в вакуумном состоянии.

Гауссовские каналы, разрушающие сцепленность, исследованы в работе Холево [43].

6. Принцип "пропускная способность гауссовского канала достигается для гауссовского сигнала" выполняется в классической теории информации [72]. Гипотеза об оптимальных гауссовских ансамблях экстраполирует аналогичный принцип на квантовые гауссовские каналы, однако в этом направлении пока получены лишь частичные результаты. Имеются основания полагать, что подтверждение этой гипотезы для классической пропускной способности квантового гауссовского канала по крайней мере частично зависит от решения проблемы аддитивности. В работе Вольфа, Гидке и Сирака [159] показано, что если гипотеза аддитивности верна, то среднее состояние $\bar{S}_\pi$ оптимального ансамбля должно быть гауссовским.

Теорема 11.6.1, доказанная в статье Холево и Вернера [105], означает выполнение "гауссовского" принципа для классической пропускной способности с использованием сцепленного состояния. Предложение 11.6.2, подтверждающее этот принцип для квантовой пропускной способности

деградируемых гауссовских каналов, основано на результатах Вольфа, Перес-Гарсиа и Гидке [160].

7. Классификация одномодовых гауссовских каналов получена Холево [43], см. также [69].

Пропускные способности аттенюатора/усилителя изучались в работе [105]. Предложение 11.7.2 для аттенюатора с вакуумным шумом доказано в статье Джованнетти, Гуха, Ллойда, Макконе, Шапиро и Юна [88].

Утверждение о том, что квантовая пропускная способность аттенюатора/усилителя дается гауссовским выражением, полученным в [105], было установлено Вольфом и Перес-Гарсиа в [160]. Области нулевой квантовой пропускной способности описаны в статье Карузо, Джованнетти и Холево [69]. Решения задач из раздела 11.7.4, в частности, подробное описание канала, комплементарного к каналу с аддитивным классическим гауссовским шумом см. в работе Холево [41].

Литература

1. Бурнашев М. В., Холево А. С. *О функции надежности квантового канала связи*, Пробл. предачи информ. **34**, 1-13 1998; e-print quant-ph/9703013, 1997.
2. К. А. Валиев, А. А. Кокин, Квантовые компьютеры: надежды и реальность, М.-Ижевск: РХД, 2001
3. Г. Вейль, Теория групп и квантовая механика, М.: Наука, 1986
4. Э. П. Вигнер, Теория групп и ее приложения к квантово-механической теории атомных спектров, М.: Физматгиз, 1961
5. Р. Галлагер, Теория информации и надежная связь, М.: Советское радио, 1974.
6. И. М. Глазман, Ю. И. Любич, Конечномерный линейный анализ в задачах, М.: Наука, 1969
7. Р. Глаубер, Оптическая когерентность и статистика фотонов, в. с. Квантовая оптика и квантовая радиофизика, М. Мир, 1966
8. П. А. М. Дирак, Принципы квантовой механики, М, Наука, 1970
9. Б. Б. Кадомцев, Динамика и информация, М.: УФН, 1999.
10. А. Китаев, А. Шень, М. Вялый, Классические и квантовые вычисления, М.: МЦНМО, 1999
11. Китаев А. Ю., Квантовые вычисления: алгоритмы и исправление ошибок, УМН, **52** No. 6, стр. 53–112 (1997)
12. Клаудер Дж. и Сударшан Э., Основы квантовой оптики, Москва, Мир, (1970)
13. А. И. Кострикин, Ю. И. Манин, Линейная алгебра и геометрия, М.: Наука, 1986
14. А. А. Курикша, Квантовая оптика и оптическая локация, М.: Советское радио, 1973
15. У. Люиселл, Излучение и шумы в квантовой оптике, М. Наука, 1972
16. Дж. Макки, Лекции по математическим основам квантовой механики, М., Мир, 1965
17. Г. Г. Магарил-Ильяев, В. М. Тихомиров, Выпуклый анализ и его приложения, М.: Едиториал УРСС, 2003
18. И. А. Малкин, В. И. Манько, Динамические симметрии и когерентные состояния квантовых систем, М.: Наука, 1979
19. Митюгов В. В. Физические основы теории информации, М.: Советское радио, 1976
20. Е. А. Морозова, Н. Н. Ченцов, "Марковская инвариантная геометрия на многообразиях состояний. Итоги науки и техники," Совр. пробл. матем., Новейшие достижения, т. 36, 69-102, ВИНИТИ, 1990.
21. фон Нейман Дж., Математические основы квантовой механики, Москва, Наука, (1964).
22. Нильсен М. А., Чанг И., Квантовые вычисления и квантовая информация, М.: Мир, 2006

