А.С.Холево

ВВЕДЕНИЕ В КВАНТОВУЮ ТЕОРИЮ ИНФОРМАЦИИ

М.: МЦНМО, 2002. —128с.

Лекции посвящены изложению основных понятий и ряда строгих результатов новой научной дисциплины - квантовой теории информации. Возможности квантовых систем передачи и преобразования информации проиллюстрированы на примерах сверхплотного кодирования, квантовой телепортации и квантовых алгоритмов.

Рассматриваются энтропийные и информационные характеристики квантовых систем. Подробно обсуждается понятие квантового канала связи, его классическая и квантовая пропускные способности, а также передача классической информации с помощью сцепленного состояния. Сформулировано несколько принципиальных открытых проблем, решение которых явилось бы существенным вкладом в квантовую теорию информации.

В лекциях приведены необходимые сведения из классической теории информации и дается подробное введение в статистическую структуру квантовой теории, поэтому для их понимания достаточно владения основными общематематическими дисциплинами.

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Недавно отмечалось 50-летие двух революционных научных открытий, которые по существу предопределили облик современного мира в той важнейшей его части, которая касается процессов хранения, обработки и передачи информации. Это — изобретение транзистора, открывшее путь к миниатюризации [1] и радикальному снижению материальных и энергетических затрат при создании систем обработки информации, и создание математических основ теории информации, заложивших принципы рационального и помехоустойчивого дизайна таких систем и обрабатываемых ими массивов данных [2].

В настоящее время происходит создание теоретических и экспериментальных основ *квантовой теории информации*. Построены демонстрационные системы, основанные на принципах квантовой криптографии. Обсуждается идея квантового компьютера, сулящая в перспективе немыслимые ранее возможности [3]. Независимо от того, как скоро могут быть практически реализованы подобные проекты, квантовая теория информации является новым направлением, дающим ключ к пониманию фундаментальных закономерностей Природы, до недавних пор остававшихся вне поля зрения исследователей. Она также стимулирует развитие экспериментальной физики, значительно расширяющее возможности манипулирования состояниями микросистем и потенциально важное для новых эффективных приложений.

Центральным результатом классической теории информации являются *теоремы кодирования*, устанавливающие возможность помехоустойчивой передачи и обработки информации при скоростях, не превышающих некоторую вполне определенную величину, характеризующую данную систему преобразования информации (для определенности обычно говорят про *канал связи*) и называемую *пропускной способностью*. Почти одновременно с появлением пионерских работ Шеннона и с созданием математических основ

---

[1] О степени миниатюризации в современных вычислительных устройствах говорит следующий факт: микрочип операционной памяти емкостью 16 Mb содержит 33 млн. транзисторов на площади менее 1 см^2 (данные 1999 г.).

[2] В качестве примеров эффективного применения идей теории информации можно привести коды Лемпеля — Зива, используемые для сжатия файлов в UNIX и PC, а также коды Рида — Соломона, применяемые для исправления ошибок при воспроизведении CD.

[3] Практическая реализация этого проекта предполагает технологическую революцию, по значимости сопоставимую по крайней мере с изобретением транзистора. Компетентное обсуждение возможных подходов см. в книге: *Валиев К. А., Кокин А. А.* Квантовые компьютеры: надежды и реальность. — М. — Ижевск: РХД, 2001.

теории информации встал вопрос о фундаментальных ограничениях на возможности передачи сообщений, накладываемых природой физического носителя информации. Проблема определения пропускной способности квантового канала связи сформировалась в 60-е годы и восходит к более ранним классическим трудам Габора и Бриллюэна, поставившим вопрос о квантово-механических пределах точности и скорости передачи информации. Эти работы заложили физические основы и подняли проблему адекватного математического рассмотрения всего данного круга вопросов. Принципиальные шаги в этом направлении были сделаны в 70-е годы, когда была построена некоммутативная теория статистических решений, найдено первое доказательство квантовой энтропийной границы и обнаружена строгая супераддитивность шенноновской информации в квантовом канале связи без памяти.

Существенный прогресс был достигнут в последние годы, когда были доказаны прямые теоремы кодирования, устанавливающие достижимость квантовой энтропийной границы, и понято, что квантовый канал характеризуется целым набором пропускных способностей, в зависимости от рода передаваемой информации и специфики используемых ресурсов. В значительной мере этот прогресс стимулирован современным развитием идей квантовой теории информации и квантового компьютинга. С другой стороны, вопрос о пропускной способности квантового канала связи представляет большой интерес в связи с квантовыми кодами, исправляющими ошибки, исследованием эффективности и сложности квантовых алгоритмов, и в целым рядом других вопросов. Настоящие лекции посвящены в основном систематическому изложению строгих результатов, относящихся к понятию квантового канала связи и пропускных способностей. В них также дается подробное введение в статистическую структуру квантовой теории, составляющую базис для квантовой теории информации, однако их не следует рассматривать как всеобъемлющий курс по квантовой теории информации: так, здесь не затрагиваются проблемы квантовой криптографии, а также получившая значительное развитие в последнее время количественная теория *сцепленности* квантовых состояний [1]. Весьма фрагментарно рассматриваются и квантовые алгоритмы.

---

[1] Существуют определенные сложности при подборе адекватного перевода терминов квантовой теории информации на русский язык. В частности, важнейшее понятие *entanglement* было переведено как *перепутанность*, что совершенно не отвечает существу дела. В настоящих лекциях мы делаем попытку исправить положение, используя, как нам кажется, более подходящий термин *сцепленность*.

Эти разделы читатель, освоивший вводные главы настоящих лекций, сможет изучить самостоятельно по другим источникам (см. список литературы [1]). Доказательства целого ряда вспомогательных результатов и следствий оставлены читателю в качестве полезных упражнений. С другой стороны, в лекциях сформулировано несколько принципиальных открытых проблем, решение которых явилось бы существенным вкладом в квантовую теорию информации.

В основу настоящих лекций положен курс, прочитанный автором в Математическом институте им. В. А. Стеклова РАН для слушателей Колледжа математической физики Московского Независимого Университета. Автор благодарен слушателям лекций А. В. Булинскому и Е. Е. Дьяконовой за помощь с конспектом, а также М. М. Деминову и Д. М. Белову, осуществившим предварительный компьютерный набор.

Работа частично поддержена грантом INTAS 00–738.

[1] Одним из основных источников новейшей информации является Лос-Аламосский электронный архив по квантовой физике (quant-ph), представленный на сервере ИТЭФ: http://xxx.itep.ru. Ссылка quant-ph/9809023 означает электронный препринт № 023, опубликованный в сентябре 1998 г., и т. п.

ГЛАВА 1

ОСНОВНЫЕ ПОНЯТИЯ КЛАССИЧЕСКОЙ ТЕОРИИ ИНФОРМАЦИИ

§ 1.1. Энтропия случайной величины и сжатие данных

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

$$H(X):=-\sum_{x\in\mathcal{X}}p_x\log p_x, \tag{1.1}$$

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

Задача 1. Докажите, что $0\leqslant H(X)\leqslant\log|\mathcal{X}|$, причем минимальное значение принимается на вырожденных распределениях, а максимальное — на равномерном.

Обычно $H(X)$ интерпретируется как мера неопределенности, изменчивости или информационного содержания случайной величины X. Поясним последнее утверждение. В этом параграфе мы следуем в основном [9].

Рассмотрим случайный источник, который порождает последовательность независимых одинаково распределенных случайных величин с распределением p. Последовательность $w=(x_1,\ldots,x_n)$ букв алфавита $\mathcal{X}$ называется *словом* длины n. Общее количество таких слов $|\mathcal{X}|^n=2^{n\log|\mathcal{X}|}$. Поэтому можно закодировать все эти слова, используя двоичные последовательности длины $n\log|\mathcal{X}|$, т. е. $n\log|\mathcal{X}|$ *бит*. Однако, используя то обстоятельство, что p в общем случае неравномерное распределение, можно предложить лучший способ кодирования. Возможность сжатия данных тесно связана со свойством *асимптотической равнораспределенности*, которое является прямым следствием закона больших чисел:

ТЕОРЕМА 1. *Если* $X_1, \ldots, X_n, \ldots$ *— независимые и одинаково распределенные случайные величины с распределением* $p = \{p_x\}$, *то*

$$-\frac{1}{n}\sum_{i=1}^{n} \log p_{x_i} \longrightarrow H(X) \qquad \text{по вероятности.} \tag{1.2}$$

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

$$\mathsf{P}\left\{\left| -\frac{1}{n}\sum_{i=1}^{n} \log p_{x_i} - H(X)\right| < \delta\right\} > 1 - \varepsilon. \tag{1.3}$$

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

$$p_w = p_{x_1} \cdot \ldots \cdot p_{x_n} = 2^{-n\left(-\frac{1}{n}\sum\limits_{i=1}^{n} \log p_{x_i}\right)} \tag{1.4}$$

мы теперь можем использовать соотношение (1.3) чтобы ввести понятие *типичного слова*: слово w, имеющее вероятность p_w, называется δ-типичным, если

$$2^{-n(H(X)+\delta)} < p_w < 2^{-n(H(X)-\delta)}. \tag{1.5}$$

Непосредственно устанавливаются следующие свойства типичных слов:

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

2) для достаточно больших n существует, по крайней мере, $(1-\varepsilon)2^{n(H(X)-\delta)}$ типичных слов;

3) множество нетипичных слов имеет вероятность $\leqslant \varepsilon$.

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

Задача 2. Докажите последнее утверждение.

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

§ 1.2. Пропускная способность канала с шумом

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

$$I(X;Y) = H(X) - H(X \mid Y), \tag{1.6}$$

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

$$H(X \mid Y) = \sum_y p_y H(X \mid Y = y) := -\sum_y p_y \sum_x \frac{p_{x,y}}{p_y} \log \frac{p_{x,y}}{p_y} =$$

$$= -\sum_{x,y} p_{x,y} \log p_{x,y} + \sum_y p_y \log p_y = H(X,Y) - H(Y).$$

Здесь $H(X,Y)$ *совместная энтропия* пары случайных величин (X,Y), соответствующая совместному распределению $p_{x,y} = p(y|x)p_x$. Подставляя эту формулу в определение шенноновского количества информации (1.6), мы видим, что оно симметрично по X и Y, и поэтому может быть также названо *взаимной информацией*

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = H(Y) - H(Y \mid X), \tag{1.7}$$

где в последней формуле уже $H(Y)$ может быть интерпретирована, как информационное содержание выхода, а $H(Y \mid X)$ как его бесполезная составляющая, обусловленная *шумом*. Взаимная информация всегда неотрицательна: тот факт, что $H(X) \geqslant H(X|Y)$ легко вытекает из вогнутости функции $-x \log x$ (задача 3).

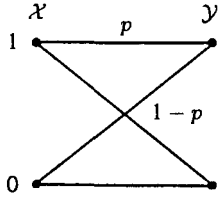


Рис. 1. Двоичный симметричный канал

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

Если посылается последовательность букв, и канал $p(y \mid x)$ действует независимо на каждую посланную букву, то он называется

каналом без памяти. Пропускная способность такого канала определяется как

$$C = \max_{\{p_x\}} I(X; Y), \tag{1.8}$$

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

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

$$h(p) = -p \log p - (1-p) \log(1-p), \tag{1.9}$$

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

$$C = 1 - h(p), \tag{1.10}$$

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

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

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

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

ЛЕММА.

$$I(X^n; Y^n) \leqslant \sum_{i=1}^{n} I(X_i; Y_i). \tag{1.11}$$

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

$$H(X_1, \ldots, X_n) = \sum_{i=1}^{n} H(X_i \mid X_1, \ldots, X_{i-1}), \tag{1.12}$$

которое легко доказать по индукции, используя формулу:

$$H(X, Y) = H(X) + H(Y \mid X). \tag{1.13}$$

Тогда взаимная информация равна

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

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

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

Взяв максимум выражения (1.11), получаем аддитивность, в частности, $C_n = nC$.

ОПРЕДЕЛЕНИЕ. *Кодом* (W, V) *размера* N для канала $p(y \mid x)$ называется совокупность N слов $w^{(1)}, \ldots, w^{(N)}$ длины n вместе с разбиением множества $\mathcal{Y}^n$ на N непересекающихся подмножеств $V^{(0)}, V^{(1)}, \ldots, V^{(N)} \subset \mathcal{Y}^n$.

Подмножества $V^{(1)}, \ldots, V^{(N)}$ интерпретируются как области принятия решения: если на выходе принято значение $y^n \in V^{(j)}$, $j = 1, \ldots, N$, то принимается решение, что было послано слово $w^{(j)}$; если же принято $y^n \in V^{(0)}$, то никакого определенного решения не принимается. Таким образом, *максимальная вероятность ошибки* такого кода есть

$$P_e(W, V) = \max_{1 \leqslant j \leqslant N} \big(1 - p(V^{(j)} \mid w^{(j)})\big), \tag{1.14}$$

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

$$\overline{P}_e(W, V) = \frac{1}{N} \sum_{i=1}^{N} \big(1 - p(V^{(j)} \mid w^{(j)})\big) \leqslant P_e(W, V), \tag{1.15}$$

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

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

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

Задача 4. Показать, что максимальная вероятность ошибки удовлетворяет неравенствам:

$$P_e(W, V) \leqslant \| \underbrace{\delta_{ji} - p(V^{(j)} \mid w^{(i)})}_{a_{ji}} \|_1 \leqslant 2P_e(W, V), \tag{1.16}$$

где

$$\|a_{ji}\|_1 = \sup_{p_i} \frac{\sum\limits_j |\sum\limits_i a_{ji} p_i|}{\sum\limits_i |p_i|}. \tag{1.17}$$

Это дает аналитическую характеристику точности воспроизведения, удобную для перехода к квантовым каналам.

ЛЕММА (неравенство Фано). *Пусть X, Y случайные величины и $\widehat{X} = \widehat{X}(Y)$ — оценка случайной величины X с вероятностью ошибки $p_e = \mathbf{P}\{\widehat{X}(Y) \neq X\}$. Тогда*

$$H(X \mid Y) \leqslant h(p_e) + p_e \log(|\mathcal{X}| - 1) \leqslant 1 + p_e \log |\mathcal{X}|. \tag{1.18}$$

Доказательство. Пусть E — индикатор ошибки оценивания,

$$E = \begin{cases} 0, & \text{если } \widehat{X}(Y) = X, \\ 1, & \text{в противном случае.} \end{cases} \tag{1.19}$$

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

$$H(E \mid X, Y) = H(E, X \mid Y) - H(X \mid Y) = 0, \tag{1.20}$$

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

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

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

ТЕОРЕМА (о кодировании для канала с шумом). *Пусть*

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

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

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

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

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

$$nC = C_n \geqslant I(Z; Y^n) = H(Z) - H(Z \mid Y^n) \geqslant$$

$$\geqslant \log N - 1 - \mathsf{P}\{\widehat{Z}(Y^n) \neq Z\} \log N = \log N - 1 - \overline{P}_e(W, V).$$

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

$$nC \geqslant nR - 1 - p_e(n, 2^{nR}) nR,$$

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

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

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

Основная идея доказательства *прямой теоремы кодирования*, восходящая к работе Шеннона [15], состоит в использовании *случайного кодирования*. Рассмотрим N слов $w^{(1)}, \ldots, w^{(N)}$, выбираемых случайным образом независимо с распределением вероятностей

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

где однобуквенное распределение $\{p_x\}$ выбрано так, что оно максимизирует $I(X;Y)$. Заметим, что имеется примерно $2^{nH(X)}$ типичных слов на входе (и $2^{nH(Y)}$ на выходе), и в среднем $2^{nH(Y|X)}$ типичных слов на выходе для каждого входного слова w.

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

$$N \approx \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}. \qquad (1.21)$$

Таким образом, $N \approx 2^{nC}$. Конечно, это рассуждение в высшей степени эвристично; строгое доказательство, реализующее эту идею, можно найти, например, в [9].

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

ГЛАВА 2

СОСТОЯНИЯ И НАБЛЮДАЕМЫЕ

§ 2.1. Соглашения и обозначения

Прежде чем перейти к квантовой теории информации, необходимо изложить предварительные сведения о *статистической структуре квантовой теории*. Цель состоит не только в том, чтобы ввести определения и зафиксировать обозначения, но и в том, чтобы глубже разобраться в основах квантовой теории и ее вероятностной интерпретации (гораздо более полное изложение этих вопросов читатель найдет в [4; 11]).

Мы будем иметь дело с квантово-механическими системами, которые описываются конечномерными гильбертовыми пространствами. С одной стороны, уже в этом случае, причем наиболее наглядно, проявляются радикальные отличия квантовой статистики. С другой, именно системы с конечным числом уровней представляют интерес с точки зрения квантового компьютинга (впрочем, в квантовой теории передачи информации в последнее время большое внимание привлекли «системы с непрерывными переменными», которые описываются бесконечномерными пространствами).

Пусть $\mathcal{H}$ — гильбертово (унитарное) пространство, $\dim \mathcal{H} = d < \infty$. Мы будем использовать дираковские обозначения: вектор ψ из $\mathcal{H}$ (который удобно представлять себе как вектор-столбец) часто будет обозначаться $|\psi\rangle$; соответственно, $\langle\psi|$ обозначает эрмитово-сопряженный вектор-строку. При этом $\langle\varphi \mid \psi\rangle$ естественно обозначает скалярное произведение. Эти обозначения позволяют удобно записывать и операторы, например, $A = |\psi\rangle\langle\varphi|$ — оператор ранга 1, действующий на вектор $|\chi\rangle$ по формуле $A|\chi\rangle = |\psi\rangle\langle\varphi \mid \chi\rangle$. Если $\langle\psi \mid \psi\rangle = 1$, то $|\psi\rangle\langle\psi|$ — проектор на единичный вектор $|\psi\rangle$.

§ 2.2. Квантовые состояния

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

В выпуклых множествах особо важны *крайние точки*, не представимые в виде нетривиальной смеси других точек, т. е. $S = pS_1 + (1-p)S_2$, $0 < p < 1$, влечет $S = S_1 = S_2$. Крайние точки множества квантовых состояний $\mathcal{S}(\mathcal{H})$, называемые *чистыми состояниями*, это в точности одномерные проекторы $S_\psi = |\psi\rangle\langle\psi|$ (задача 5). В частности, спектральное разложение

$$S = \sum_{i=1}^{d} s_i |e_i\rangle\langle e_i|, \quad s_i \geqslant 0, \quad \sum_i s_i = 1, \tag{2.1}$$

где s_i — собственные числа, $|e_i\rangle$ — собственные векторы оператора S, показывает, что всякое состояние является смесью не более чем d чистых состояний, где $d = \dim \mathcal{H}$. 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 чистом состоянии.

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

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

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

Задача 6. Доказать, что если $\dim \mathcal{H} = d$, то $\mathcal{S}(\mathcal{H})$ погружается в вещественное пространство размерности $n = d^2 - 1$. Если же $\mathcal{H}$ евклидово (вещественное) пространство, то $n = d(d+1)/2 - 1$.