23. М. А. Наймарк, Спектральные функции симметричного оператора, Изв. АН СССР (сер. матем.) 4, 3, 277-318 1940
24. Рид М., Саймон Б., Методы современной математической физики, Т. 1, М.: Мир, Пер. с англ. (1985).
25. Рокафеллар Р. Выпуклый анализ, М.:Мир, Пер. с англ. 1973
26. Сарымсаков Т. А., Введение в квантовую теорию вероятностей, ФАН, Ташкент, 1985.
27. Сигал И., Математические проблемы релятивистской физики, Москва, Мир, стр. 191, (1968)
28. Стратонович Р. Л., Ванцян А. Г., Об асимптотически безошибочном декодировании в чистых квантовых каналах, Пробл. управл. и теор. информ. 7 No. 3, стр. 161–174 (1978)
29. Л. Д. Фаддеев, О. А. Якубовский, Лекции по квантовой механике для студентов-математиков, Изд-во ЛГУ 1980
30. Хелстром К., Квантовая теория проверки гипотез и оценивания, пер. с англ., М.: Мир, 1978
31. А. С. Холево, К математической теории квантовых каналов связи, Пробл. передачи информ. 8, 1, 1972, 63-71.
32. Холево А. С. *О квазиэквивалентности локально-нормальных состояний*, ТМФ, **13**, N2, 184-199, 1972.
33. Холево А. С. Информационные аспекты квантовых измерений, Пробл. передачи информ. 9 No. 2, стр. 31–42 (1973)
34. Некоторые оценки для количества информации, передаваемого квантовым каналом связи, Пробл. передачи информ. 9 No. 3, стр. 3–11 (1973)
35. Холево А. С. Квантовые теоремы кодирования // УМН. 1998. Т. 53. С. 193-230; arXiv: quant-ph/9809023.
36. Холево А. С. *Статистическая структура квантовой теории*, М.-Ижевск: 2003.
37. Холево А. С., Вероятностные и статистические аспекты квантовой теории, 2-е изд., М.-Ижевск: ИКИ, 2004
38. Холево А. С., Введение в квантовую теорию информации, М.: МЦНМО, 2003
39. Холево А. С., Широков М. Е., Вернер Р. Ф., О понятии сцепленности в гильбертовых пространствах// УМН. 2005. Т. 60, no. 2, С. 153-154; Separability and entanglement-breaking in infinite dimensions, e-print quant-ph/0504204.
40. Холево А. С. *Комплементарные каналы и проблема аддитивностиO*, Теор. вероят. и ее примен., **51**, 133-134 2006; e-print quant-ph/0509101.
41. Холево А. С. Одномодовые квантовые гауссовские каналы: структура и квантовая пропускная способность// Проблемы передачи информации. 2007. Т. 43, С. 1-11; e-print quant-ph/0607051.
42. Холево А. С. Классические пропускные способности квантового канала // Теор. вероят. и ее примен. 2003. Т. 48. С. 359-374; e-print quant-ph/0211170.
43. A. S. Holevo, Entanglement-breaking channels in infinite dimensions, quant-ph/0802.0235
44. Холево А. С., Широков М. Е., Непрерывные ансамбли и пропускная способность квантовых каналов бесконечной размерности// Теор. вероят. и ее примен. 2005, Т. 50, no. 1, С. 98-114; e-print quant-ph/0408176.
45. Чисар И., Кернер Я., Теория информации, М.: Мир 1985.
46. C. Shannon, W. Weaver: *The mathematical theory of communication*, Univ. Illinois Press, Urbana Ill. 1949. Шеннон К. Математическая теория связи. В кн.: теория передачи электрических сигналов при наличии помех. М.: ИЛ 1953.