Спектральное разложение (2.1) показывает, что в случае множества квантовых состояний (как и для других выпуклых множеств с гладкой границей) теорема Каратеодори дает завышенное значение n. С другой стороны, для множества классических состояний (распределений вероятности на некотором фазовом пространстве), представляющего собой симплекс, эта теорема дает точное значение. Это наводит на мысль интерпретировать квантовую теорию как классическую вероятностную модель, в статистической структуре которой зашифрованы некие неклассические ограничения (теорию со скрытыми параметрами). Для одиночной квантовой системы такая точка зрения возможна, но до сих пор не оказалась плодотворной. При переходе же к составным системам она приводит к неустранимым противоречиям с физическими принципами локальности и причинности (см. далее § 2.5).

Наиболее простым, но важным примером является *q-бит* — двухуровневая квантовая система, $\dim \mathcal{H} = 2$. Будем использовать канонический базис: $\begin{bmatrix}1\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\end{bmatrix}$. Удобно ввести базис Паули в вещественном пространстве эрмитовых матриц:

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

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

$$S = \tfrac{1}{2}(I + a_x\sigma_x + a_y\sigma_y + a_z\sigma_z) = \tfrac{1}{2}\begin{bmatrix}1 + a_z & a_x - ia_y\\ a_x + ia_y & 1 - a_z\end{bmatrix}. \qquad (2.2)$$

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

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

Таким образом, $\mathcal{S}(\mathcal{H})$ как выпуклое множество изоморфно единичному шару в $\mathbb{R}^3$. Чистые состояния характеризуются условием $a_x^2 + a_y^2 + a_z^2 = 1$ и составляют *сферу Блоха*. Вводя углы Эйлера θ и φ так, что $a_z = \cos\theta$ и $a_x + ia_y = \sin\theta e^{i\varphi}$, имеем $S = |\psi(\vec{a})\rangle\langle\psi(\vec{a})|$, где

$$|\psi(\vec{a})\rangle = \begin{bmatrix}\cos(\theta/2)\, e^{-i\varphi/2}\\ \sin(\theta/2)\, e^{i\varphi/2}\end{bmatrix}. \qquad (2.3)$$

В квантовых системах со спином $1/2$ вектор $\psi(\vec{a})$ описывает ансамбль (пучок частиц) со спином в направлении $\vec{a}$. Хаотическим

называется смешанное состояние с $a_x=a_y=a_z=0$ (все направления спинов равновероятны), описываемое оператором плотности $S=I/2$.

Другим важным примером двухуровневой системы является поляризация поперечного фотона.

§ 2.3. Квантовые наблюдаемые

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

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

$$M_x\geqslant 0,\quad \sum_{x\in\mathcal{X}} M_x=I, \tag{2.4}$$

и

$$\mu_S(x)=\operatorname{Tr} SM_x. \tag{2.5}$$

Семейство, обладающее свойствами (2.4), называется *разложением единицы* в $\mathcal{H}$.

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

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

т. е. линейная оболочка $\operatorname{Lin} \mathcal{S}(\mathcal{H})$ совпадает с вещественным линейным пространством $\mathcal{B}(\mathcal{H})_h$ всех эрмитовых операторов в $\mathcal{H}$. Пусть $f(S)$ — аффинная функция на $\mathcal{S}(\mathcal{H})$, продолжим ее на операторы A, полагая $f(A) = t_1 f(S_1) - t_2 f(S_2)$. Надо проверить, что благодаря аффинности, такое продолжение однозначно и вещественно линейно (задача 7). Далее, f однозначно и комплексно линейно продолжается на алгебру всех операторов $\mathcal{B}(\mathcal{H})$.

Следовательно, если $A = [a_{ij}]$ в некотором базисе, то $f(A) = \sum_{ij} m_{ij} a_{ij} = \operatorname{Tr} AM$, где M — некоторый оператор. Применяя это к функциям $\mu_S(x)$, получаем $\mu_S(x) = \operatorname{Tr} SM_x$. Поскольку $\operatorname{Tr} SM_x$ — распределение вероятностей для всех S, то отсюда следуют соотношения (2.4).

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

В стандартных учебниках по квантовой механике под наблюдаемой понимают ортогональное разложение единицы, т. е. разложение, для которого

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

Задача 8. Ортогональное разложение единицы характеризуется тем, что все M_x — проекторы: $M_x = M_x^2$.

Будем называть *стандартной наблюдаемой* ортогональное разложение единицы в $\mathcal{H}$. Пусть $x \in \mathcal{X}$ вещественные числа. Всякая такая наблюдаемая однозначно определяется эрмитовым оператором

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

который также называется (вещественной) наблюдаемой. Среднее значение такой наблюдаемой дается обычной формулой Борна

$$\sum x \mu_s^E(x) = \operatorname{Tr} SA.$$

Чтобы прояснить значение неортогональных разложений единицы, рассмотрим ортонормированный базис $\{|\omega\rangle\}$ в $\mathcal{H}$ и операторы, диагональные в этом базисе. Оператор плотности

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

задает классическое состояние — распределение вероятностей на «фазовом пространстве» $\Omega = \{\omega\}$. Эрмитов оператор $A = \sum_{\omega} x_\omega |\omega\rangle\langle\omega|$ может быть записан в виде

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

Классическим наблюдаемым A соответствуют случайные величины x_ω на Ω. Проекторам E_x отвечают индикаторы подмножеств Ω, а ортогональному разложению единицы — разбиение пространства Ω.

Рассмотрим неортогональное разложение единицы с элементами $M_x = \sum_{\omega} M(x|\omega)|\omega\rangle\langle\omega|$. Тогда собственные числа удовлетворяют условиям $0 \leqslant M(x|\omega) \leqslant 1$ и $\sum_x M(x|\omega) \equiv 1$, т. е. определяют переходные вероятности из Ω в $\mathcal{X}$. Таким образом, в классическом случае разложения единицы описывают рандомизованные (нечеткие) наблюдаемые, задающие только вероятности исходов x в каждой точке ω фазового пространства. Для «четких» наблюдаемых, удовлетворяющих условию $M_x^2 = M_x$, эти вероятности принимают значения 0 или 1. Множество всех переходных вероятностей является выпуклым.

Задача 9. Показать, что крайние точки этого множества соответствуют в точности ортогональным разложениям единицы (см. [4]).

Однако такая простая картина имеет место только в классике. Рассмотрим следующий пример.

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

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

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

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

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

Задача 10. Система $\{|\psi_j\rangle\}$ является переполненной тогда и только тогда, когда

1) система полна, т. е. $\{|\psi_j\rangle\}^\perp=\{0\}$;

2) матрица $P=[\langle\psi_j\,|\,\psi_k\rangle]$ идемпотентна, т. е. $P=P^2$.

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

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

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

С каждой переполненной системой связано разложение единицы $M_j=|\psi_j\rangle\langle\psi_j|$. Это разложение единицы является крайней точкой выпуклого множества всех разложений единицы тогда и только тогда, когда операторы M_j линейно независимы (см. § 4.3). Переполненные неортогональные системы не имеют аналога в классической статистике.

Математический смысл неортогональных разложений единицы проясняет теорема Наймарка.

Теорема 5. *Пусть $\{M_x\}_{x\in\mathcal{X}}$ — разложение единицы в гильбертовом пространстве $\mathcal{H}$, $\dim\mathcal{H}=d$, $|\mathcal{X}|=n$. Тогда существует гильбертово пространство $\widetilde{\mathcal{H}}$, $\dim\widetilde{\mathcal{H}}\leqslant n\cdot d$, изометрический оператор $V\colon\mathcal{H}\to\widetilde{\mathcal{H}}$ и ортогональное разложение единицы $\{E_x\}$ в $\widetilde{\mathcal{H}}$, такие, что*

$$M_x=V^*E_xV.$$

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

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

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

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

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

в которой определим псевдоскалярное произведение формулой

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

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

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

Задача 11. После факторизации эта формула корректно определяет оператор V из $\mathcal{H}$ в $\widetilde{\mathcal{H}}$.

Этот оператор изометричен так как

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

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

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

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

§ 2.4. Составные квантовые системы

Своеобразие квантовой теории информации и возможности квантового компьютинга в значительной мере обусловлены необычными свойствами составных квантовых систем. Пусть $\mathcal{H}_i$, $i = 1, 2$, гильбертовы пространства двух квантовых систем со скалярными

произведениями $\langle\cdot\,|\,\cdot\rangle_i$. Их совокупность описывается тензорным произведением гильбертовых пространств, которое строится следующим образом. Рассмотрим векторное пространство $\mathcal{L}$ конечных формальных линейных комбинаций $\sum_j c_j \psi_1^j \times \psi_2^j$. Введем псевдоскалярное произведение на $\mathcal{L}$, полагая на порождающих элементах

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

и далее продолжая по линейности на $\mathcal{L}$. Полученное псевдоскалярное произведение будет вырожденным на подпространстве $\mathcal{L}_0$ (заведомо содержащем все элементы вида $-c\varphi_1 \times \psi_2 - c'\varphi_1' \times \psi_2 + (c\varphi_1 + c'\varphi_1') \times \psi_2$). Тензорным произведением $\mathcal{H}_1 \otimes \mathcal{H}_2$ гильбертовых пространств называется факторпространство $\mathcal{L}/\mathcal{L}_0 = \mathcal{H}$ со скалярным произведением, порожденным формой $\langle\cdot\,|\,\cdot\rangle$. Образы порождающих элементов $\psi_1 \times \psi_2$ при этой факторизации обозначаются $\psi_1 \otimes \psi_2$.

Задача 12. Пусть $\{e_1^j\}, \{e_2^k\}$ — ортонормированные базисы в $\mathcal{H}_1, \mathcal{H}_2$, тогда $\{e_1^j \otimes e_2^k\}$ — ортонормированный базис в $\mathcal{H}_1 \otimes \mathcal{H}_2$ и $\dim \mathcal{H} = \dim \mathcal{H}_1 \cdot \dim \mathcal{H}_2$.

Таким образом, реализуя $\mathcal{H}_{1,2}$ как пространства ℓ^2 числовых последовательностей $\{c_j^1\}, \{c_k^2\}$, получим реализацию $\mathcal{H}$ в виде пространства матриц $[c_{jk}]$ с величиной $\sum_{j,k} |c_{jk}|^2$ в качестве нормы. Заметим, что всякий вектор $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ однозначно записывается в виде

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

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

Для операторов X_j в пространствах $\mathcal{H}_j$ зададим их тензорное произведение в пространстве $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, полагая

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

и продолжая по линейности.

Задача 13. Если S_j — операторы плотности в $\mathcal{H}_1$, то $S_1 \otimes S_2$ — оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2$.

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

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

Задача 14. Определение корректно (не зависит от выбора ортонормированного базиса $\{e_2^k\}$). Если $T = T_1 \otimes T_2$, то $\mathrm{Tr}_{\mathcal{H}_2}(T_1 \otimes T_2) = (\mathrm{Tr}\, T_2) T_1$.

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

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

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

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

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

где $\mathcal{H}_0 = \ell^2$ — гильбертово пространство размерности d_0, причем $\mathcal{H}$ отождествляется с первым слагаемым в прямой сумме, или с подпространством $\mathcal{H} \otimes |\psi_0\rangle\langle\psi_0|$, где

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

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

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

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

В классической статистике рандомизация, т. е. добавление «рулетки», хотя и может оказаться полезным приемом (например, в теории игр), никогда не увеличивает информации о состоянии наблюдаемой системы. В главе 4 мы покажем, что в квантовой

механике это уже не так: парадоксальным образом, квантовая рандомизация позволяет извлекать больше информации о наблюдаемой системе, нежели содержится в стандартных наблюдаемых, не использующих вспомогательной системы.

§ 2.5. Парадокс ЭПР. Неравенство Белла

Ключевой пример необычного (с классической точки зрения) поведения составной квантовой системы рассмотрели Эйнштейн, Подольский и Розен (ЭПР) в 1935 г. В более отчетливой форме, использующей спиновые степени свободы, его представил Бом в 50-х, и полную ясность внес Белл в 60-х годах. Рассмотрим составную систему из двух q-битов, например, две частицы со спином $1/2$, каждая из которых описывается гильбертовым пространством $\mathcal{H}$ с $\dim \mathcal{H} = 2$. В начальный момент частицы взаимодействуют таким образом, что конечное состояние их спинов, называемое состоянием Белла, описывается вектором

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

где векторы

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

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

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

а в квантовых вычислениях предпочитают обозначение

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

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

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

1) как и в классической механике, (чистое) состояние описывает внутренние свойства системы. Тогда приходится допустить мгновенное дальнодействие, противоречащее принципу локальности;

2) вектор состояния — это лишь выражение информационного содержания процедуры приготовления системы. При таком понимании никакого противоречия с локальностью или причинностью не возникает, и обстоятельство, что вторая частица «мгновенно» оказывается в состоянии с противоположным спином, не более удивительно, чем то, что у наугад выбранной пары носков оказывается одинаковый цвет.

Однако внимательное рассмотрение этого мысленного эксперимента приводит к более глубокому и неожиданному выводу, на который обратил внимание Белл: если пытаться описывать корреляции измерений спинов двух частиц классически и в соответствии с принципом локальности, то оказывается невозможным достичь такого характера и уровня коррелированности, который соответствует предсказаниям квантовой механики. Более того, этот уровень коррелированности может быть количественно сформулирован и проверен экспериментально. Дадим точную формулировку. Пусть вектор $\vec{a} = (a_x, a_y, a_z)$ задает некоторое направление, тогда $\sigma(\vec{a}) = a_x\sigma_x + a_y\sigma_y + a_z\sigma_z$ — наблюдаемая спина в направлении $\vec{a}$ (с точностью до множителя $\hbar/2$.) Оператор $\sigma(\vec{a})$ имеет собственные значения ± 1 (спин вдоль и против направления $\vec{a}$). Таким образом

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

Напомним, что $\vec{a}$ имеет углы Эйлера (θ, φ), при этом вектор (2.3) отвечает чистому состоянию со спином в направлении $\vec{a}$. Соответствующий оператор плотности равен

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

Рассмотрим эксперимент, в котором производятся совместные измерения наблюдаемой $\sigma(\vec{a})$ для одной системы и $\sigma(\vec{b})$ — для другой (см. рис. 2).

Задача 15. Для состояния Белла двух q-битов корреляция спинов дается формулой

$$\langle\psi|\sigma(\vec{a})\otimes\sigma(\vec{b})|\psi\rangle = -\vec{a}\cdot\vec{b}. \tag{2.6}$$

Оказывается, что такая корреляция не может быть смоделирована никакой классической моделью составной системы, удовлетворяющей принципу локальности. Это вытекает из следующего неравенства Белла — Клаузера — Хорна — Шимони. Пусть X_j, Y_k, $j, k = 1, 2$, — случайные величины на произвольном вероятностном пространстве Ω, такие что $|X_j| \leqslant 1$, $|Y_k| \leqslant 1$. Тогда для любого распределения вероятностей на Ω корреляции этих величин удовлетворяют неравенству

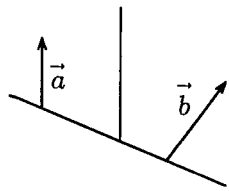


Рис. 2. Направления спинов

$$|\mathbf{E}X_1Y_1 + \mathbf{E}X_1Y_2 + \mathbf{E}X_2Y_1 - \mathbf{E}X_2Y_2| \leqslant 2, \tag{2.7}$$

где $\mathbf{E}$ — соответствующее математическое ожидание.

Доказательство получается усреднением элементарного неравенства

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

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

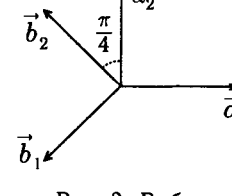


Рис. 3. Выбор векторов a_j и b_k

Вернемся теперь к системе из двух q-битов и рассмотрим четыре эксперимента, когда в первом q-бите измеряется наблюдаемая

спина $\sigma(\vec{a}_j)$, $j=1,2$, а во втором $\sigma(\vec{b}_k)$, $k=1,2$, где направления $\vec{a}_j$, $\vec{b}_k$, $j, k=1,2$, образуют конфигурацию, изображенную на рис. 3.

При этом система приготавливается в одном и том же состоянии Белла. Подстановка соответствующих значений корреляций из формулы (2.6) в левую часть формулы (2.7) дает значение $2\sqrt{2}$, нарушающее неравенство. Отсюда следует, что либо квантовая механика дает неправильные выражения для корреляций, либо для данной составной системы не существует классического вероятностного описания, удовлетворяющего условию локальности. После первого эксперимента (Аспек, 1981–1982 гг.) был проделан целый ряд аналогичных экспериментов по измерению ЭПР-корреляций, результаты которых с определенностью свидетельствуют в пользу квантовой механики.

ГЛАВА 3

ПРИМЕНЕНИЯ СЦЕПЛЕННЫХ СОСТОЯНИЙ

§ 3.1. Квантовое состояние как информационный ресурс

В этом параграфе нам потребуются элементарные сведения об эволюциях квантовой системы. В дальнейшем, в главе 6 этот вопрос будет рассмотрен углубленно и с общих позиций теории открытых квантовых систем. Математически настроенному читателю мы советуем перейти к главе 4 и вернуться к этому материалу после главы 6. Пока же достаточно знать следующее:

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

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

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

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

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

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

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

Таким образом, для фиксированного измерения мы получаем обычный канал связи. Это дает возможность поставить вопрос о максимальном количестве классической информации, которое может быть передано по данному квантовому каналу связи и о его пропускной способности. Этот вопрос будет детально рассмотрен в главе 5. Отметим здесь лишь один факт, имеющий принципиальное значение:

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

Рассмотрим теперь следующий вопрос. Нелокальный, с классической точки зрения, характер ЭПР-корреляций наводит на мысль попытаться использовать их для мгновенной передачи информации. Покажем, что этого невозможно достичь, находясь в рамках квантовой механики (с точки зрения которой ЭПР-корреляции не противоречат локальности). Рассмотрим две квантовые системы A

и B, в пространствах $\mathcal{H}_A$ и $\mathcal{H}_B$ соответственно, которые находятся в сцепленном состоянии S_{AB}. В случае, представляющем интерес, системы пространственно разделены, хотя формально это ни в чем не выражается. Система A получает классическую информацию, содержащуюся в значениях параметра x, которая может быть использована для выполнения произвольных унитарных операций U_x в пространстве $\mathcal{H}_A$. При этом состояние системы AB переходит в $S_x = (U_x \otimes I_B) S_{AB} (U_x \otimes I_B)^*$, таким образом, классическая информация записывается в квантовом состоянии составной системы. В свою очередь, над системой B может быть произведено произвольное измерение, описываемое ортонормированным базисом $|e_y\rangle$ в $\mathcal{H}_B$. Легко видеть, что результирующая переходная вероятность (3.2) не зависит от x, а значит количество передаваемой информации в самом деле равно нулю.

§ 3.2. Сверхплотное кодирование

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

Рассмотрим системы A и B, каждая из которых представляет собой q-бит, между которыми имеется идеальный квантовый канал связи. Из того что было сказано выше в § 3.1, 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», предложенный Виснером в 1992 г., имеет в своей основе простой математический факт: базис 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 $1/\sqrt{2}$) может быть получен из одного вектора действием «локальных» унитарных операторов, т. е. операторов, действующих нетривиально только в пространстве q-бита A, например

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

Таким образом, если AB изначально находится в состоянии $|e_+\rangle$, A может закодировать 2 бита классической информации в 4 состояния базиса Белла, производя только локальные операции, а затем (физически) послать свой q-бит B по идеальному квантовому каналу. Тогда, производя измерение в базисе Белла, B получает 2 бита классической информации. Конструкции протоколов сверхплотного кодирования и телепортации допускают обобщение на случай пространства произвольной конечной размерности (ср. далее в главе 8).

§ 3.3. Квантовая телепортация

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

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

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

универсального прибора, который бы размножал произвольное квантовое состояние.

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

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

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

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

2) C посылает A произвольное чистое состояние

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

Совокупность трех систем CAB описывается состоянием

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

3) Затем

а) A производит некоторое обратимое преобразование состояния системы CA;

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

---

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

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

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

Производимые преобразования являются характерными примерами логических операций, используемых в квантовом компьютинге. На 3-м шаге над системой CA производится операция CNOT (контролируемое «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 четырехмерном 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 в системе CA, проецирующее на один из четырех базисных векторов $|00\rangle, \ldots, |11\rangle$. Результат измерения посылается от A к B по классическому (идеальному) каналу

связи. В зависимости от полученного результата B применяет к своему состоянию один из унитарных операторов

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

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

Возможность телепортации состояния поляризации фотона была продемонстрирована экспериментально Целлингером в 1997 г. Одна из технически наиболее сложных проблем в таком эксперименте — совместные манипуляции с состояниями фотонов C и A.

§ 3.4. Квантовые алгоритмы

Идея квантового компьютера была предложена Фейнманом для моделирования квантовомеханических систем. Впоследствии был поставлен вопрос: не может ли квантовый компьютер решать какие-либо задачи более эффективно, чем классический. Простейшие, но довольно искусственные примеры таких задач рассмотрели Дейч и Джоза. Их усовершенствованием является алгоритм Саймона, который лежит в основе и алгоритма Шора, эффективно решающего важную и практически интересную (по крайней мере, с точки зрения криптографии) задачу разложения большого натурального числа на простые множители.

3.4.1. Алгоритм Саймона. Обозначим $B = \{0, 1\}$, $B^n = B^{\times n}$. Пусть задана функция $f\colon B^n \to B^n$. Известно, что функция f является периодической, т. е. $f(x) = f(y) \iff y = x \oplus \xi$, где $\xi \in B^n$ — двоичный (булев) вектор. Здесь $\oplus$ обозначает покомпонентное двоичное сложение векторов.

Требуется найти период ξ за наименьшее возможное число шагов (принимая за шаг каждый акт вычисления функции f). Классическое решение задачи сводится к перебору и требует число шагов $O(2^{n/2})$, растущее экспоненциально с n. (После вычисления s значений функции f, сравнивая значения в двух точках, мы можем исключить не более $s(s-1)/2$ из 2^n значений ξ, откуда $s \sim 2^{n/2}$.)

Квантовый алгоритм требует всего $O(n)$ шагов, если считать за шаг квантовое вычисление функции f. Для описания квантового алгоритма нам понадобится n-мерное обобщение операции Адамара $H_n = \underbrace{H \otimes \ldots \otimes H}_{n}$. Рассмотрим квантовый регистр — физическую

систему из n q-битов; информация будет задаваться состоянием этой системы. Если x — набор нулей или единиц длины n, то

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

где $x \cdot y$ — скалярное произведение векторов $x \in B^n$, $y \in B^n$ по модулю 2, поскольку

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

Алгоритм Саймона состоит из следующих шагов:

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

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

В результате получается суперпозиция всевозможных базисных состояний с одинаковыми коэффициентами.

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

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

Предполагается, что такой унитарный оператор дан «свыше» (поэтому его принято называть «оракулом»). Отметим, что в алгоритме Шора соответствующее вычисление описывается эффективно. В принципе, он может быть составлен из некоторых элементарных операций, если известно, как сама f составлена из элементарных логических операций. Здесь $|z\rangle$ состояние вспомогательного регистра, который введен, чтобы сделать операцию вычисления функции обратимой. Если исходно этот регистр находится в основном состоянии, то

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

3) Вновь применяя операцию Адамара,получаем состояние

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

4) Измеряя оба регистра, получаем состояние $|y\rangle|f(x)\rangle$ с вероятностью

$$\left(\frac{1}{2^n}\right)^2 [(-1)^{x\cdot y} + (-1)^{(x_0+\xi)\cdot y}]^2,$$

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

Таким образом, получается случайный равномерно распределенный вектор $y(\omega)$ из булевой «гиперплоскости» $y\cdot\xi = 0$.

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

$$\mathbf{P}\{y_1(\omega), \ldots, y_{n-1}(\omega) \text{ линейно независимы}\} \geqslant e^{-1}.$$

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

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

и

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

Следовательно

$$\mathbf{P}\{y_1(\omega), \ldots, y_{n-1}(\omega) \text{ линейно независимы}\} =$$
$$= \left(1 - \frac{1}{2^{n-1}}\right)\ldots\left(1 - \frac{1}{2}\right) = \exp\left[\sum_{k=1}^{n-1} \log\left(1 - \frac{1}{2^k}\right)\right] \geqslant$$
$$\geqslant \exp\left[-\sum_{k=1}^{n-1} \frac{1}{2^k}\right] \geqslant e^{-1}.$$

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

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

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

Квантовый алгоритм Шора имеет полиномиальную сложность $O(n^2 \log n \log\log n)$. Представление о его эффективности дает следующая грубая оценка: задача факторизации числа $N \sim 2^{800}$ не решаема за разумное время на классическом компьютере, тогда как применение квантового алгоритма при частоте переключений 1 Мгц потребовало бы пару дней. Алгоритм использует сведение задачи факторизации к нахождению периода функции $f(x) = a^x (mod\ N)$, где a выбирается случайным образом. Можно показать, что в большинстве случаев период r является четным и число $a^{r/2} \pm 1$ имеет общий множитель с N, который находится с помощью классического алгоритма Евклида. Алгоритм Шора включает детальное описание эффективного выполнения операции U_f. Нахождение периода $f(x)$ использует квантовую модификацию быстрого преобразования Фурье (роль которого в гораздо более простой задаче Саймона выполняло преобразование Адамара H_n). Подробнее об алгоритме Шора и квантовых вычислениях см. в [13, 16, 2].

[1)] См. *Shor P.* Introduction to quantum algorithms. — LANL Report no. quant-ph/0005003.

3.4.3. Алгоритм Гровера. Этот алгоритм решает задачу поиска. Более точно, предполагается, что задана булева функция F: $B^n \to B$, такая что $F(x_0)=1$, $F(x)=0$, $x \neq x_0$. Требуется найти x_0, причем вычисление значения функции F в любой заданной точке принимается за один шаг. Классический алгоритм сводится к перебору значений x и проверки для них равенства $F(x)=1$, что в наименее благоприятном случае требует $N \sim 2^n$ шагов. Квантовый алгоритм Гровера позволяет решить задачу за $\approx \sqrt{N} = 2^{n/2}$ шагов.

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

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

Алгоритм состоит из следующих шагов:

1) К основному состоянию применяется операция Адамара

$$|0\ldots0\rangle \xrightarrow{H_n} \frac{1}{\sqrt{N}} \sum_x |x\rangle = |\psi\left(\arcsin \frac{1}{\sqrt{N}}\right)\rangle,$$

где введено обозначение

$$|\psi(\theta)\rangle = \sin\theta|x_0\rangle + \frac{\cos\theta}{\sqrt{N-1}} \sum_{x \neq x_0} |x\rangle.$$

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

Задача 16. Проверьте, что

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

где $\sin\varphi = 2\sqrt{N-1}/N$, т. е. U осуществляет поворот на угол φ в плоскости, натянутой на вектор $|x_0\rangle$ и вектор $\frac{1}{\sqrt{N-1}} \sum_{x \neq x_0} |x\rangle$.

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

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

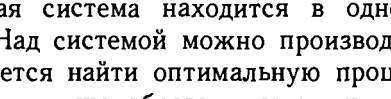
ГЛАВА 4

ОПТИМАЛЬНОЕ РАЗЛИЧЕНИЕ КВАНТОВЫХ СОСТОЯНИЙ

§ 4.1. Постановка задачи

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

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



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

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

а средняя вероятность ошибки равна $1 - \mathcal{P}\{M\}$, и задача состоит в ее минимизации, или же в максимизации $\mathcal{P}\{M\}$. В статистике

применяется и минимаксный критерий, когда минимизируется $\max_j p_M(j \mid j)$, но мы его не будем здесь затрагивать.

Другой важный критерий, который мы рассмотрим позже, — шенноновская информация. Согласно формуле (1.7) количество взаимной информации между входом J (j — номер входного состояния) и выходом K (k — номер решения) дается формулой

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

$$= -\sum_k \sum_j p_M(k \mid j)\pi_j \log \sum_l p_M(k \mid l)\pi_l + \sum_j \pi_j \sum_k p_M(k \mid j) \log p_M(k \mid j) =$$

$$= \sum_j \pi_j \sum_k p_M(k \mid j) \log \left[\frac{p_M(k \mid j)}{\sum_l p_M(k \mid l)\pi_l} \right].$$

§ 4.2. Различение по максимуму правдоподобия

Будем максимизировать вероятность правильного решения

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

Множество наблюдаемых, по которым ведется оптимизация

$$\mathfrak{M}_n = \left\{ M = \{M_k\}_{k=1,\ldots,n} \;\middle|\; M_k \geqslant 0, \sum_{k=1}^{n} M_k = I \right\}$$

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

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

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

ТЕОРЕМА 6. *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 такой эрмитов оператор Λ^0, что*

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

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

При этом имеет место соотношение двойственности:

$$\max\{\mathcal{P}\{M\} \mid M \in \mathfrak{M}_n\} = \min\{\operatorname{Tr}\Lambda \mid \Lambda \geqslant W_k,\ k = 1, \ldots, n\}. \quad (4.1)$$

Доказательство. Докажем достаточность условий теоремы.

Пусть наблюдаемая M^0 удовлетворяет этим условиям, $M \in \mathfrak{M}_n$ — произвольная наблюдаемая, тогда в силу 1) и 2) имеем

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

Здесь был использован следующий простой факт:

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

Докажем необходимость условий теоремы.

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

$$\operatorname{Tr}\sum_k W_k X_k - \operatorname{Tr}\Lambda(\sum_k X_k^2 - I), \quad (4.2)$$

где Λ — эрмитов оператор, по всевозможным наборам эрмитовых операторов X_k. Пусть X_k^0 оптимальный набор, положим $X_k = X_k^0 + \varepsilon Y_k$, и рассмотрим (4.2) как функцию от ε. Рассматривая коэффициенты при ε и ε^2, получаем условия

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

для произвольных эрмитовых Y_k, т. е.

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

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

Задача 18. Доказать, что операторный множитель Лагранжа Λ является единственным решением двойственной задачи в правой части (4.1).

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

ПРИМЕР 1. Пусть операторы W_k (пропорциональные S_k) коммутируют, тогда существует общий ортонормированный базис, где они все диагонализуются, т. е.

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

Тогда можно взять

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

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

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

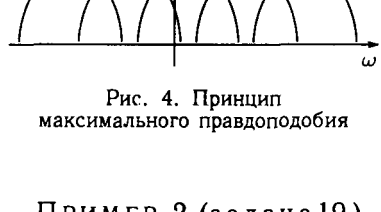


Рис. 4. Принцип максимального правдоподобия

ПРИМЕР 2 (задача 19). *Различение двух квантовых состояний.*. Произвольная наблюдаемая с двумя значениями имеет вид $M = \{M_0, M_1\}$, $M_{0,1} \geqslant 0$, $M_1 = I - M_0$, причем стандартные наблюдаемые характеризуются условием $M_0^2 = M_0$, которое в точности соответствует крайним точкам «некоммутативного отрезка» $\mathfrak{M}_2 = \{0 \leqslant M_0 \leqslant I\}$ (задача 20). Таким образом, для различения двух состояний достаточно стандартных наблюдаемых.

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

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

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

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

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

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

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

Пусть $S_0 = |\psi_0\rangle\langle\psi_0|$, $S_1 = |\psi_1\rangle\langle\psi_1|$. В этом случае оптимум дается ортонормированным базисом $\{|e_0\rangle, |e_1\rangle\}$, так что $M_0 = |e_0\rangle\langle e_0|$, $M_1 = |e_1\rangle\langle e_1|$. Вектор $|e_0\rangle$ отвечает положительному собственному числу λ_0 оператора $\pi_0|\psi_0\rangle\langle\psi_0| - \pi_1|\psi_1\rangle\langle\psi_1|$, причем $\max \mathcal{P}\{M\} = \pi_1 + \lambda_0$. Диагонализуя оператор $\pi_0|\psi_0\rangle\langle\psi_0| - \pi_1|\psi_1\rangle\langle\psi_1|$, можно дать явное решение задачи (см. [3]). Пусть для простоты $\pi_0 = \pi_1 = 1/2$, тогда оптимальный базис расположен симметрично по отношению к $|\psi_0\rangle, |\psi_1\rangle$ (см. рис. 5) и

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

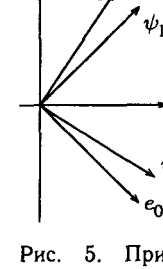


Рис. 5. Принцип максимального правдоподобия

Задача 21. Показать, что для различения n чистых состояний с линейно независимыми векторами $|\psi_j\rangle$, $j = 1, \ldots, n$, достаточно стандартных наблюдаемых. В этом случае оптимальная наблюдаемая дается векторами некоторой ортонормированной системы $|e_j\rangle$, $j = 1, \ldots, n$ (см. [3]).

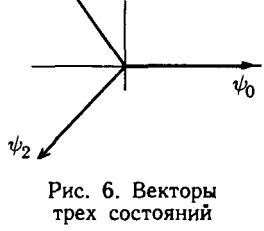


Рис. 6. Векторы трех состояний

ПРИМЕР 3. На плоскости (рассматриваемой как вещественное подпространство двумерного унитарного пространства) рассмотрим «равноугольную» конфигурацию трех векторов (см. рис. 6)

$$|\psi_j\rangle = \begin{bmatrix} \cos\frac{2j\pi}{3} \\ \sin\frac{2j\pi}{3} \end{bmatrix}, \quad j = 0, 1, 2. \qquad (4.3)$$

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

Имеем

$$S_j = \begin{bmatrix} \cos^2\frac{2j\pi}{3} & \cos\frac{2j\pi}{3}\sin\frac{2j\pi}{3} \\ \cos\frac{2j\pi}{3}\sin\frac{2j\pi}{3} & \sin^2\frac{2j\pi}{3} \end{bmatrix} = \\ = \frac{1}{2}\left(I + \begin{bmatrix} \cos\frac{4j\pi}{3} & \sin\frac{4j\pi}{3} \\ \sin\frac{4j\pi}{3} & \cos\frac{4j\pi}{3} \end{bmatrix}\right). \quad (4.4)$$

Поскольку $\sum\limits_{j=0}^{2} e^{i\frac{4j\pi}{3}} = 0$, то $\sum\limits_{j=0}^{2} S_j = \frac{3}{2}I$, т. е. $M_k^0 = 2S_k/3$ является разложением единицы.

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

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

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

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

Таким образом условие 1) также выполнено.

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

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

Таким образом, использование в квантовой статистике неортогональных разложений единицы в качестве наблюдаемых (т. е. использование квантовой рандомизации — дополнительной независимой квантовой системы в фиксированном состоянии) может приводить к выигрышу при различении состояний исходной системы!

Подчеркнем, что в классическом случае никакая рандомизация не может улучшить качество процедуры различения состояний.

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

§ 4.3. Максимум информации

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

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

$$\mathcal{J}\{M\} = \sum_j \pi_j \sum_k p_M(k \mid j) \Big[\log p_M(k \mid j) - \log \sum_l p_M(k \mid l)\pi_l\Big], \quad (4.5)$$

где π_j — априорные вероятности состояний.

Лемма 5. *Функция $\mathcal{J}\{M\}$ выпукла на $\mathfrak{M}$, т. е.*

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

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

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

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

Доказательство. Рассмотрим множество переходных вероятностей $p(k \mid j) \geqslant 0$, $\sum_k p(k \mid j) = 1$. Имеем

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

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

$$\sum_j p(k \mid j)\pi_j \Big[\log p(k \mid j) - \log\Big(\sum_l p(k \mid l)\pi_l\Big)\Big],$$

поскольку количество информации является суммой слагаемых вида

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

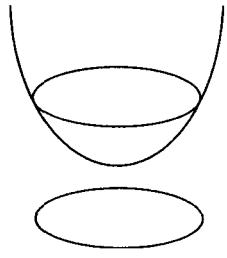


Рис. 7. Максимум выпуклой функции

Дифференцируя по x_j, получаем

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

(здесь для простоты log — натуральный логарифм) и

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

Согласно неравенству Коши — Буняковского имеем

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

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

Задача 22. Максимум непрерывной выпуклой функции на компактном выпуклом множестве достигается в крайней точке этого множества (см. рис. 7).

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

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

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

$$p_M(j_1 \mid i)\Big[\log p_M(j_1 \mid i) - \log \sum_l p_M(j_1 \mid l)\pi_l\Big] +$$

$$+ p_M(j_2 \mid i)\Big[\log p_M(j_2 \mid i) - \log \sum_l p_M(j_2 \mid l)\pi_l\Big] \geqslant$$

$$\geqslant \underbrace{\{p_M(j_1 \mid i) + p_M(j_2 \mid i)\}}_{\text{Отвечает одному исходу в } M'}\Big[\log\Big(p_M(j_1 \mid i) + p_M(j_2 \mid i)\Big) -$$

$$- \log \sum_l \pi_l \Big(p_M(j_1 \mid l) + p_M(j_2 \mid l)\Big)\Big].$$

Введем множитель $1/2$:

$$\tfrac{1}{2}p_M(j_1 \mid i)[\ldots] + \tfrac{1}{2}p_M(j_2 \mid i)[\ldots] \geqslant$$
$$\geqslant \frac{p_M(j_1 \mid i) + p_M(j_2 \mid i)}{2}\left\{\log\frac{[\ldots]}{2} - \log\sum\frac{[\ldots]}{2}\right\}$$

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

ТЕОРЕМА 7[1]. *Пусть дан набор квантовых состояний $S_1, \ldots S_m$ с определенными вероятностями $\pi_1, \ldots, \pi_m$, тогда существует наблюдаемая M^0, для которой $\max_{\mathfrak{M}} \mathcal{J}\{M\} = \mathcal{J}\{M^0\}$, и такая, что ее компоненты — линейно независимые операторы ранга 1, т. е. $M_j^0 = |\psi_j\rangle\langle\psi_j|_{j=1,\ldots,n}$, и число компонент $n \leqslant d^2$, где $d = \dim \mathcal{H}$. Если все операторы S_j имеют вещественные матрицы в некотором базисе, то $n \leqslant d(d+1)/2$.*

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

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

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

Согласно лемме 6 и задаче 2, можно считать, что M^0 крайняя точка, имеющая компоненты $M_j^0 = |\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|$ $(j = 1, \ldots, n)$. Отсюда следует, что операторы $|\psi_j\rangle\langle\psi_j|$ линейно независимы. В самом деле, предположим, что

$$\sum_j c_j |\psi_j\rangle\langle\psi_j| = 0. \tag{4.6}$$

Для достаточно малого $\varepsilon > 0$ имеем

$$M_j^{\pm} = (1 \pm \varepsilon c_j) M_j^0 \geqslant 0, \qquad j = 1, \ldots, n.$$

[1] *Davies E. B.* Information and quantum measurement // IEEE Trans. Inform. Theory. — 1978. — V. 24, № . 6. — P. 596–599.