47. Широков М.Е. Энтропийные характеристики подмножеств состояний I // Известия РАН. Серия математическая, 2006, Т. 70, N.6, с. 193-222.
48. Широков М.Е. О свойствах квантовых каналов, связанных с классической пропускной способностью// Теория вероятностей и ее применения. 2007. Т.52. N.2. С. 301-335.
49. A. Abeyesinghe, I. Devetak, P. Hayden, A. Winter, The mother of all protocols: Restructuring quantum information's family tree quant-ph/0606225
50. C. Adami, N. J. Cerf, *Capacity of noisy quantum channels*, Phys. Rev. A **56**, pp. 3470-3485 1997.
51. R. Ahlswede, I. Csiszár, *Common randomness in information theory and cryptography – Part II: CR capacity*, IEEE Trans. Inform. Theory, bf 44, 225-240, 1998.
52. G. Alber, T. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. F. Werner, *Quantum information. An introduction to basic theoretical concepts and experiments*, Springer, Berlin, 2001
53. H. Araki, On a characterization of the state space in quantum mechanics, Commun.Math. Phys., v. 75, n. 1, 1-24, 1980
54. K. M. R. Audenaert *A sharp Fannes-type inequality for the von Neumann entropy*, e-print quant-ph/0610146.
55. H. Barnum, E. Knill, M. A. Nielsen: *On quantum fidelities and channel capacities*, IEEE Trans. Inform. Theory **46** N4, pp. 1317-1329 1998; e-print quant-ph/9809010.
56. H. Barnum, M. A. Nielsen, B. Schumacher: *Information transmission through a noisy quantum channel*, Phys. Rev. A **57**, pp. 4153-4175 1998.
57. J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Physics, 1, 195-200, 1964
58. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A Peres and W. K. Wootters, Teleporting an unknown quantum state via dual classical and EPR channels, Phys. Rev. Lett. 70, pp. 1895Ц1899 (1993).
59. C. H. Bennett, I. Devetak, P. W. Shor, J. A. Smolin, Inequalities and separations among assisted capacities of quantum channels, quant-ph/0406086
60. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters, Mixed state entanglement and quantum error correction, quant-ph/9604024
61. C. H. Bennett, C. A. Fuchs, J. A. Smolin, *Entanglement-enhanced classical communication on a noisy quantum channel*, in Quantum Communication, Computing and Measurement, Proc. QCM96, ed. by O. Hirota, A. S. Holevo and C. M. Caves, New York: Plenum 1997, pp. 79-88; e-print quant-ph/9611006.
62. C. H. Bennett, P. W. Shor: *Quantum Information Theory*, IEEE Trans. Inform. Theory **44** N6, pp. 2724-2742 1998.
63. C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal: *Entanglement-assisted capacity and the reverse Shannon theorem*, IEEE Trans. Inform. Theory; e-print quant-ph/0106052.
64. C. H. Bennett and S. J. Wiesner, Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states, Phys. Rev. Lett. 69, pp. 2881Ц 2884 (1992).
65. R. Bhatia: *Matrix Analysis*, Springer, New York 1997.
66. I. Bjelaković, R. Siegmund-Schultze, *A new proof of the monotonicity of quantum relative entropy for finite quantum systems*, e-print quant-ph/0307170.
67. S. L. Braunstein, *Squeezing as an irreducible resource*, e-print quant-ph/9904002.
68. R. Calderbank, P. W. Shor: *Good quantum error-correcting codes exist*, Phys. Rev. A **54**, pp 1098-1105 1996.Good Quantum Error-Correcting Codes Exist, Phys. Rev. A, Vol. 54, No. 2, pp. 1098-1106, 1996; quant-ph/9512032

69. F. Caruso, V. Giovannetti, A. S. Holevo, One-mode Bosonic Gaussian channels: a full weak-degradability classification, New Journal of Physics 8, 310 (2006); quant-ph/0609013
70. C. M. Caves, P. B. Drummond, *Quantum limits of bosonic communication rates*, Rev. Mod. Phys. **66**, 481-538 1994.
71. M.-D. Choi. Completely positive maps on complex matrices // Linear Alg. and Its Appl. 1975. V. 10. P. 285-290.
72. T. M. Cover, J. A. Thomas: *Elements of Information Theory*, Wiley Series in Telecommunications, John Wiley & Sons, New York 1991).
73. T. S. Cubitt, M.-B. Ruskai, G. Smith, The structure of degradable quantum channels, arXiv:0802.1360
74. E. B. Davies: *Information and quantum measurement*, IEEE Trans. Inform. Theory **24** N6, pp. 596-599 1978.
75. E. B. Davies: *Quantum theory of open systems*, Academic Press, London 1976.
76. Dell'Antonio G.F., On the limits of sequences of normal states, Commun. Pure Appl. Math., 20, 413-430, 1967;
77. B. Demoen, P. Vanheuverzwijn, A. Verbeure, *Completely positive quasi-free maps on the CCR algebra*, Rep. Math. Phys. **15**, 27-39, 1979.
78. I. Devetak, *The private classical information capacity and quantum information capacity of a quantum channel*, e-print no. quant-ph/0304127, 2003.
79. I. Devetak, P. Shor, *The capacity of a quantum channel for simultaneous transition of classical and quantum information*, quant-ph/0311131.
80. D. Di Vincenzo, P. W. Shor, J. Smolin, *Quantum channel capacities of very noisy channels*, Phys. Rev. A **57**, 830-839, 1998.
81. J. Eisert, M. M. Wolf, Gaussian quantum channels, quant-ph/0505151
82. E. G. Effros, A matrix convexity approach to some celebrated quantum inequalities, math-ph/0802.1234
83. M. Fannes, A continuity property of quantum entropy for spin lattice systems, Commun. Math. Phys., 31, 291-294, 1973
84. *Foundations of quantum mechanics and ordered linear spaces*. Eds. A. Hartkamper and H. Neumann, Lect. Notes Phys. **29**, Springer-Verlag, New York-Heidelberg-Berlin 1974.
85. C. A. Fuchs, *Quantum mechanics as quantum information (and only a little more)*, e-print quant-ph/0205039.
86. M. Fukuda, C. King, D. Moser, Comments on Hastings' Additivity Counterexamples, e-print arXiv:0905.3697
87. M. Fukuda, M. M. Wolf, *Simplifying additivity problems using direct sum constructions,* J. Math. Phys., vol. 48, 072101, 2007.
88. V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro and H. P. Yuen, *Classical capacity of the lossy bosonic channel: the exact solution*, e-print quant-ph/0308012.
89. J. P. Gordon, *Noise at optical frequencies; information theory*, In: Quantum Electronics and Coherent Light, Proc. Int. School Phys. "Enrico Fermi", Course XXXI, ed. P. A. Miles, Academic Press, New York 1964, pp.156-181.
90. M. B. Hastings, A counterexample to additivity of minimum output entropy, quant-ph/0809.3972
91. P. Hausladen, R. Josza, B. Schumacher, M. Westmoreland, W. Wootters: *Classical information capacity of a quantum channel*, Phys. Rev. A **54**, pp. 1869-1876 1996.
92. M. Hayashi, *Quantum information: an introduction*, Springer, Berlin 2006.
93. M. Hayashi and H.Nagaoka, *General formulas for capacity of classical-quantum channels*, e-print quant-ph/0206186.