Тогда $M^{\pm}$ являются наблюдаемыми и $M^0 = \frac{1}{2}M^+ + \frac{1}{2}M^-$. Следовательно $M_j^0 = M_j^+ = M_j^-$, поскольку M^0 экстремальна. Итак, из равенства (4.6) следует, что $c_j = 0$, т. е. компоненты наблюдаемой M^0 линейно независимы. Но максимальное число линейно независимых эрмитовых операторов в d-мерном унитарном пространстве равно d^2 ($d(d+1)/2$ в вещественном случае).

Явное решение возможно в случаях, когда есть некоторая симметрия.

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

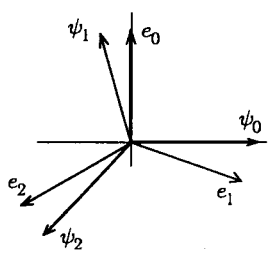


Рис. 8. Информационный оптимум для трех равноугольных состояний

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

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

$$\max_M \mathcal{J}\{M\} = 1 - h\left(\frac{1+\sqrt{1-\varepsilon^2}}{2}\right) \tag{4.7}$$

и максимум информации достигается на базисе, расположенном симметрично по отношению к векторам ψ_0, ψ_1 (рис. 5), оптимальном по критерию максимального правдоподобия.

ПРИМЕР 2. Случай трех равновероятных «равноугольных» чистых состояний (4.4) с углами $2\pi/3$ между направлениями спинов. Согласно теореме, $m \leqslant d(d+1)/2 = 3$. Используя симметрию задачи, можно доказать, что информационно-оптимальная наблюдаемая имеет вид $M_k = \frac{2}{3}|e_k\rangle\langle e_k|_{k=0,1,2}$, где $e_k \perp \psi_k$ (см. рис. 8; задача 23). Таким образом, она не совпадает с наблюдаемой, оптимальной по максимуму правдоподобия, для которой $e_k = \psi_k$. Более того, можно показать, что последняя является наихудшей с точки зрения информационного критерия. Максимум информации по всевозможным наблюдаемым: $\max_M \mathcal{J}\{M\} \approx 0,585$, тогда как $\max_{M\text{-стандартные}} \mathcal{J}(M) \approx 0,428$.

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

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

[1] *Shor P. W.* On the number of elements needed in a POVM attaining the accessible information // LANL Report no. quant-ph/0009077.

ГЛАВА 5

КЛАССИЧЕСКАЯ ПРОПУСКНАЯ СПОСОБНОСТЬ КВАНТОВОГО КАНАЛА СВЯЗИ

§ 5.1. Формулировка и обсуждение квантовой теоремы кодирования

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

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

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

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

На выходе производится измерение некоторой наблюдаемой $M = \{M_{\widehat{w}}^{(n)}\}$ в пространстве $\mathcal{H}^{\otimes n}$ (получив исход измерения $\widehat{w}$, считаем, что было послано $\widehat{w}$). В итоге приемник выдает ответ

о принятом решении; таким образом, разложение единицы в пространстве $\mathcal{H}^{\otimes n}$ описывает статистику всей решающей процедуры, которая включает в себя физическое измерение и последующую классическую обработку его результатов. Выбор наблюдаемой M формально аналогичен выбору решающей процедуры в классическом случае, но как мы увидим, играет здесь гораздо более важную роль. После того, как M выбрана, мы получаем классический канал $p_M(y \mid x) = \operatorname{Tr} S_x M_y$ в однобуквенном случае, и $p_{M^{(n)}}(\widehat{w} \mid w) = \operatorname{Tr} S_w M^{(n)}_{\widehat{w}}$ — в n-буквенном.

Определим шенноновскую взаимную информацию между входом и выходом. Если есть априорное распределение вероятностей π на $\mathcal{X}$ и выбрана процедура измерения M на выходе, то шенноновская информация между входом и выходом дается формулой

$$I_1(\pi, M) = \sum_x \pi_x \sum_y p_M(y \mid x) \Big[\log p_M(y \mid x) - \log \sum_z p_M(y \mid z) \pi_z \Big],$$

а максимальное количество информации, допустимое законами квантовой механики, равно

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

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

$$I_n(\pi, M) =$$

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

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

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

Можно сказать, что это есть двойственное проявление корреляций Эйнштейна — Подольского — Розена. Последние возникают,

когда рассматривается сцепленное (т. е. неразделимое) состояние составной квантовой системы, а измерения разделимы. Строгая супераддитивность информации имеет место для разделимых состояний и обусловлена существованием коллективных (сцепленных) измерений.

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

ОПРЕДЕЛЕНИЕ. *Кодом размера* N называется набор слов $w^{(1)}, \ldots, w^{(N)}$ вместе с разложением единицы $M = \{M_j\}$ в $\mathcal{H}^{\otimes n}$ с исходами $j = 0, 1, \ldots, N$; исход 0 означает уклонение от принятия решения.

Средняя ошибка кода равна

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

Обозначим $\min\limits_{W, M} \overline{P}_e(W, M) = p_e(n, N)$ минимальную среднюю ошибку по всем кодам размера N, использующим слова длины n. Чтобы сформулировать теорему кодирования и понятие классической пропускной способности для квантового канала C, введем следующие величины: для оператора плотности $S = \sum s_j |e_j\rangle\langle e_j|$ рассмотрим энтропию фон Неймана

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

В главе 8 мы подробно рассмотрим квантовые энтропийные количества, а пока нам понадобятся элементарные свойства квантовой энтропии:

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

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

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

Введем обозначение:

$$C = \max_{\pi} \left\{ H\left(\sum_x \pi_x S_x\right) - \sum_x \pi_x H(S_x) \right\}.$$

ТЕОРЕМА 8 (квантовая теорема кодирования). *При* $n \to \infty$ *имеем*:

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

2) $p_e(n, 2^{nR}) \not\to 0$, *если* $R > C$.

Эта теорема оправдывает название *классическая пропускная способность* для величины C. В самом деле, определим C_∞ как $\lim\limits_n C_n/n$, где $C_n = \max I_n(\pi, M)$. Из классической теоремы кодирования (теорема 2) вытекает, что утверждение теоремы 8 выполняется с заменой C на C_∞. Таким образом, утверждение теоремы 8 состоит в том, что $C_\infty = C$.

Задача 24. Рассмотрим двоичный квантовый канал с чистыми состояниями ψ_0, ψ_1. Докажите, что

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

а максимум информации по измерениям и априорным распределениям за один шаг

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

дается формулой (4.7). Поскольку $C_1 < C$ при $0 < \varepsilon < 1$, отсюда следует свойство супераддитивности $nC_1 < C_n$ для достаточно больших n.

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

$$C = \max_{\pi} H\left(\sum_x \pi_x |\psi_x\rangle\langle\psi_x|\right).$$

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

Рассмотрим конфигурацию (4.3) из трех равновероятных «равноугольных» векторов ψ_0, ψ_1, ψ_2. Тогда

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

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

Задача 25. Величина $C_1 \approx 0,645$ достигается для неравномерного распределения $\pi_0 = \pi_1 = 1/2$, $\pi_2 = 0$ и соответствующего оптимального измерения для двух равновероятных состояний (см. пример 1 в § 4.3).

Это еще раз иллюстрирует феномен супераддитивности классической информации в квантовом канале связи.

§ 5.2. Квантовая энтропийная граница и доказательство обратной теоремы

ТЕОРЕМА 9 (квантовая энтропийная граница).

$$I_1(\pi, M) \leqslant H\left(\sum \pi_x S_x\right) - \sum \pi_x H(S_x), \qquad (5.1)$$

причем имеет место строгое неравенство, если операторы $\pi_x S_x$ не коммутируют.

Первое, прямое доказательство этого неравенства (А. С. Холево, 1973), опирающееся на исследование свойств выпуклости квантовой энтропии, достаточно сложно, однако его можно довольно просто получить из полученного позднее свойства монотонности относительной энтропии (доказательство которого, однако, по крайней мере, столь же сложно), см. далее § 7.2 и Приложение.

Докажем слабое обращение теоремы кодирования, используя классическое неравенство Фано и квантовую энтропийную границу. Возьмем $N = 2^{nR}$, $R > C$, и рассмотрим произвольный набор кодовых слов $w^{(1)}, \ldots, w^{(N)}$ на входе, а на выходе — произвольное разложение единицы $M = \{M_j, j = 0, 1, \ldots, N\}$. Рассмотрим классическую случайную величину X со значениями $1, \ldots, N$ (номер посланного слова), которые имеют равные вероятности

$1/N$. На выходе после измерения получим классическую случайную величину Y со значениями $0, 1, \ldots, N$. Взаимная информация равна $I(X;Y) = H(X) - H(X \mid Y)$, где $H(X) = \log N = nR$ — энтропия равномерного распределения. Условная энтропия оценивается с помощью неравенства Фано: $H(X \mid Y) \leqslant 1 + \mathbf{P}(X \neq Y) \log N$. Таким образом, $\max I(X, Y) \geqslant nR(1 - p_e(n, 2^{nR})) - 1$ (это повторение доказательства слабого обращения классической теоремы Шеннона).

Из квантовой энтропийной границы (5.1) вытекает неравенство

$$I(X;Y) \leqslant \max_{\pi} \Big[H\Big(\sum_w \pi_w S_w\Big) - \sum_w \pi_w H(S_w) \Big] = \overline{C}_n.$$

Лемма 8. *Последовательность $\overline{C}_n$ аддитивна*: $\overline{C}_n = n\overline{C}_1 \equiv nC$.

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

$$\max_{\pi_{ij}} \Big[H\Big(\sum_{ij} \pi_{ij} S_i^1 \otimes S_j^2\Big) - \sum_{ij} \pi_{ij} H(S_i^1 \otimes S_j^2) \Big] =$$

$$= \max_{\pi_i^1} \Big[H\Big(\sum_i \pi_i^1 S_i^1\Big) - \sum_i \pi_i^1 H\Big(S_i^1\Big) \Big] + \max_{\pi_j^2} \Big[H\Big(\sum_j \pi_j^2 S_i^2\Big) - \sum_j \pi_j^2 H\Big(S_j^2\Big) \Big]$$

Очевидно, что $\max\limits_{\pi_{ij}} \geqslant \max\limits_{\pi_{ij} = \pi_i \pi_j}$, тогда в силу свойства аддитивности квантовой энтропии

$$H\Big(\sum_i \pi_i^1 S_i^1 \otimes \sum_j \pi_j^2 S_j^2\Big) = H\Big(\sum_i \pi_i^1 S_i^1\Big) + H\Big(\sum_j \pi_j^2 S_j^2\Big),$$

откуда $\overline{C}_n \geqslant n\overline{C}_1$.

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

$$H\Big(\sum \pi_{ij} S_i^1 \otimes S_j^2\Big) \leqslant H\Big(\sum \pi_i^1 S_i^1\Big) + H\Big(\sum \pi_i^2 S_i^2\Big),$$

где $\pi_i^1 = \sum\limits_j \pi_{ij}$, $\pi_j^2 = \sum\limits_i \pi_{ij}$ — маргинальные распределения.

Окончательно, $nC \geqslant nR[1 - p_e(n, 2^{nR})] - 1$, и если $R > C$, то не может быть $p_e(n, 2^{nR}) \to 0$ при $n \to \infty$.

§ 5.3. Доказательство прямой теоремы для канала с чистыми состояниями

Доказательство прямого утверждения теоремы кодирования дадим в простейшем случае чистых состояний $S_x = |\psi_x\rangle\langle\psi_x|$. В этом случае классическая пропускная способность квантового канала $x \to S_x$ дается выражением

$$C = \max_{\pi}\left[H\left(\sum \pi_x S_x\right)\right].$$

Доказательство. Пусть $R < C$. Докажем, что $p_e(n, 2^{nR}) \to 0$. Рассмотрим среднюю вероятность ошибки кода

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

которая зависит от выбора слов W и наблюдаемой M. Для ее минимизации желательно выбрать M_j как можно ближе к $|\psi_{w(j)}\rangle\langle\psi_{w(j)}|$. С этой целью введем переполненную систему

$$M_j = G^{-1/2}|\psi_{w(j)}\rangle\langle\psi_{w(j)}|G^{-1/2} = |\widehat{\psi}_{w(j)}\rangle\langle\widehat{\psi}_{w(j)}|,$$

где $G = \sum\limits_{j=1}^{N} |\psi_{w(j)}\rangle\langle\psi_{w(j)}|$ — оператор Грама, $G^{-1/2}|\psi_{w(j)}\rangle = |\widehat{\psi}_{w(j)}\rangle$.

Оператор G имеет областью значений подпространство $\operatorname{supp} G = \mathcal{L}$, порожденное кодовыми векторами $|\psi_{w(j)}\rangle$; G^{-1} — это обобщенный обратный, равный 0 на ортогональном дополнении к $\mathcal{L}$. Обозначим P проектор на подпространство $\mathcal{L}$, тогда $\sum\limits_j M_j = P \leqslant I$.

Если $\psi_{w(j)}$ — линейно зависимы, то получится действительно «переполненная» система $\widehat{\psi}_{w(j)}$, если же независимы — то ортонормированная система. Оператор Грама всегда можно обратить на подпространстве $\mathcal{L}$.

Подставляя такую наблюдаемую M, получим:

$$\overline{P}_e(W; M) = \frac{1}{N}\sum_{j=1}^{N}(1 - |\langle\psi_{w(j)} \mid \widehat{\psi}_{w(j)}\rangle|^2),$$

$$\langle\psi_{w(j)} \mid \widehat{\psi}_{w(j)}\rangle = \langle\widehat{\psi}_{w(j)}|G^{1/2}|\widehat{\psi}_{w(j)}\rangle.$$

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

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

$$= \frac{2}{N}(N - \operatorname{Tr} G^{1/2}) = \frac{2}{N}\operatorname{Tr}(G - G^{1/2}).$$

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

$$2(x-\sqrt{x}) = (1-\sqrt{x})^2 + (x-1) = \frac{(1-x)^2}{(1+\sqrt{x})^2} + (x-1) \leqslant$$

$$\leqslant (1-x)^2 + (x-1) = x^2 - x, \quad x \geqslant 0.$$

Отсюда следует, что $2(G-G^{1/2}) \leqslant G^2 - G$. Теперь применим метод случайных кодов. Надо доказать существование кода, для которого вероятность ошибки стремится к нулю. Идея, предложенная Шенноном, состоит в том, чтобы рассмотреть случайное распределение на всевозможных словах, тогда минимальная ошибка оценивается сверху средним по ансамблю случайных слов.

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

$$P(w = (x_1, \ldots, x_n)) = \pi_{x_1} \cdot \ldots \cdot \pi_{x_n}$$

(буквы берутся также независимо с одинаковым распределением на алфавите π.) Усреднение по такому случайному ансамблю слов будет обозначаться через $\langle\!\langle\ \rangle\!\rangle$. Отметим, что

$$\langle\!\langle G \rangle\!\rangle = \sum \langle\!\langle |\psi_{w^{(i)}}\rangle\langle\psi_{w^{(i)}}| \rangle\!\rangle = N \overline{S}_\pi^{\otimes n}.$$

Тогда

$$\langle\!\langle \overline{P}_e(W, M) \rangle\!\rangle \leqslant \frac{1}{N} \operatorname{Tr}(\langle\!\langle G^2 - G \rangle\!\rangle) =$$

$$= \frac{1}{N} \operatorname{Tr} \Big\langle\!\Big\langle \sum_j |\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}| \sum_k |\psi_{w^{(k)}}\rangle\langle\psi_{w^{(k)}}| - \sum_j |\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}| \Big\rangle\!\Big\rangle =$$

$$= \frac{1}{N} \operatorname{Tr} \sum_{j \neq k} \langle\!\langle |\psi_{w^{(j)}}\rangle\langle\psi_{w^{(j)}}| \rangle\!\rangle \langle\!\langle |\psi_{w^{(k)}}\rangle\langle\psi_{w^{(k)}}| \rangle\!\rangle =$$

$$= (N-1) \operatorname{Tr} \Big(\sum_{x_1, \ldots, x_n} \pi_{x_1} \ldots \pi_{x_n} |\psi_{x_1}\rangle\langle\psi_{x_1}| \otimes \cdots \otimes |\psi_{x_n}\rangle\langle\psi_{x_n}| \Big)^2 =$$

$$= (N-1) \operatorname{Tr} \big(\overline{S}_\pi^{\otimes n} \big)^2,$$

где $\overline{S}_\pi = \sum \pi_x |\psi_x\rangle\langle\psi_x|$. Итак

$$\langle\!\langle \overline{P}_e(W, M) \rangle\!\rangle \leqslant (N-1) \operatorname{Tr}_{\mathcal{H}^{\otimes n}} (\overline{S}_\pi^{\otimes n})^2 = (N-1) \big[\operatorname{Tr}_{\mathcal{H}} (\overline{S}_\pi^2) \big]^n.$$

Так как $N - 1 \leqslant 2^{nR}$, и абсолютный минимум не превосходит среднего по ансамблю, то

$$p_e(n, 2^{nR}) \leqslant \langle\!\langle \overline{P}_e(W, M) \rangle\!\rangle \leqslant 2^{nR} 2^{n \log \operatorname{Tr}(\bar{S}_\pi)^2} = 2^{-n(\max_\pi(-\log \operatorname{Tr} \bar{S}_\pi)^2 - R)}. \quad (5.2)$$

Итак, $p_e(n, 2^{nR}) \to 0$ при $R < \widetilde{C} = \max_\pi(-\log \operatorname{Tr} \overline{S}_\pi^2)$.

В ряде случаев величину

$$\widetilde{C} = \max_{\pi}\Big(-\log\sum_{x,y}|\langle\psi_x\,|\,\psi_y\rangle|^2\pi_x\pi_y\Big)$$

можно вычислить. Так, для двоичного канала связи

$$\widetilde{C} = -\log\frac{1+\varepsilon^2}{2} > C_1$$

при $0<\varepsilon<1$ (см. рис. 9). Если $\widetilde{C} > R > C_1$, то ошибка стремится к 0. Эта оценка доказывает, что $C_\infty \geqslant \widetilde{C} > C_1$, так что в этом примере имеет место строгая супераддитивность: $C_n > nC_1$.