94. P. Hayden, The maximal p-norm multiplicativity conjecture is false quant-ph/0707.3291
95. P. Hayden, Entanglement in Random Subspaces, Proc. of QCMC04. AIP conference procedings vol. 734, pp. 226-229, 2004; quant-ph/0409157
96. P. Hayden, P. W. Shor, A. Winter, Random quantum codes from Gaussian ensembles and an uncertainty relation, Open Syst. Inf. Dyn. **15**, pp. 71-89 2008; quant-ph/0712.0975
97. T. Hiroshima, Additivity and multiplicativity properties of some Gaussian channels for Gaussian inputs, Phys. Rev. A 73, 012330 (2006); quant-ph/0511006
98. W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Amer. Statist. Assoc. **58**, N31, pp. 13-30, 1963.
99. A. S. Holevo, Statistical definition of observable and the structure of statistical models. Rep. Math. Phys., 1985, v.22, N3, 385-407.
100. A. S. Holevo, *The capacity of quantum communication channel with general signal states*, e-print quant-ph/9611023, 1996; IEEE Trans. Inform. Theory **44**, 269-272 1998.
101. A. S. Holevo, *On entanglement-assisted classical capacity*, J. Math. Phys., **43**, 2002; e-print quant-ph/0106075.
102. A. S. Holevo, M. Sohma, O. Hirota, *Capacity of quantum Gaussian channels*, Phys. Rev. A **59**, 1820-1828, 1999.
103. A. S. Holevo, M. Sohma, O. Hirota, *Error exponents for quantum channels with constrained inputs*, Rep. Math. Phys. **46**, 343-358, 2000.
104. A. S. Holevo, M. E. Shirokov, On Shor's channel extension and constrained channels. Commun. Math. Phys. v. 249, 417-430, 2004. Arxiv e-print quant-ph/0306196
105. A. S. Holevo, R. F. Werner, *Evaluating capacities of Bosonic Gaussian channels*, Phys. Rev. A **63**, 032312, 2001,; LANL e-print quant-ph/9912067, 1999.
106. A. S. Holevo, *The additivity problem in quantum information theory*, Proc. ICM, Madrid, Spain, 2006, pp. 999-1018.
107. M. Horodecki, J. Oppenheim, A. Winter, Quantum state merging and negative information, Comm. Math. Phys. **269**, 107 (2006); quant-ph/0512247
108. M. Horodecki, P.W. Shor, M.B. Ruskai, *General entanglement breaking channels*, Rev. Math. Phys. **15**, 629-641, (2003); quant-ph/0302031.
109. R. Josza, B. Schumacher: *A new proof of the quantum noiseless coding theorem*, J. Modern Optics **41**, pp. 2343-2349 1994.
110. K. Kato, M. Osaki, O. Hirota, *Derivation of classical capacity of quantum channels for discrete information sources*, e-print quant-ph/9811085.
111. K. Kato, M. Osaki, T. Suzuki, M. Ban, O. Hirota, *Upper bound of the accessible information and lower bound of the Bayes cost in quantum signal detection processes*, Phys. Rev. A **54** 2718-2727 1996.
112. M. Keyl: *Fundamentals of Quantum Information*, e-print quant-ph/0202122, Feb. 2002.
113. C. King, *Additivity for a class of unital qubit channels*, e-print quant-ph/0103156.
114. C. King, *The capacity of quantum depolarizing channel*, e-print quant-ph/0204172.
115. C. King, K. Matsumoto, M. Natanson and M. B. Ruskai, *Properties of conjugate channels with applications to additivity and multiplicativity*, quant-ph/0509126.
116. E. Knill, R. Laflamme: *Theory of quantum error-correcting codes*, Phys. Rev. A **55**, pp. 900-911 1997.

117. K. Kraus: *States, Effects and Operations*, Springer Lecture Notes in Physics **190** 1983.
118. D. Kretschmann, R. Werner, *Tema con variazioni: quantum channel capacities*, e-print quant-ph/0311037.
119. D. Kretschmann, R. Werner, *Quantum channels with memory*, e-print quant-ph/0502106.
120. A. Lesniewski, M. B. Ruskai: *Monotone Riemannian metrics and relative entropy on noncommutative probability spaces*, J. Math. Phys. **40**, pp. 5702-5724 1999.
121. L. B. Levitin, *Optimal quantum measurement for two pure and mixed states*, In: Quantum Communications and Measurement, Proc. QCM94, ed. by V. P. Belavkin, O. Hirota, R. L. Hudson, Plenum, New York 1995, pp. 439-448.
122. E. H. Lieb, M. B. Ruskai, *Proof of the strong subadditivity of quantum mechanical entropy*, J. Math. Phys., **14**, 1938-1941 1973.
123. G. Lindblad: *Entropy, information and quantum measurements*, Commun. Math. Phys. **33**, pp. 305-322 1973.
124. G. Lindblad: *Expectations and entropy inequalities for finite quantum systems*, Commun. Math. Phys. **39**, pp. 111-119 1974.
125. G. Lindblad: *Completely positive maps and entropy inequalities*, Commun. Math. Phys. **40**, pp. 147-151 1975.
126. 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, pp. 71-80.
127. S. Lloyd, *Capacity of noisy quantum channel*, Phys. Rev. **A 55**, 1613-1622 1997.
128. G. Ludwig: *Foundations of Quantum Mechanics I*, Springer, Berlin-Heidelberg-New York 1983.
129. M. Ohya, D. Petz: *Quantum Entropy and Its Use*, Texts and Monographs in Physics, Springer, New York 1993.
130. M. Ozawa, *On information gain by quantum measurement of continuous observable*, J. Math. Phys. **27**, 759-763 1986.
131. M. Ozawa, *Quantum measuring process of continuous observables*, J. Math. Phys. **18**, 412-421 1987.
132. K. R. Parthasarathy, Probability measures on metric spaces, Academic Press, New York and London, 1967
133. D. Petz, Quantum information theory and quantum statistics, Berlin: Springer, 2008
134. M. B. Ruskai, S. Szarek, E. Werner, *A characterization of completely-positive trace-preserving maps on* $\mathcal{M}_2$, e-print quant-ph/0005004.
135. 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; e-print quant-ph/9812062.
136. B. Schumacher, M. D. Westmoreland, *Sending classical information via noisy quantum channel*, Phys. Rev. A. **56**, 131-138 1997.
137. B. Schumacher, Westmoreland M. D., Quantum privacy and quantum coherence// Phys. Rev. Lett., 1998, vol. 80, pp. 5695-5697; e-print quant-ph/9709058.
138. B. Schumacher, M. D. Westmoreland, *Optimal signal ensembles*, e-print quant-ph/9912122.
139. B. Schumacher, M. D. Westmoreland, *Approximate quantum error correction*, e-print quant-ph/0112106.