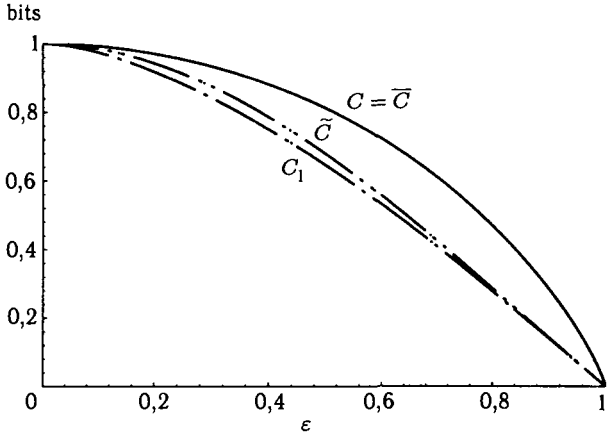


Рис. 9. Информационные характеристики двоичного канала как функции $\varepsilon = |\langle\psi_0|\psi_1\rangle|$

Однако эта часть доказательства, известная с 1979 г., не позволяет еще доказать, что $C_\infty \geqslant C$, а в этом и состоит прямое утверждение теоремы кодирования. Только в 1996 г. прямая теорема кодирования была доказана для квантового канала в случае чистых состояний [1]. Улучшить оценку (5.2) удалось, используя метод проекции на типичное подпространство оператора $\overline{S}_\pi^{\otimes n}$ (см. следующий параграф). На случай произвольных состояний доказательство обобщили Холево в 1996 г., а также Шумахер

[1] *Hausladen P., Jozsa R., Schumacher B., Westmoreland M., Wootters W.* Classical information capacity of a quantum channel // Phys. Rev. A. — 1996. — V. 54, № 3. — P. 1869–1876.

и Вестморленд в 1997 г. (подробно об истории доказательства теоремы кодирования см. в [6]). Здесь мы приведем другое доказательство, не использующее проекции на типичное подпространство и позволяющее получить значительно более точную оценку. Оно является обобщением на квантовый случай известного метода Галлагера [1], основанного на некотором точном неравенстве для вероятности ошибки.

Имеем

$$\frac{2}{N}\langle\!\langle G - G^{1/2}\rangle\!\rangle \leqslant \begin{cases} \frac{1}{N}\langle\!\langle G^2 - G\rangle\!\rangle = (N-1)(\overline{S}_\pi^{\otimes n})^2, \\ \frac{2}{N}\langle\!\langle G\rangle\!\rangle = 2\overline{S}_\pi^{\otimes n}. \end{cases}$$

Рассмотрим спектральное разложение $\overline{S}_\pi^{\otimes n} = \sum \lambda_J |e_J\rangle\langle e_J|$, тогда

$$\langle e_J|\langle\!\langle G - G^{1/2}\rangle\!\rangle|e_J\rangle \leqslant \lambda_J \min(2, (N-1)\lambda_J).$$

Используем неравенство $\min(a, b) \leqslant a^s b^{1-s}$, $0 \leqslant s \leqslant 1$. Тогда получим

$$\langle e_J|\langle\!\langle G - G^{1/2}\rangle\!\rangle|e_J\rangle \leqslant \lambda_J 2[(N-1)\lambda_J]^s,$$

откуда

$$\overline{P}_e(W, M)\rangle\!\rangle \leqslant \frac{2}{N}\sum_J \langle e_J|(G - G^{1/2})|e_J\rangle \leqslant 2(N-1)^s \sum_J \lambda_J^{s+1} =$$

$$= 2(N-1)^s \operatorname{Tr}\left[(\overline{S}_\pi^{\otimes n})^{1+s}\right].$$

Окончательно,

$$p_e(n, 2^{nR}) \leqslant 2 \max_{0 \leqslant s \leqslant 1} 2^{-n[\max_\pi(-\log \operatorname{Tr}(\bar{S}_\pi)^{1+s} - sR)]}.$$

Заметим теперь (см. рис. 10), что

$$\frac{d}{ds}\Big|_{s=0}\left[-\log \operatorname{Tr} \overline{S}_\pi^{\,1+s}\right] = H(\overline{S}_\pi).$$

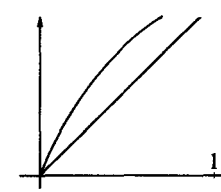


Рис. 10. График функции $-\log \operatorname{Tr} \bar{S}_\pi^{\,1+s}$ и прямой с угловым коэффициентом $R < C$

Возьмем π, которое максимизирует $H(\overline{S}_\pi)$, т. е. $H(\overline{S}_\pi) = C$. Тогда при $R < C$ и достаточно малых s коэффициент при n в показателе экспоненты будет положительным, откуда следует, что $p_e(n, 2^{nR}) \to 0$ при $n \to \infty$ с экспоненциальной скоростью. Более подробное рассмотрение этого показателя позволяет количественно оценить эту скорость, т. е. *функцию надежности* квантового канала [6].

Одной из интересных открытых проблем является обобщение этого доказательства на случай каналов со смешанными состояниями. Существующее доказательство квантовой теоремы кодирования, которое опирается на метод типичного подпространства, лишь устанавливает, что $p_e(n, 2^{nR}) \to 0$ при $R < C$, но не позволяет оценить функцию надежности.

§ 5.4. Сжатие квантовой информации

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

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

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

чем ближе она к единице, тем точнее воспроизведение.

ТЕОРЕМА 10[1]. *Пусть* $\overline{S}_p = \sum_{x=1}^{a} p_x |\psi_x\rangle\langle\psi_x|$. *Тогда*

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

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

[1] *Jozsa R., Schumacher B.* A new proof of the quantum noiseless coding theorem // J. Modern Optics. — 1994. — V. 41, № 12. — P. 2343–2349.

ЗАМЕЧАНИЕ. Это утверждение раскрывает информационный смысл квантовой энтропии, подобно тому как идея сжатия данных раскрывала смысл классической энтропии. Для смеси чистых квантовых состояний $\sum_{x=1}^{a} p_x|\psi_x\rangle\langle\psi_x| = \overline{S}_p$ энтропия оператора плотности $\overline{S}_p$ является мерой квантовой информации, содержащейся в ансамбле, поскольку $2^{nH(\overline{S}_p)}$ есть критическое значение размерности гильбертова пространства. (Напомним классический результат: пусть имеется источник, посылающий символы $1, \ldots, a$ с вероятностями $p_1, \ldots, p_a$, тогда количество слов, асимптотически безошибочно пересылаемых источником, есть $N \sim 2^{nH(p)}$, где $H(p) = -\sum_x p_x \log p_x$.)

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

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

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

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

Выделим в множестве всевозможных значений J подмножество

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

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

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

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

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

Тогда точность воспроизведения равна

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

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

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

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

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

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

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

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

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

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

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

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

для достаточно больших n.

Задача 26. Дайте другое доказательство прямой теоремы кодирования для канала с чистыми состояниями, используя проекцию на типичное подпространство. Указание. Используйте «сжатые» векторы

$$|\widetilde{\psi}_j\rangle = \frac{P_{\delta,n}|\psi_{w(j)}\rangle}{\|P_{\delta,n}|\psi_{w(j)}\rangle\|}$$

и соответствующую переполненную систему как субоптимальную наблюдаемую для получения $\varepsilon - \delta$-оценки для вероятности ошибки.

До сих пор не решена проблема обобщения квантового сжатия данных на источники, производящие произвольные (смешанные) состояния $S_1, \dots, S_a$. Имеется предположение, что в этом случае критическая размерность равна $2^{n\left(H\left(\sum_x \pi_x S_x\right) - \sum_x \pi_x H(S_x)\right)}$.

ГЛАВА 6

КВАНТОВЫЕ КАНАЛЫ

§ 6.1. Эволюции квантовой системы

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

Каковы минимальные требования к теоретическому описанию квантового канала связи? Всякий канал должен преобразовывать квантовые состояния из $\mathcal{S}(\mathcal{H}_1)$ в квантовые состояния (вообще говоря, в другом пространстве $\mathcal{S}(\mathcal{H}_2)$). Необходимое требование для согласованности со статистической интерпретацией квантовой механики состоит в том, чтобы смеси входных состояний переходили в такие же смеси выходных состояний, т. е. канал должен задаваться аффинным отображением

$$\Phi\Big(\sum p_i S_i\Big) = \sum p_i \Phi(S_i).$$

ЛЕММА 9. *Аффинное отображение* Φ: $\mathcal{S}(\mathcal{H}_1) \to \mathcal{S}(\mathcal{H}_2)$ *однозначно продолжается до отображения* $\mathcal{B}(\mathcal{H}_1) \to \mathcal{B}(\mathcal{H}_2)$, *обладающего следующими свойствами*:

1) $\Phi\big(\sum c_j T_j\big) = \sum c_j \Phi(T_j)$ (*линейность*);

2) $T \geqslant 0 \Longrightarrow \Phi(T) \geqslant 0$ (*положительность*);

3) $\operatorname{Tr} \Phi(T) = \operatorname{Tr} T$ (*сохранение следа*).

Эти свойства следуют из того, что выпуклое множество $\mathcal{S}(\mathcal{H})$ порождает $\mathcal{B}(\mathcal{H})$, отсюда с помощью аффинности доказывается линейность, а тот факт, что состояния переходят в состояния, влечет свойства 2 и 3 (см. доказательство теоремы 4 в § 2.3.).

Описание квантовой эволюции в терминах состояний соответствует картине Шредингера, двойственная ей картина Гейзенберга дает описание в терминах наблюдаемых. Рассмотрим сопряженное

отображение Φ^* из $\mathcal{B}(\mathcal{H}_2)$ в $\mathcal{B}(\mathcal{H}_1)$, определяемое с помощью формы двойственности $\operatorname{Tr} TX$:

$$\operatorname{Tr} \Phi[T]X = \operatorname{Tr} T\Phi^*[X],$$

тогда отображение Φ^* имеет свойства:

1) линейность;

2) положительность;

3) сохранение единицы: $\Phi^*[I] = I$.

Здесь имеется прямая аналогия с марковскими (переходными) отображениями в теории вероятностей.

В рассматриваемых далее примерах $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$.

Пример 1 (обратимая эволюция). Рассмотрим отображение $\Phi[S] = USU^*$, где U — унитарный оператор в $\mathcal{H}$. Имеем

$$\operatorname{Tr} \Phi[T]X = \operatorname{Tr} (UTU^*)X = \operatorname{Tr} T(U^*XU),$$

так что $\Phi^*[X] = U^*XU$.

Теорема 11 (Вигнер). *Если Φ — аффинное взаимно однозначное отображение $\mathcal{S}(\mathcal{H})$ на $\mathcal{S}(\mathcal{H})$, то $\Phi[S] = USU^*$, где U — унитарный или антиунитарный оператор.*

Для антиунитарного оператора U выполняется $\|U\psi\| = \|\psi\|$, но $U\left(\sum c_j\psi_j\right) = \sum \overline{c}_j U\psi_j$. Примером антиунитарного оператора является комплексное сопряжение в каком-либо фиксированном базисе. В физике такие операторы связаны с обращением времени. Линейное продолжение на $\mathcal{B}(\mathcal{H})$ в этом случае дается формулой

$$\Phi[T] = UT^*U^*.$$

Пример 2 (релаксация к состоянию S_{fin}). Рассмотрим отображение, переводящее любое состояние в фиксированное конечное $S \to S_{\text{fin}}$. Продолжая по линейности на $\mathcal{B}(\mathcal{H})$, получаем $\Phi[T] = S_{\text{fin}} \operatorname{Tr} T$. Сопряженное отображение действует на наблюдаемые по формуле $\Phi^*[X] = (\operatorname{Tr} S_{\text{fin}} X)I$.

Пример 3 (условное ожидание, связанное с идеальным измерением). Пусть $\{E_i\}$ — ортогональное разложение единицы в $\mathcal{H}$, т. е. $E_iE_j = \delta_{ij}E_j$, $\sum\limits_j E_j = I$. Рассмотрим отображение

$$\Phi[S] = \sum_j E_j S E_j. \tag{6.1}$$

Это аффинное отображение, переводящее состояния в состояния. Оно возникает в связи с проекционным постулатом Людерса—фон Неймана: идеальное измерение описывается ортогональным разложением единицы $\{E_j\}$, так что после измерения получается один из исходов j, с вероятностью $p_j = \operatorname{Tr} SE_j$, а состояние той доли статистического ансамбля, в которой получен этот исход (апостериорное состояние), описывается оператором плотности

$$S_j = \frac{E_j S E_j}{\operatorname{Tr} E_j S E_j} = \frac{E_j S E_j}{\operatorname{Tr}(S E_j)}.$$

В частности, полное идеальное измерение соответствует случаю $E_i = |e_i\rangle\langle e_i|$, где $\{e_i\}$ — ортонормированный базис в $\mathcal{H}$. Тогда

$$\Phi[S] = \sum |e_i\rangle\langle e_i|\langle e_i|S|e_i\rangle,$$

т. е. после измерения система оказывается в состоянии $|e_i\rangle\langle e_i|$ с вероятностью $\langle e_i|S|e_i\rangle$. При чистом начальном состоянии $S = |\psi\rangle\langle\psi|$ эта вероятность равна $|\langle e_i \mid \psi\rangle|^2$.

Если смешать подансамбли, полученные после измерения, то получится новое состояние ансамбля

$$\sum_j \frac{E_j S E_j}{\operatorname{Tr}(S E_j)} \operatorname{Tr}(S E_j) = \sum_j E_j S E_j,$$

которое, вообще говоря, отличается от исходного состояния S. Таким образом, в квантовой механике, в отличие от классической, идеальное измерение не сводится к простому считыванию значения наблюдаемой величины, а всегда предполагает воздействие на наблюдаемую систему, изменяющее ее состояние.

Отображение $\Phi^*[X] = \sum_j E_j X E_j$ обладает свойствами:

$$\Phi^*(\Phi^*[X]) = \Phi^*[X], \quad \Phi^*[X\Phi^*[Y]] = \Phi^*[X]\Phi^*[Y],$$

характеризующими условное ожидание в теории вероятностей.

Идеальное измерение обладает свойством воспроизводимости: повторив измерение сразу после получения исхода j, мы должны с достоверностью получить исход j. Ряд парадоксов и тупиков в стандартной формулировке квантовой механики связан с тем, что в качестве математической модели измерений рассматриваются только описанные выше идеальные процедуры, удовлетворяющие условию воспроизводимости. Эти трудности в значительной мере снимаются в обобщенной формулировке (см. ниже), позволяющей

охватить «приближенные» измерения, воздействие которых на систему может быть произвольно слабым.

Пример 4 (эволюция открытой квантовой системы, взаимодействующей с окружением). Пусть $\mathcal{H}$ — гильбертово пространство системы, $\mathcal{H}_0$ — пространство окружения. Эволюция составной системы является обратимой и описывается унитарным оператором U.

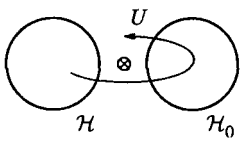


Рис. 11. Открытая квантовая система

В начальный момент система и окружение находятся в состоянии $S \otimes S_0$, затем взаимодействие «подкручивает» состояние (см. рис. 11). Усредняя по окружению, получаем необратимую эволюцию самой системы

$$\Phi[S] = \mathrm{Tr}_{\mathcal{H}_0} U(S \otimes S_0)U^*, \tag{6.2}$$

или

$$\Phi^*[X] = \mathrm{Tr}_{\mathcal{H}_0}(I \otimes S_0)U^*(X \otimes I_0)U.$$

Рассмотрим процесс косвенного измерения, когда система взаимодействует с пробной системой $\mathcal{H}_0$, а затем над пробной системой производится идеальное измерение наблюдаемой $\{E_j^0\}$. Тогда согласно примеру 3, вероятность получить исход j равна $p_j = \mathrm{Tr}\, U(S \otimes S_0)U^*(I \otimes E_j)$, а состояние той доли статистического ансамбля, в которой получен этот исход, описывается оператором плотности S_j, таким что

$$p_j S_j = \mathrm{Tr}_{\mathcal{H}_0}(I \otimes E_j)U(S \otimes S_0)U^*(I \otimes E_j).$$

Таким образом, формула (6.2) может быть записана как

$$\Phi[S] = \sum_j p_j S_j.$$

Пусть для простоты состояние $S_0 = |\psi_0\rangle\langle\psi_0|$ чистое, а измерение полное, т. е. $E_j = |e_j^0\rangle\langle e_j^0|$, где e_j — ортонормированный базис в $\mathcal{H}_0$. Вводя операторы V_j, определяемые формулой $\langle\varphi \mid V_j\psi\rangle = \langle\varphi \otimes e_j^0 \mid U\psi \otimes \psi_0\rangle$, $\varphi, \psi \in \mathcal{H}$, получаем

$$p_j = \mathrm{Tr}\, V_j S V_j^*, \quad S_j = \frac{V_j S V_j^*}{p_j}, \tag{6.3}$$

так что

$$\Phi[S] = \sum_j V_j S V_j^*.$$

Эта формула обобщает соотношение (6.1) для идеального измерения, и описывает статистику (вероятности исходов и апостериорные состояния) более общего процесса, включающего измерения, не удовлетворяющие требованию воспроизводимости (в частности, приближенные измерения). Подробное обсуждение понятия квантового измерения с этих позиций см. в [10, 5, 11].

§ 6.2. Вполне положительные отображения

Оказывается, что во всех четырех примерах отображение Φ обладает очень важным свойством, помимо линейности, положительности и сохранения следа. (В картине Гейзенберга отображение Φ^* линейно, положительно и сохраняет I.)

ОПРЕДЕЛЕНИЕ. Отображение Φ^* называется *вполне положительным*, если выполняется условие:

1) для любого $m = 1, 2, \ldots$, отображение $\Phi^* \otimes \mathrm{Id}_m$, где Id_m — тождественное отображение в $\mathcal{B}(\mathcal{H}_0)$, $\dim \mathcal{H}_0 = m$, положительно.

Дадим другую формулировку, используя изоморфизм

$$\mathcal{H} \otimes \mathcal{H}_0 \approx \underbrace{\mathcal{H} \oplus \ldots \oplus \mathcal{H}}_{m}, \quad X \approx \left[X_{ij}\right]_{i,j=1,\ldots,m},$$

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

Тогда условие 1) можно переформулировать в виде:

1′) если матрица $\left[X_{ij}\right] \geqslant 0$, то $\left[\Phi^*[X_{ij}]\right] \geqslant 0$.

Задача 27. Доказать, что условие 1) равносильно условию:

2) для любых конечных наборов $\{\psi_j\} \subset \mathcal{H}_1$, $\{X_j\} \subset \mathcal{B}(\mathcal{H}_2)$ выполнено неравенство:

$$\sum_{j,k} \langle \psi_j | \Phi^*[X_j^* X_k] | \psi_k \rangle \geqslant 0.$$

Преимущество условия 2) в простой проверяемости. Скажем, в примере 4 для проверки надо рассмотреть

$$\sum_{j,k} \langle \psi_j | \operatorname{Tr}_{\mathcal{H}_0} (I \otimes S_0) U^* (X_j^* X_k \otimes I_0) U | \psi_k \rangle.$$

Достаточно доказать положительность для чистого $S_0 = |\psi_0\rangle\langle\psi_0|$ и применить спектральное разложение в общем случае. Для

$S_0 = |\psi_0\rangle\langle\psi_0|$ имеем

$$\sum_{j,k}\langle\psi_0| \otimes \langle\psi_k|U^*(X_j^* \otimes I_0)(X_k \otimes I_0)U|\psi_k\rangle \otimes |\psi_0\rangle = \\ = \sum_k \|(X_k \otimes I_0)U|\psi_k\rangle \otimes |\psi_0\rangle\|^2 \geqslant 0.$$

ПРИМЕР. Не все положительные отображения вполне положительны. Фиксируем базис в $\mathcal{H}$, тогда всякий оператор X задается матрицей $[X_{ij}]$. Рассмотрим преобразование транспонирования $\Phi[X] = X^T$, которое совпадает с комплексным сопряжением на эрмитовых матрицах.

Задача 28. Доказать, что условие 1 нарушается уже при $m = 2$.

Физический смысл нарушения условия 1: в одной системе происходит обращение времени, в другой — нет, таким образом получаем нефизическое преобразование составной системы.

Свойство полной положительности было введено Стайнспрингом, который доказал теорему, обобщающую теорему Наймарка. Мы рассмотрим теорему Стайнспринга в важном частном случае.

ТЕОРЕМА 12. *Пусть Φ^* — вполне положительное отображение из $\mathcal{B}(\mathcal{H}_2)$ в $\mathcal{B}(\mathcal{H}_1)$, сохраняющее единицу. Существует представление*

$$\Phi^*[X] = V^*\pi[X]V,$$

где V — изометрическое отображение из $\mathcal{H}_1$ в $\widetilde{\mathcal{H}}$, а π является $*$-представлением алгебры[1] *$\mathcal{B}(\mathcal{H}_2)$ в $\widetilde{\mathcal{H}}$.*

В нашем случае Φ^*: $\mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$. Если же отображение Φ^* действует из алгебры $\mathcal{B}_2$ в алгебру $\mathcal{B}_1$, причем хотя бы одна из алгебр $\mathcal{B}_1$ или $\mathcal{B}_2$ коммутативна, то свойство полной положительности следует из положительности. Таким образом, свойство полной положительности существенно некоммутативно (задача 29).

Доказательство. В качестве заготовки для $\widetilde{\mathcal{H}}$ рассмотрим алгебраическое тензорное произведение линейных пространств $\mathcal{L} = \mathcal{H}_1 \otimes \mathcal{B}(\mathcal{H}_2)$. Оно порождается элементами вида $\Psi = \psi \otimes X$, $\psi \in \mathcal{H}_1$, $X \in \mathcal{B}(\mathcal{H}_2)$. Скалярное произведение вводится согласно формуле для квадрата нормы:

$$\|\sum_j \psi_j \otimes X_j\|^2 = \sum_{j,k}\langle\psi_j|\Phi^*[X_j^* X_k]|\psi_k\rangle,$$

[1] $*$-Представлением называется линейное отображение, сохраняющее алгебраические операции и инволюцию: $\pi[XY] = \pi[X]\pi[Y]$, $\pi[X^*] = \pi[X]^*$.

которая неотрицательна. После факторизации по подпространству $\mathcal{L}_0$ нулевой нормы, получим $\widetilde{\mathcal{H}} = \mathcal{L}/\mathcal{L}_0$. Введем $V\psi = \psi \otimes I$ и

$$\pi[Y]\Psi = \pi[Y](\psi \otimes X) = \psi \otimes YX.$$

Задача 30. Проверить, что требуемое представление имеет место для данной конструкции.

Изометричность V следует из сохранения единицы. Обратно, условие 2 выполнено, если есть такое представление, так как

$$\sum_j \langle \psi_j | V^* \pi[X_j^* X_k] V | \psi_k \rangle = \sum_j \langle \psi_j | V^* \pi[X_j]^* \pi[X_k] | \psi_k \rangle \geqslant 0.$$

Такого типа рассуждения, когда по некоторому объекту, обладающему каким-либо свойством положительности, строится гильбертово пространство и конкретное представление в этом пространстве, называется конструкцией Гельфанда — Наймарка — Сигала (ГНС). Аналогичная теорема верна для отображений более общих $*$-алгебр. В нашем случае можно получить более конкретное представление.

Следствие 1 (представление Крауса). *Отображение* Φ^*: $\mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$ *является вполне положительным и сохраняющим единицу тогда и только тогда, когда*

$$\Phi^*[X] = \sum_j V_j^* X V_j, \tag{6.4}$$

или

$$\Phi[S] = \sum_j V_j S V_j^*, \tag{6.5}$$

где V_j *операторы из* $\mathcal{B}(\mathcal{H}_1)$ *в* $\mathcal{B}(\mathcal{H}_2)$, *такие что* $\sum_j V_j^* V_j = I$.

Это представление удобно тем, что не использует вспомогательного пространства $\mathcal{H}_0$.

Доказательство. Любое $*$-представление алгебры $\mathcal{B}(\mathcal{H})$ унитарно эквивалентно кратному тождественного представления $\pi[X] = X \otimes I_0$, где I_0 — единица в $\mathcal{H}_0$, т. е. можно считать, что $\widetilde{\mathcal{H}} = \mathcal{H}_2 \otimes \mathcal{H}_0$ [10, разд. 9.2]. Записывая $I_0 = \sum |e_j^0\rangle\langle e_j^0|$, где e_j — ортонормированный базис в $\mathcal{H}_0$, введем операторы V_j, определяемые формулой $\langle \varphi \mid V_j \psi \rangle = \langle \varphi \otimes e_j^0 \mid V\psi \rangle$, $\varphi \in \mathcal{H}_2$, $\psi \in \mathcal{H}_1$. Представление (6.4) следует из теоремы Стайнспринга.

Обращаем внимание на аналогию с рассуждениями в примере 4. На самом деле, как показывает следующий результат, этот пример является порождающим.

СЛЕДСТВИЕ 2. *Всякое вполне положительное отображение* Φ^* *алгебры* $\mathcal{B}(\mathcal{H})$, *сохраняющее единицу, можно продолжить до обратимой эволюции составной системы в* $\mathcal{H}\otimes\mathcal{H}_0$, *где вторая система (окружение) находится в чистом состоянии* $|\psi_0\rangle\langle\psi_0|$.

Доказательство. Исходя из представления Стайнспринга, рассмотрим действие изометрического оператора V на $\psi\in\mathcal{H}$. Имеем $V|\psi\rangle\in\mathcal{H}\otimes\mathcal{H}_0$. Зафиксируем в $\mathcal{H}$ ортонормированный базис $\{e_j\}$. Тогда $V|\psi\rangle=\sum\limits_j|\psi_j\rangle\langle e_j\mid\psi\rangle$, где $\psi_j=Ve_j$. Поскольку V изометрично, имеем $\langle\psi_j\mid\psi_k\rangle=\delta_{jk}$, но система ψ_j может быть неполна. Дополним ее до ортонормированного базиса; пусть ψ_{jl} — такой ортонормированный базис в $\mathcal{H}\otimes\mathcal{H}_0$,, что $\psi_{j0}=\psi_j$. Тогда оператор $U=\sum\limits_{jl}|\psi_{jl}\rangle\langle e_j|\otimes\langle e_l^0|$ унитарный, так как он переводит ортонормированный базис в ортонормированный базис. Полагаем $|\psi_0\rangle=|e_0^0\rangle$, тогда получаем

$$\Phi^*[X]=V^*(X\otimes I_0)V=\mathrm{Tr}_{\mathcal{H}_0}(I\otimes|\psi_0\rangle\langle\psi_0|)U^*(X\otimes I_0)U,$$

т. е. приходим к ситуации примера 4. В самом деле,

$$\langle V\psi|(X\otimes I_0)|V\psi\rangle=\langle\psi|\otimes\langle\psi_0|U^*(X\otimes I_0)U|\psi\rangle\otimes|\psi_0\rangle,$$

так как

$$U|\psi\rangle\otimes|e_0^0\rangle=\sum_j|\psi_{j0}\rangle\langle e_j\mid\psi\rangle=V|\psi\rangle.$$

§ 6.3. Определение канала

Предыдущее обсуждение показывает, что в определение канала следует включить свойство полной положительности.

ОПРЕДЕЛЕНИЕ. *Каналом (в пространстве состояний)* называется линейное отображение Φ из $\mathcal{B}(\mathcal{H}_1)$ в $\mathcal{B}(\mathcal{H}_2)$, сохраняющее след и такое, что Φ^* — вполне положительное отображение. *Каналом в пространстве наблюдаемых* называется линейное вполне положительное отображение из $\mathcal{B}(\mathcal{H}_2)$ в $\mathcal{B}(\mathcal{H}_1)$, переводящее единицу в единицу.

Рассмотрим два примера, позволяющие установить связь этого определения с теми каналами, которые рассматривались при определении классической пропускной способности.

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

$$\Phi[S]=\sum_j S_j\langle e_j|S|e_j\rangle,$$

где S_j — состояния в $\mathcal{B}(\mathcal{H}_2)$, e_j — ортонормированный базис в $\mathcal{H}_1$.

Если S — состояние на входе, то $\langle e_j|S|e_j\rangle=p_j$ — распределение вероятностей на наборе состояний S_j и $\Phi[S]=\sum_j p_jS_j$. Если $p_j=\delta_{kj}$, то получим на выходе состояние S_k, а в общем случае — смесь. Такой канал переводит классическое состояние $\{p_i\}$ в квантовое состояние (можно рассматривать это также как отображение диагональных матриц в произвольные матрицы плотности. Фактически, именно такие каналы рассматривались в гл. 5, где для них была доказана теорема кодирования.

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

$$\Phi[S]=\sum_j(\operatorname{Tr} SM_j)|e_j\rangle\langle e_j|$$

где $\{e_j\}$ — ортонормированный базис в $\mathcal{B}(\mathcal{H}_2)$, а $\{M_j\}$ — наблюдаемая в $\mathcal{B}(\mathcal{H}_1)$. При произвольном входном квантовом состоянии получаем классический выход (диагональную матрицу плотности).

Простейший нетривиальный пример канала, не являющегося c-q или q-c, дает *деполяризующий канал* ($\mathcal{H}_1=\mathcal{H}_2=\mathcal{H}$, $\dim\mathcal{H}=d$):

$$\Phi[S]=(1-p)S+p\,\frac{I}{d}\operatorname{Tr}S. \tag{6.6}$$

При $0\leqslant p\leqslant 1$ он представляет собой смесь идеального и полностью деполяризующего каналов. Последний просто отображает любое состояние в хаотическое $\overline{S}=I/d$.

Задача 31. Покажите, что деполяризующий канал может быть охарактеризован свойством унитарной ковариантности: $\Phi[USU^*]=U\Phi[S]U^*$ для произвольного унитарного оператора U в $\mathcal{H}$.

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

Задача 32. Доказать полную положительность и найти представление Крауса для c-q, q-c и деполяризующего каналов.

Сформулируем теперь важную нерешенную проблему. Она называется проблемой аддитивности. Пусть задан канал

Φ: $\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$. Процесс передачи классической информации по такому квантовому каналу описывается диаграммой

$$i \longrightarrow S \xrightarrow{\Phi} S' \xrightarrow{M} j.$$

Для передачи классической информации используется блочное кодирование, т. е. рассматривается составной канал $\Phi^{\otimes n} = \Phi \otimes \cdots \otimes \Phi$ в пространстве $\mathcal{H}^{\otimes n}$, на выходе которого производится измерение $M^{(n)}$. Пусть w слово (последовательность символов), тогда передача классической информации через составной канал описывается новой диаграммой

$$w \longrightarrow \underbrace{S}_{\substack{\text{входное} \\ \text{состояние}}} \longrightarrow \underbrace{\Phi^{\otimes n}[S]}_{\substack{\text{выходное} \\ \text{состояние}}} \xrightarrow{M^{(n)}} j.$$

При этом переходная вероятность равна

$$\mathsf{P}(j \mid w) = \operatorname{Tr} \Phi^{\otimes n}[S_w^{(n)}] M_j^{(n)}.$$

Если задано распределение на входных словах, то можно найти совместное распределение входа и выхода и взаимную информацию $I^{(n)}(\pi_w^{(n)}, S_w^{(n)}, M^{(n)})$, причем оптимизации подлежит не только наблюдаемая $M^{(n)}$ и распределение $\pi_w^{(n)}$, но и сами состояния $S_w^{(n)}$ на входе канала Φ.

Согласно теореме кодирования для c-q каналов, классическая пропускная способность канала Φ равна

$$C \equiv \lim_n \frac{1}{n} \max_{\{\pi^{(n)}, S^{(n)}, M^{(n)}\}} I^{(n)}(\pi_w^{(n)}, S_w^{(n)}, M^{(n)}) =$$

$$= \lim_n \frac{1}{n} \max_{\{\pi_w, S_w\}} \left[H\Big(\sum \pi_w \Phi^{\otimes n}[S_w]\Big) - \sum \pi_w H(\Phi^{\otimes n}[S_w]) \right].$$

Вопрос заключается в следующем: верно ли, что

$$C = \max_{\pi_i, S_i} \left[H\Big(\sum \pi_i \Phi[S_i]\Big) - \sum \pi_i H(\Phi[S_i]) \right]. \tag{6.7}$$

Или, более общим образом, верно ли, что величина $C^1 = C^1(\Phi)$, определенная правой частью формулы (6.7), является аддитивной, т. е. выполняется ли равенство

$$C^1(\Phi_1 \otimes \ldots \otimes \Phi_n) = \sum_{i=1}^{n} C^1(\Phi_i).$$

Очевидно, что C^1 супераддитивна. Если бы нашелся пример, в котором C^1 строго супераддитивна, то это означало бы, что не только коллективные измерения на выходе, но и использование сцепленных состояний на входе позволяет увеличить количество классической информации при передаче по квантовым каналам. Несмотря на интенсивный численный поиск, такого примера пока найти не удалось. В то же время, не удается найти и сколь-нибудь общее аналитическое доказательство аддитивности. Обзор частичных результатов см. в работе Амосова, Холево и Вернера[1]. Для некоторого класса бистохастических каналов в $\mathcal{H}_2$, включающего деполяризующий канал, доказательство дано в работе Кинга[2].

Задача 33. Покажите, что для деполяризующего канала (6.6) величина

$$C^1 = \log d + \left(1 - p\frac{d-1}{d}\right)\log\left(1 - p\frac{d-1}{d}\right) + p\frac{d-1}{d}\log\frac{p}{d}, \qquad (6.8)$$

достигается для ансамбля равновероятных чистых состояний, отвечающих ортонормированному базису в $\mathcal{H}$.

§ 6.4. Каналы в $\mathcal{H}_2$

Рассмотрим каналы в пространстве q-бита $\mathcal{H}_2$, т. е. аффинные отображения единичного шара в пространстве $\mathbb{R}^3$ в себя, которые удовлетворяют некоторым дополнительным ограничениям, вытекающим из условия полной положительности. Используя полярное разложение, можно доказать[3], что всякий канал в $\mathcal{H}_2$ допускает представление

$$\Phi[S] = U_2\Lambda[U_1 S U_1^*]U_2^*, \qquad (6.9)$$

где U_1, U_2 унитарные операторы, а Λ имеет канонический вид

$$\Lambda[I] = I + \sum_{\gamma=x,y,z} t_\gamma \sigma_\gamma, \quad \Lambda[\sigma_\gamma] = \lambda_\gamma \sigma_\gamma, \quad \gamma = x, y, z, \qquad (6.10)$$

где λ_γ, t_γ вещественные числа.

---

[1] *Amosov G. G., Holevo A. S., Werner R. F.* On some additivity problems in quantum information theory // Probl. Inform. Transm. — 2000. — V. 36, № 4. LANL report math-ph/0003002.

[2] *King C.* Aditivity for a class of unital qubit channels // LANL e-print quant-ph/0103156.

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

Задача 34. Докажите возможность представления (6.9). Покажите, что унитарные эволюции в $\mathcal{H}_2$ соответствуют вращениям единичного шара в $\mathbb{R}^3$.

Сосредоточим наше внимание на каналах вида (6.10). Разумеется, полная положительность налагает нетривиальные ограничения на параметры λ_γ, t_γ [1]. Наиболее прост случай, когда $t_\gamma \equiv 0$, т. е. отображение представляет собой сжатие единичного шара вдоль осей x, y, z с коэффициентами $\lambda_x, \lambda_y, \lambda_z$. Это имеет место тогда и только тогда, когда канал Φ, и следовательно Λ, бистохастичен.

Задача 35. Покажите, что

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

где

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

и что неотрицательность этих чисел необходима и достаточна для полной положительности отображения Λ, а значит, и Φ. В частности, отображение (6.6) вполне положительно тогда и только тогда, когда $0 \leqslant p \leqslant 4/3$.

Вычислим величину $C^1 = C^1(\Phi)$ для произвольного бистохастического канала.

Лемма 10. *Пусть Φ бистохастический канал в $\mathcal{H}_2$, тогда*

$$C^1(\Phi) = 1 - \min_{S \in \mathfrak{S}(\mathcal{H})} H(\Phi(S)). \tag{6.11}$$

Доказательство. Неравенство $\leqslant$ в (6.11) вытекает из того, что для любого канала

$$C^1(\Phi) \leqslant \log \dim \mathcal{H} - \min_{S \in \mathfrak{S}(\mathcal{H})} H(\Phi(S)), \tag{6.12}$$

так что остается доказать неравенство $\geqslant$. Поскольку энтропия — вогнутая функция состояния, минимум достигается на чистом состоянии S. Беря равновероятно чистые состояния $S_0 = S$, $S_1 = I - S$, получаем

$$C^1(\Phi) \geqslant H(\tfrac{1}{2}\Phi(I)) - \tfrac{1}{2}[H(\Phi(S)) + H(\Phi(I - S))].$$

[1] См. там же.

Поскольку канал бистохастический, правая часть равна $H(\frac{1}{2}I) - \frac{1}{2}[H(\Phi(S)) + H(I - \Phi(S))]$, а в силу двумерности пространства, это равно правой части в (6.11).

При вычислении $\min_{S \in \mathfrak{S}(\mathcal{H})} H(\Phi(S))$, в силу унитарной инвариантности энтропии достаточно рассмотреть случай $\Phi = \Lambda$. Отметим также, что собственные значения оператора плотности (2.2) в $\mathcal{H}_2$ равны $(1 \pm |\vec{a}|)/2$, а значит, энтропия равна

$$H(S) = h\left(\frac{1 - |\vec{a}|}{2}\right). \tag{6.13}$$

Поскольку единичный шар отображается каналом Λ в эллипсоид с полуосями $|\lambda_\gamma|$, $\gamma = x, y, z$, минимум энтропии достигается на конце самой длинной полуоси, соответствующем оператору плотности с собственными значениями $(1 \pm \max_\gamma |\lambda_\gamma|)/2$. Отсюда получаем

$$C^1(\Phi) = 1 - h\left(\frac{1 - \max_\gamma |\lambda_\gamma|}{2}\right). \tag{6.14}$$

ГЛАВА 7

ЭНТРОПИЙНЫЕ ХАРАКТЕРИСТИКИ КВАНТОВЫХ СИСТЕМ

§ 7.1. Квантовая относительная энтропия

В классической статистике часто используется относительная энтропия двух распределений вероятностей $P=\{p_j\}$, $Q=\{q_j\}$:

$$H(P;\ Q)=\begin{cases}\sum\limits_j p_j(\log p_j-\log q_j), & \text{если из } q_j=0 \Longrightarrow p_j=0;\\ +\infty, & \text{в противном случае.}\end{cases}$$

В классической статистике $H(P;\ Q)$ играет роль (асимметричного) расстояния между распределениями [9].