140. M. E. Shirokov, The Holevo capacity of infinite dimensional channels and the additivity problem, Commun. Math. Phys. 2006. 262, N.2, 137-159; e-print quant-ph/0408009;
141. P. W. Shor, *Introduction to quantum algorithms*, e-print quant-ph/0005003.
142. P. W. Shor, *On the number of elements needed in a POVM attaining the accessible information*, e-print quant-ph/0009077.
143. P. W. Shor, *Additivity of the classical capacity of entanglement-breaking quantum channels*, J. Math. Phys., 2002, vol. 43, pp. 4334-4340; e-print quant-ph/0201149.
144. P. W. Shor, *The adaptive classical capacity of a quantum channel, or information capacity of 3 symmetric pure states in three dimensions*, e-print quant-ph/0206058.
145. P. W. Shor, Equivalence of additivity questions in quantum information theory // Commun. Math. Phys. 2004. V. 246. P. 453-472; arXiv: quant-ph/0305035.
146. P. W. Shor, The classical capacity achievable by a quantum channel assisted by limited entanglement, in: Quantum Information, Statistics, Probability, Ed. by O. Hirota, Rinton Press, Inc., Princeton, New Jersey 2004; quant-ph/0402129.
147. G. Smith and J.A. Smolin, Degenerate quantum codes for Pauli channels Phys. Rev. Lett. 98, 030501 (2007).
148. A. Steane: *Quantum computing*, Rept. Prog. Phys. **61**, pp. 117-173 1997.
149. W. F. Stinespring. Positive functions on C^*-algebras // Proc. Amer. Math. Soc. 1955. V. 6. P. 211-316.
150. A. Uhlmann: *The "transition probability" in the state space of a $*$-algebra*, Rept. Math. Phys. **9**, pp. 273-279 1976.
151. A. Uhlmann, *Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory*, Commun. Math. Phys., vol. 54, 21-32 1977.
152. A. Wehrl: *General properties of entropy*, Rev. Mod. Phys. **50**, pp. 221-260 1978.
153. Werner R.F., Wolf M. M., Bound entangled Gaussian states// e-print quant-ph/0009118.
154. A. S. Wightman, Hilbert's sixth problem: mathematical treatment of the axioms of physics, Proc. Symp. in pure math. **28**, pt. 1, 147-240 1977
155. J. Williamson: *On the algebraic problem concerning the normal forms of linear dynamical systems*, Am. J. Math **58**, 141 (1936); Am. J. Math **59**, 599 (1937); **61**, 897 (1939)
156. A. Winter, *Coding theorem and strong converse for quantum channels*, IEEE Trans. Inform. Theory, **45**, 2481-2485, 1999.
157. A. Winter, *Compression of sources of probability distributions and density operators*, e-print quant-ph/0208131.
158. A. Winter, The maximum output p-norm of quantum channels is not multiplicative for any p>2, e-print quant-ph/ 0707.0402
159. M. Wolf, G. Giedke, J. I. Cirac, *Extremality of Gaussian quantum states*, quant-ph/0509154.
160. M. Wolf, D. Pérez-Garsía, G. Giedke, Quantum Capacities of Bosonic Channels , Phys. Rev. Lett. 98, 130501 (2007) quant-ph/0606132.