В квантовом случае *относительная энтропия* вводится следующим образом: пусть S, T — операторы плотности, тогда

$$H(S;\ T)=\begin{cases}\operatorname{Tr} S(\log S-\log T), & \text{если } \operatorname{supp} S\subseteq \operatorname{supp} T,\\ +\infty, & \text{в остальных случаях.}\end{cases}$$

Более общим образом, для любых положительных операторов S, T с собственными значениями λ_j, μ_k и собственными векторами e_j, h_k положим

$$H(S;\ T)=\sum_{j,k}|\langle e_j\mid h_k\rangle|^2(\lambda_j\log\lambda_j-\lambda_j\log\mu_k+\mu_k-\lambda_j). \qquad (7.1)$$

Отметим, что слагаемое $\mu_k-\lambda_j$ можно опустить в случае операторов плотности.

Рассмотрим ряд свойств относительной энтропии в квантовом случае.

Лемма 11. $H(S;\ T)\geqslant\frac{1}{2}\operatorname{Tr}(S-T)^2$.

Доказательство. Используя формулу

$$\eta(\lambda) - \eta(\mu) = (\lambda - \mu)\eta'(\mu) + \tfrac{1}{2}(\lambda - \mu)^2\eta''(\xi),$$

где ξ — число между λ и μ, получаем неравенство

$$\lambda \log \lambda - \lambda \log \mu + \mu - \lambda \geqslant \tfrac{1}{2}(\mu - \lambda)^2,$$

подставляя которое в (7.1), мы и получаем результат.

Таким образом, относительная энтропия мажорируется снизу среднеквадратичным отклонением, в частности, она положительна, если $S \neq T$.

Следствие 1 (субаддитивность квантовой энтропии). *Пусть S_{12} — оператор плотности в $\mathcal{H}_1 \otimes \mathcal{H}_2$, $S_1 = \mathrm{Tr}_{\mathcal{H}_2} S_{12}$, $S_2 = \mathrm{Tr}_{\mathcal{H}_1} S_{12}$. Тогда $H(S_{12}) \leqslant H(S_1) + H(S_2)$, причем равенство имеет место тогда и только тогда, когда $S_{12} = S_1 \otimes S_2$.*

Доказательство.

$$H(S_1) + H(S_2) - H(S_{12}) = H(S_{12}; S_1 \otimes S_2) \geqslant 0.$$

Весьма важным является следующее свойство монотонности относительной энтропии (Линдблад, 1975; Ульманн, 1976).

Теорема 13. *Пусть Φ — произвольный канал, тогда*

$$H(S; T) \geqslant H(\Phi[S]; \Phi[T]).$$

Иногда это свойство называют «обобщенной H-теоремой». Заметим, что при унитарном (обратимом) преобразовании относительная энтропия не меняется.

Следствие 2 (H-теорема). *Если канал Φ оставляет инвариантным хаотическое состояние $\overline{S} = I/d$, т. е. $\Phi[I] = I$, то для всех состояний S имеет место $H(\Phi[S]) \geqslant H(S)$.*

Таким образом, H-теорема справедлива для динамик, сохраняющих хаотическое состояние.

Доказательство. Имеем

$$H(S; \overline{S}) = \mathrm{Tr}\, S(\log S - \log \tfrac{1}{d} I) = -H(S) + \log d. \qquad (7.2)$$

Тогда в силу монотонности

$$H(\Phi[S]; \Phi[\overline{S}]) \leqslant H(S; \overline{S}) \Longrightarrow -H[S] + \log d \geqslant -H(\Phi[S]) + \log d.$$

СЛЕДСТВИЕ 3 (квантовая энтропийная граница). *Пусть задан c-q-канал $i \to S_i$, распределение вероятностей $\pi = \{\pi_i\}$ на входе и наблюдаемая $M = \{M_j\}$ на выходе канала. Обозначим $I_1(\pi, M)$ шенноновскую взаимную информацию между переменными i и j. Тогда*

$$I_1(\pi, M) \leqslant H\Big(\sum_i \pi_i S_i\Big) - \sum_i \pi_i H(S_i).$$

Доказательство. Обозначим $\overline{S}_\pi = \sum \pi_i S_i$, тогда правая часть равна

$$H(\overline{S}_\pi) - \sum \pi_i H(S_i) = \sum \pi_i H(S_i; \overline{S}_\pi),$$

где $H(S_i; \overline{S}_\pi) = \operatorname{Tr} S_i(\log S_i - \overline{S}_\pi)$, поскольку суммирование по i с коэффициентами π_i дает:

$$\sum_i \pi_i H(S_i; \overline{S}_\pi) = -\underbrace{\operatorname{Tr} \sum \pi_i S_i \log \overline{S}_\pi}_{H(\overline{S}_\pi)} + \underbrace{\sum \pi_i \operatorname{Tr} S_i \log S_i}_{-\sum \pi_i H(S_i)}.$$

Сконструируем q-c-канал $\Psi[S] = \sum_j \operatorname{Tr} SM_j |e_j\rangle\langle e_j|$, тогда

$$\Psi[S_i] = \operatorname{diag}[p(j \mid i)]; \quad \Psi[\overline{S}_\pi] = \operatorname{diag}\Big[\sum_i \pi_i p(j \mid i)\Big].$$

(Базис $\{e_j\}$ произвольный и его выбор роли не играет.) При этом квантовая относительная энтропия переходит в классическую, и применяя теорему к каналу Ψ, получаем требуемое неравенство.

Все известные доказательства свойства монотонности достаточно сложны. В Приложении мы приводим недавнее доказательство Лесьневского и Рускаи, которое представляется наиболее прямым. Свойство монотонности связано также с фундаментальной теоремой выпуклости (Либ, 1973), и со следующим свойством сильной субаддитивности (Либ, Рускаи, 1974) в теории квантовой энтропии, см. [17, 14].

Пусть имеются три квантовые системы в гильбертовом пространстве $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, находящиеся в совместном состоянии S_{123}. Символ S с каким-либо набором индексов будет обозначать соответствующее частичное состояние. Тогда

$$H(S_{123}) + H(S_2) \leqslant H(S_{12}) + H(S_{23}).$$

ТЕОРЕМА 14. *Следующие свойства эквивалентны:*

1) *монотонность относительной квантовой энтропии*;

2) *сильная субаддитивность квантовой энтропии*;

3) *совместная выпуклость относительной квантовой энтропии по аргументам* S_1, S_2.

Доказательство. 1) $\Rightarrow$ 2). Обозначая $\overline{S}_\alpha = (\dim \mathcal{H}_\alpha)^{-1} I_\alpha$ хаотическое состояние в пространстве $\mathcal{H}_\alpha$, имеем

$$H(S_{23}; \overline{S}_{23}) = H(S_{23}; S_2 \otimes \overline{S}_3) + H(S_2; \overline{S}_2),$$
$$H(S_{123}; \overline{S}_{123}) = H(S_{123}; S_{12} \otimes \overline{S}_3) + H(S_{12}; \overline{S}_{12}).$$

Вычитая и используя (7.2), получаем

$$H(S_{12}) + H(S_{23}) - H(S_{123}) - H(S_2) = H(S_{123}; S_{12} \otimes \overline{S}_3) - H(S_{23}; S_2 \otimes \overline{S}_3).$$

Рассмотрим канал $\Phi[S_{123}] = \overline{S}_1 \otimes S_{23}$. Применяя теорему 13, получаем, что правая часть неотрицательна, т. е. свойство сильной субаддитивности.

2)$\Rightarrow$3). Рассмотрим положительный оператор 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 $\mathcal{H}_2$, который при подходящей нормировке является оператором плотности. Тогда

$$S_{12} = \sum_k |e_k\rangle\langle e_k| \otimes \sum_l A_{kl}, \quad S_{23} = \sum_l \sum_k A_{kl} \otimes |h_l\rangle\langle h_l|, \quad S_2 = \sum_{kl} A_{kl},$$

и свойство сильной субаддитивности влечет следующее неравенство:

$$-\sum_{kl} H(A_{kl}) + \sum_k H\left(\sum_l A_{kl}\right) + \sum_l H\left(\sum_k A_{kl}\right) - H\left(\sum_{kl} A_{kl}\right) \geqslant 0,$$

где $H(A) = -\operatorname{Tr} A \log A$ для положительного ненормированного оператора A. В частности, для любых положительных операторов A_1, A_2, B_1, B_2,

$$H(A_1 + A_2 + B_1 + B_2) - H(A_1 + A_2) - H(B_1 + B_2) \leqslant$$
$$\leqslant H(A_1 + B_1) - H(A_1) - H(B_1) + H(A_2 + B_2) - H(A_2) - H(B_2),$$

что равносильно совместной выпуклости функции

$$\Delta(A, B) = H(A + B) - H(A) - H(B)$$

по аргументам A и B.

С другой стороны, для операторов плотности A и B имеет место формула

$$H(A;B)=\sup_{\lambda>0}\lambda^{-1}[\Delta(\lambda A,(1-\lambda)B)+h(\lambda)]$$

(задача 36), из которой видно, что $H(A;B)$ выпукла, как верхняя грань семейства выпуклых функций.

3)⇒1). Рассмотрим состояние S_{12} в пространстве $\mathcal{H}_1\otimes\mathcal{H}_2$. Имеет место формула

$$S_1\otimes\overline{S}_2=\int(I_1\otimes U_2^*)S_{12}(I_1\otimes U_2)dU_2,$$

где dU_2 — нормированная инвариантная мера на группе унитарных операторов в $\mathcal{H}_2$ (задача 37). Совместная выпуклость (вместе с непрерывностью) относительной энтропии тогда влечет

$$H(S_1;S_1')\leqslant H(S_{12};S_{12}'),$$

т. е. монотонность относительно взятия частичного следа. Но согласно следствию 2 теоремы 12, всякий канал можно представить как суперпозицию обратимого преобразования (сохраняющего относительную энтропию) с последующим взятием частичного следа.

§ 7.2. Разложение Шмидта и очищение состояния

Рассмотрим состояние S_{12} в пространстве составной системы $\mathcal{H}_1\otimes\mathcal{H}_2$. Мы будем часто использовать следующий простой, но неожиданный результат [1]:

ЛЕММА 12. *Пусть $S_{12}=|\psi\rangle\langle\psi|$, тогда частичные состояния S_1 и S_2 имеют одинаковые ненулевые собственные значения (с учетом кратности), а следовательно, и одинаковую энтропию.*

Доказательство. Предположим без ограничения общности, что $\dim\mathcal{H}_1\leqslant\dim\mathcal{H}_2$. Рассмотрим спектральное разложение

$$S_1=\sum_j\lambda_j|e_j^1\rangle\langle e_j^1|.$$

[1] См., например, *Lindblad G.* Quantum entropy and quantum measurements // Quantum Aspects of Optical Communication. Ed. by C. Benjaballah, O. Hirota, S. Reynaud. — Lect. Notes Phys. — 1991 — V. 378. — P. 71–80.

Разложим вектор $|\psi\rangle$ по базису $|e_j^1\rangle$,

$$|\psi\rangle = \sum_j |e_j^1\rangle \otimes |h_j^2\rangle,$$

где $|h_j^2\rangle$ некоторые векторы из $\mathcal{H}_2$. Взяв частичный след, получаем

$$\mathrm{Tr}_{\mathcal{H}_2} |\psi\rangle\langle\psi| = \sum_j \lambda_j |e_j^1\rangle\langle e_j^1| = \sum_{j,k} |e_j^1\rangle\langle e_k^1|\langle h_k^2|h_j^2\rangle.$$

Следовательно, $\langle h_j^2|h_k^2\rangle = \lambda_j \delta_{jk}$. Для $\lambda_j \neq 0$ положим $|e_j^2\rangle = \lambda_j^{-1/2}|h_j^2\rangle$, и дополним до ортонормированного базиса в $\mathcal{H}_2$. Для вектора ψ имеем *разложение Шмидта*

$$|\psi\rangle = \sum_j \lambda_j |e_j^1\rangle \otimes |e_j^2\rangle.$$

Отсюда следует, что λ_j являются также собственными значениями для S_2.

Более того, из этих рассуждений вытекает, что для любого смешанного состояния S_1 в $\mathcal{H}_1$ всегда существует чистое состояние в $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ ($\mathcal{H}_1 \approx \mathcal{H}_2$), для которого $S_1 = \mathrm{Tr}_{\mathcal{H}_2} |\psi\rangle\langle\psi|$. Подобное расширение произвольного состояния до чистого называется *очищением*.

Задача 38. Докажите, что очищение квантового состояния в существенном единственно: любые два очищения с пространствами $\mathcal{H}_2$, $\mathcal{H}_2'$, такими, что $\dim \mathcal{H}_2 \leqslant \dim \mathcal{H}_2'$, связаны изометрическим вложением $\mathcal{H}_2$ в $\mathcal{H}_2'$.

Рассмотрим состояние S и неидеальное измерение (6.3), описываемое семейством операторов V_k, удовлетворяющим условию $\sum_k V_k^* V_k = I$, так что исход k получается с вероятностью $p_k = \mathrm{Tr}\, SV_k^* V_k$, приводя к апостериорному состоянию системы $S_k = V_k S V_k^* / p_k$. Имеет место следующее неравенство Гроневольда — Линдблада — Озава, связывающее энтропии априорного и апостериорных состояний [1)]

$$H(S) \geqslant \sum_k p_k H(S_k).$$

Для доказательства положим $\mathcal{H} = \mathcal{H}_1$ и рассмотрим очищение $|\psi\rangle$ состояния S в $\mathcal{H}_1 \otimes \mathcal{H}_2$. Тогда

$$p_k S_k = \mathrm{Tr}_{\mathcal{H}_2} (V_k \otimes I)|\psi\rangle\langle\psi|(V_k \otimes I)^*,$$

[1)] *Ozawa M.* On information gain by quantum measurement of continuous observable // J. Math. Phys. — 1986. — V. 27. — P. 759–763.

и $S_2 = \mathrm{Tr}_{\mathcal{H}_1} S = \sum_k p_k S_2^k$, где

$$p_k S_2^k = \mathrm{Tr}_{\mathcal{H}_1}(V_k \otimes I)|\psi\rangle\langle\psi|(V_k \otimes I)^*.$$

Используя лемму 12, а также вогнутость энтропии, получаем

$$H(S) = H(S_2) = H\Big(\sum_k p_k S_2^k\Big) \geqslant \sum_k p_k H\Big(S_2^k\Big) = \sum_k p_k H(S_k).$$

Частным случаем является следующее полезное неравенство для составной системы. Пусть S_{AB} состояние системы AB и пусть $\{|e_k^B\rangle\}$ ортонормированный базис в $\mathcal{H}_B$. Полагая $V_k = I_A \otimes |e_k^B\rangle\langle e_k^B|$, имеем

$$H(S_{AB}) \geqslant \sum_k p_k H\left(S_A^k\right), \tag{7.3}$$

где $p_k S_A^k = \langle e_k^B|S_{AB}|e_k^B\rangle$. Это неравенство может быть использовано [1] для доказательства следующего частного случая гипотезы аддитивности (см. § 6.3), именно

$$C^1\left(\Phi_A \otimes \mathrm{Id}_B\right) = C^1\left(\Phi_A\right) + C^1\left(\mathrm{Id}_B\right) = C^1\left(\Phi_A\right) + \log\dim\mathcal{H}_B \tag{7.4}$$

для произвольного канала Φ_A в $\mathcal{H}_A$. В самом деле, для произвольных вероятностей π_i и сигнальных состояний S_{AB}^i имеем

$$H\Big(\sum_i \pi_i(\Phi_A \otimes \mathrm{Id}_B)\left[S_{AB}^i\right]\Big) - \sum_i \pi_i H\left((\Phi_A \otimes \mathrm{Id}_B)\left[S_{AB}^i\right]\right) \leqslant$$
$$\leqslant H\Big(\sum_i \pi_i \Phi_A\left[S_A^i\right]\Big) - \sum_i \pi_i H\left((\Phi_A \otimes \mathrm{Id}_B)\left[S_{AB}^i\right]\right) + H\Big(\sum_i \pi_i S_B^i\Big)$$

в силу субаддитивности первого слагаемого в левой части. Последнее слагаемое в правой части не превосходит $\log\dim\mathcal{H}_B$. Применяя неравенство (7.3) к каждому члену во втором слагаемом, получаем, что сумма первых двух слагаемых в правой части не превосходит

$$H\Big(\sum_{ik} \pi_i p_{ik} \Phi_A\left[S_A^{ik}\right]\Big) - \sum_{ik} \pi_i p_{ik} H\Big(\Phi_A\left[S_A^{ik}\right]\Big),$$

где $p_{ik} S_A^{ik} = \langle e_k^B|S_{AB}^i|e_k^B\rangle$. Теперь мы имеем сигнальные состояния S_A^{ik} с вероятностями $\pi_i p_{ik}$ на входе канала Φ_A, так что последнее выражение не превосходит $C^1\left(\Phi_A\right)$.

---

[1] *Schumacher B., Westmoreland M. D.* Optimal signal ensembles // LANL Report quant-ph/9912122.

§ 7.3. Энтропийная корреляция и условная энтропия

В классике количество информации, передаваемое каналом $X \to Y$, измеряется величиной $I(X, Y) = H(X) + H(Y) - H(X, Y)$. Рассмотрим сначала формальный квантовый аналог этой величины. В квантовой статистике далеко не всегда можно говорить о совместном распределении. Поэтому рассмотрим ситуацию, в которой оно заведомо существует. Пусть имеются две системы в гильбертовом пространстве $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ в совместном состоянии S_{12}, и пусть $S_1 = \mathrm{Tr}_2 S_{12}$, $S_2 = \mathrm{Tr}_1 S_{12}$ частичные состояния. Определим *энтропийную корреляцию* между подсистемами как

$$c_{12} = H(S_1) + H(S_2) - H(S_{12}).$$

Лемма 13. *Энтропийная корреляция неотрицательна,* $c_{12} \geqslant 0$. *Причем* $c_{12} = 0$ *тогда и только тогда, когда* $S_{12} = S_1 \otimes S_2$.

Доказательство. Имеем

$$H(S_1) + H(S_2) - H(S_{12}) = H(S_{12}; S_1 \otimes S_2) \geqslant 0,$$

где $H(\cdot\,;\,\cdot)$ — относительная энтропия.

Рассмотрим случай, когда совместное состояние — чистое, $S_{12} = |\psi\rangle\langle\psi|$. Поскольку $H(S_{12}) = 0$, то в силу леммы 12 имеем

$$c_{12} = H(S_1) + H(S_2) = 2H(S_1).$$

Величина $H(S_1)$ является естественной мерой сцепленности чистого состояния $S_{12} = |\psi\rangle\langle\psi|$. Вопрос о мерах сцепленности смешанного состояния является куда более сложным и его изучению посвящена обширная и быстро растущая литература (см. [13]). Скажем только, что для смешанного состояния можно определить разные меры сцепленности, соотношение между которыми еще не вполне изучено. Кроме того, для смешанных состояний существуют качественно различные степени сцепленности (появляется так называемая связанная сцепленность (bound entanglement), отсутствующая для чистых состояний).

Заметим, что в классической статистике частичное (маргинальное) состояние для любого чистого состояния обязательно чистое, так что очищение не имеет классического аналога. С этим обстоятельством связано необычное свойство *квантовой условной энтропии.* Классическая условная энтропия

$H(X|Y) = H(XY) - H(Y)$ всегда неотрицательна, т. е. расширение классической системы не может привести к уменьшению энтропии состояния системы. Квантовый аналог условной энтропии $H(A \mid B) = H(S_{AB}) - H(S_B)$, где S_{AB} — совместное состояние квантовых систем A, B, а S_B — частичное состояние системы B, может принимать отрицательные значения, поскольку при очищении энтропия составной системы может быть меньше энтропии подсистемы. Заметим, однако, что квантовая условная энтропия, как и классическая, обладает важным свойством монотонности

$$H(A \mid BC) \leqslant H(A \mid B),$$

которое, очевидно, эквивалентно сильной субаддитивности энтропии (задача 39).

Лемма 14. *Условная энтропия $H(A|B)$ является вогнутой функцией совместного состояния S_{AB}.*

Доказательство. Имеем

$$H(A \mid B) = -H(S_{AB}; d_A^{-1} I_A \otimes S_B) + \log d_A.$$

Утверждение теперь следует из совместной выпуклости относительной энтропии.

Эти свойства делают квантовую условную энтропию полезным аналитическим инструментом, несмотря на отсутствие удовлетворительной статистической интерпретации.

§ 7.4. Обменная энтропия

Пусть $\mathcal{H}$ — гильбертово пространство, S — состояние. Рассмотрим квантовый канал в $\mathcal{H}$, т. е. вполне положительное отображение, сохраняющее след. Индексом $_Q$ будем обозначать входные состояния, а значком $_{Q'}$ выходные состояния системы, так что

$$\Phi\colon S_Q \longrightarrow \Phi[S_Q] = S_{Q'}.$$

Введем еще одну систему — окружение (environment), обозначаемую значком $_E$, находящуюся в исходном чистом состоянии $|\psi_E\rangle\langle\psi_E|$. Канал Φ всегда можно предоставить в виде

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

где U унитарный оператор, описывающий эволюцию составной системы.

Обменной энтропией называется величина $H(S_{E'})$, равная приращению энтропии окружения в результате взаимодействия с системой. Здесь $S_{E'}$ состояние окружения после взаимодействия,

$$S_{E'} = \mathrm{Tr}_{\mathcal{H}_Q} U(S_Q \otimes S_E)U^*. \tag{7.5}$$

Можно показать, что это определение зависит лишь от S_Q и Ф, но не зависит от способа представления канала.

Введем еще эталонную систему (reference system), обозначаемую индексом $_R$, с пространством $\mathcal{H}_R \approx \mathcal{H}_Q$. Эта система нужна для очищения состояния S_Q:

$$S_Q \longrightarrow |\psi_{QR}\rangle\langle\psi_{QR}| = S_{QR}, \quad H(S_Q) = H(S_R).$$

По определению, эталонная система не изменяется в ходе эволюции, так что $S_{R'} = S_R$. Рассмотрим S_{QR} как входное состояние канала $\Phi \otimes \mathrm{Id}_R$, и обозначим выходное состояние этого канала $S_{Q'R'} = (\Phi \otimes \mathrm{Id}_R)[S_{QR}]$.

ТЕОРЕМА 15. *Обменная энтропия дается соотношениями*

$$H(S_{E'}) = H(S_{Q'R'}) = H([\mathrm{Tr}\, S_Q V_k^* V_j]_{jk}), \tag{7.6}$$

где V_j операторы из разложения Крауса

$$\Phi(S) = \sum_{j=1}^{N} V_j S V_j^*.$$

Доказательство. Рассмотрим исходное чистое состояние

$$S_{QRE} = |\psi_{QR}\rangle\langle\psi_{QR}| \otimes |\psi_E\rangle\langle\psi_E|.$$

В результате действия канала вновь получим чистое состояние

$$S_{Q'R'E'} = (U \otimes I_R)S_{QRE}(U \otimes I_R)^*.$$

По лемме 12 получаем, что $H(S_{Q'R'}) = H(S_{E'})$. Отсюда видно также, что обменная энтропия не зависит от способа унитарного расширения канала.

Для доказательства второго равенства заметим, что

$$S_{Q'R'} = \sum_{j=1}^{N} (V_j \otimes I_R)|\psi_{QR}\rangle\langle\psi_{QR}|(V_j \otimes I_R)^*.$$

Рассмотрим матрицу

$$\left[(V_j \otimes I_R)|\psi_{QR}\rangle\langle\psi_{QR}|(V_k \otimes I_R)^*\right]_{jk}.$$

Это матрица плотности чистого состояния $|\psi\rangle\langle\psi|$ с вектором

$$|\psi\rangle = [(V_j \otimes I_R)|\psi_{QR}\rangle]_{j=1,\ldots,N} \in \oplus_{j=1}^{N} \mathcal{H}_{QR} \approx \mathcal{H}_{QR} \otimes \mathcal{H}_E,$$

где $\mathcal{H}_E$ — гильбертово пространство размерности N. Тогда можно считать,что $S_{Q'R'} = \mathrm{Tr}_{\mathcal{H}_E} |\psi\rangle\langle\psi|$. Следовательно, согласно лемме 12, получаем

$$H(S_{Q'R'}) = H(\mathrm{Tr}_{\mathcal{H}_E} |\psi\rangle\langle\psi|) = H(\mathrm{Tr}_{\mathcal{H}_{QR}} |\psi\rangle\langle\psi|) =$$
$$= H(\left[\langle\psi_{QR}|(V_k \otimes I_R)^*(V_j \otimes I_R)|\psi_{QR}\rangle\right]_{jk}) = H(\mathrm{Tr}\left[S_Q V_k^* V_j\right]_{jk}).$$

§ 7.5. Информационные количества

В классической теории информации взаимная информация между входом и выходом дается формулой Шеннона

$$I = H(X) + H(Y) - H(XY) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X),$$

где $N = H(Y \mid X)$ — шум, $L = H(X \mid Y)$ — потери в канале связи (см. рис. 12). В частности, сумма шума и потерь всегда равна энтропии источника, $I + L = H(X)$. В квантовом случае определить прямой аналог величины I не представляется возможным, поскольку не определено совместное состояние входа–выхода. Однако, если мы введем эталонную систему, которая не изменяет своего состояния в течение эволюции, ее можно рассматривать как некоторый заменитель входа в классическом случае. Напомним, что $H(S_{R'}) = H(S_R) = H(S_Q)$. Кроме того, предполагая, что начальное состояние окружения чистое, получаем, что состояние системы QRE остается чистым в ходе эволюции и поэтому

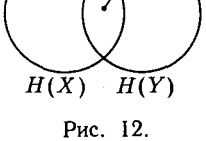


Рис. 12. Классическая диаграмма Венна

$$H(S_{Q'R'}) = H(S_{E'}), \quad H(S_{Q'E'}) = H(S_{R'}), \quad H(S_{E'R'}) = H(S_{Q'}).$$

Формальный аналог взаимной информации дается энтропийной корреляцией между выходом и эталонной системой:

$$I(Q';R') = H(S_{Q'}) + H(S_{R'}) - H(S_{Q'R'}) = \\ = H(S_{Q'}) + H(S_Q) - H(S_{E'}). \quad (7.7)$$

Естественно рассмотреть также энтропийные корреляции между выходом и окружением и между окружением и эталонной системой

$$I(Q';E') = H(S_{Q'}) + H(S_{E'}) - H(S_{Q'E'}) = \\ = H(S_{Q'}) + H(S_{E'}) - H(S_Q) = N,$$

$$I(R';E') = H(S_{R'}) + H(S_{E'}) - H(S_{R'E'}) = \\ = H(S_Q) + H(S_{E'}) - H(S_{Q'}) = L,$$

которые по своему физическому смыслу естественно ассоциировать с «шумом» и «потерями», соответственно. В квантовом случае их сумма равна удвоенной энтропии источника, $I+L=2H(S_Q)$. Отметим, что начальные значения этих трех корреляций суть

$$I(Q;R) = H(S_Q) + H(S_R) - H(S_{QR}) = 2H(S_Q),$$
$$I(Q;E) = H(S_Q) + H(S_E) - H(S_{QE}) = 0,$$
$$I(R;E) = H(S_R) + H(S_E) - H(S_{RE}) = 0.$$

Таким образом, все эти информационные количества выражаются через *входную* $H(S_Q)$, *выходную* $H(S_{Q'})$ и *обменную* $H(S_{E'})$ энтропии, и вследствие этого являются функциями только канала Φ и начального состояния $S = S_Q$. Поэтому мы вправе ввести обозначения $I = I(S,\Phi)$, $L = L(S,\Phi)$, $N = N(S,\Phi)$. Поскольку эти величины неотрицательны, входная, выходная и обменная энтропии удовлетворяют неравенствам для сторон треугольника:

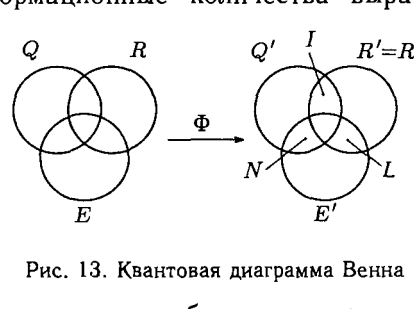


Рис. 13. Квантовая диаграмма Венна

$$H(S) - H(\Phi[S]) \leqslant H(S,\Phi), \quad (7.8)$$

и т. д.

Графическое представление всех этих величин дается квантовыми диаграммами Венна (см. рис. 13), которые, однако, отличаются

от классических тем, что области, отвечающие условным энтропиям, например

$$H(Q \mid R) = H(S_{QR}) - H(S_Q)$$

могут иметь отрицательную площадь.

Взаимная информация $I(S, \Phi)$ обладает целым рядом хороших свойств, аналогичных свойствам шенноновской информации:

1) вогнутость по S;

2) выпуклость по Φ;

3) субаддитивность: $I(S_{12}, \Phi_1 \otimes \Phi_2) \leqslant I(S_1, \Phi_1) + I(S_2, \Phi_2)$;

4) цепные свойства:

а) $I(S, \Phi_2 \cdot \Phi_1) \leqslant I(S, \Phi_1)$;

б) $I(S, \Phi_2 \cdot \Phi_1) \leqslant I(\Phi_1[S], \Phi_2)$.

Доказательства этих свойств основываются, главным образом, на лемме 12 и свойстве сильной субаддитивности квантовой энтропии [1)].

1) Используя определение (7.7) и лемму 12, имеем

$$I(S, \Phi) = H(S_{Q'E'}) + H(S_{Q'}) - H(S_{E'}) = H(S_{Q'}) + H(Q'|E'). \quad (7.9)$$

Отображение $S_Q \to S_{Q'}$ аффинно, а энтропия $H(S_{Q'})$ вогнута как функция $S_{Q'}$. С другой стороны, отображение $S_{Q'} \to S_{Q'E'}$ аффинно, а условная энтропия $H(Q'|E')$ вогнута согласно лемме 14.

2) С другой стороны,

$$I(S, \Phi) = H(S_Q) + H(S_{Q'}) - H(S_{Q'R'}) = H(S_Q) - H(R' \mid Q').$$

Здесь первое слагаемое вообще не зависит от Φ, а второе, согласно той же лемме, является выпуклой функцией от состояния $S_{Q'R'}$, которое аффинно зависит от Φ.

3) Для доказательства субаддитивности вновь используем представление (7.9). Тогда нам надо доказать, что

$$H(S'_{12}) + H(Q'_1 Q'_2 \mid E'_1 E'_2) \leqslant H(S'_1) + H(Q'_1 \mid E'_1) + H(S'_2) + H(Q'_2 \mid E'_2).$$

Но это вытекает из субаддитивности энтропии и условной энтропии. Последнее означает, что

$$H(Q'_1 Q'_2 \mid E'_1 E'_2) \leqslant H(Q'_1 \mid E'_1) + H(Q'_2 \mid E'_2),$$

[1)] *Adami C., Cerf N. J.* Capacity of noisy quantum channels // Phys. Rev. A — 1997. — V. A56. — P. 3470–3485; LANL Report no. quant-ph/9609024.

и доказывается путем двойного использования свойства сильной субаддитивности (задача 40).

4) Доказательства цепных свойств оставляются в качестве упражнений. Они основываются на свойстве монотонности условной энтропии:

$$H(A \mid BC) \leqslant H(A \mid B),$$

которое также вытекает из сильной субаддитивности квантовой энтропии.

ГЛАВА 8

ПЕРЕДАЧА КЛАССИЧЕСКОЙ ИНФОРМАЦИИ С ПОМОЩЬЮ СЦЕПЛЕННОГО СОСТОЯНИЯ

Изложенный выше феноменологический подход не раскрывает операциональный смысл величины $I(S,\Phi)$ и ей подобных, как теорема кодирования раскрывала смысл шенноновской информации. Замечательно, что эта величина действительно асимптотически характеризует некоторый протокол передачи информации через квантовый канал связи. Этот протокол, называемый передачей классической информации с помощью сцепленного состояния (entanglement-assisted classical communication) является прямым обобщением идеи сверхплотного кодирования на случай произвольного канала с шумом.

Пусть имеется система связи с передающим концом A и приемным концом B. (Предполагается, что $\mathcal{H}_A \approx \mathcal{H}_B$.) Таким образом, AB описывается гильбертовым пространством $\mathcal{H}_A \otimes \mathcal{H}_B$. Перед началом связи с помощью некоторой процедуры типа ЭПР система AB приготовляется в некотором чистом состоянии S_{AB}. Система A получает классическую информацию, содержащуюся в значениях параметра x, которая кодируется произвольными преобразованиями Φ_x в пространстве $\mathcal{H}_A$. В случае сверхплотного кодирования мы ограничились унитарными преобразованиями, но обобщение на произвольные преобразования очевидно. При этом состояние системы AB преобразуется в $(\Phi_x \otimes \mathrm{Id}_B)[S_{AB}]$. После этого A посылает свою часть состояния по данному каналу Φ, в результате чего на конце B становится доступной измерению система AB в состоянии $(\Phi\cdot\Phi_x\otimes\mathrm{Id}_B)[S_{AB}]$. 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 $\mathcal{H}_{AB}$.

Это аналогично сверхплотному кодированию (см. § 3.2), но вместо идеального канала, A использует канал Φ. Нас интересует классическая пропускная способность этого протокола, которая называется *пропускной способностью протокола передачи*

классической информации с помощью сцепленного состояния (*entanglement-assisted classical capacity или EACC*).

Максимум по измерениям B может быть оценен с помощью теоремы кодирования для классической пропускной способности из гл. 5. Сначала введем величину

$$C_{ea}^{(1)}(\Phi) = \\ = \max_{\pi_i,\, \mathcal{E}_A^i,\, S_{AB}} \left[H\left(\sum_i \pi_i (\Phi \otimes \mathrm{Id}_B) \left[S_{AB}^i \right] \right) - \sum_i \pi_i H\left((\Phi \otimes \mathrm{Id}_B) \left[S_{AB}^i \right] \right) \right].$$

Используя канал n раз и коллективные измерения в системе B, получаем величину

$$C_{ea}^{(n)}(\Phi) = C_{ea}^{(1)}(\Phi^{\otimes n}).$$

EACC тогда равна

$$C_{ea}(\Phi) = \lim_{n \to \infty} \frac{1}{n} C_{ea}^{(1)}(\Phi^{\otimes n}).$$

ТЕОРЕМА 16[1]. *Пропускная способность протокола передачи классической информации с помощью сцепленного состояния дается формулой*

$$C_{ea} = \max_S I(S, \Phi).$$

Очевидно, что всегда $C_{ea} \geqslant C$, причем $C_{ea} = 2C^1 = 2C$ для идеального канала (случай сверхплотного кодирования, см. § 3.2). Величину C_{ea} естественно сравнивать с классической пропускной способностью канала C, однако, поскольку гипотеза $C = C^1$ остается открытой (см. § 6.3), приходится рассматривать C^1.

Вычислим величину C_{ea} для деполяризующего канала (6.6). Используя унитарную ковариантность деполяризующего канала и вогнутость функции $S \to I(S, \Phi)$, получаем, что ее максимум достигается на хаотическом состоянии $\bar{S} = I/d$. Имеем

$$H(\bar{S}) = H(\Phi[\bar{S}]) = \log d.$$

Обменная энтропия $H(\bar{S}, \Phi)$ может быть вычислена по формуле (7.6) (задача 41), откуда получаем

$$C_{ea} = \log d^2 + \left(1 - p\frac{d^2-1}{d^2}\right) \log \left(1 - p\frac{d^2-1}{d^2}\right) + p\frac{d^2-1}{d^2} \log \frac{p}{d^2}.$$

[1] Этот результат был анонсирован в работе *Bennett C. H., Shor P. W., Smolin J. A., Thapliyal A. V.* Entanglement-assisted classical capacity of noisy quantum channel // Phys. Rev. Lett. — 1999. — V. 83. — P. 3081; LANL Report quant-ph/9904023. Мы приводим усовершенствованную и дополненную версию доказательства из неопубликованной работы: *Bennett C. H., Shor P. W., Smolin J. A., Thapliyal A. V.* Entanglement-assisted capacity and the reverse Shannon theorem. — 2000.

Сравнивая с величиной C^1, которая дается формулой (6.8), получаем, что $C_{ea}/C^1 \to d+1$ при $p \to 1$ (когда обе пропускные способности стремятся к нулю!) 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 не могут эту информацию передавать.

Доказательство теоремы. Докажем неравенство

$$C_{ea}(\Phi) \geqslant \max_S I(S, \Phi). \tag{8.1}$$

Сначала построим обобщение протокола сверхплотного кодирования (§ 3.2), которое позволит установить неравенство

$$C_{ea}^{(1)}(\Phi^{\otimes n}) \geqslant I\left(\frac{P}{\dim P}, \Phi^{\otimes n}\right) \tag{8.2}$$

для произвольного проектора P в $\mathcal{H}_A^{\otimes n}$. Пусть $P = \sum_{k=1}^{m} |e_k\rangle\langle e_k|$, где $\{e_k;\ k = 1, \ldots, m = \dim P\}$ ортонормированная система. Определим унитарные операторы в $\mathcal{H}_A$, действующие по формулам

$$V|e_k\rangle = \exp\left(\frac{2\pi i k}{m}\right)|e_k\rangle; \quad U|e_k\rangle = |e_{k+1(\mathrm{mod}\ m)}\rangle; \quad k = 1, \ldots, m;$$

$$W_{\alpha\beta} = U^\alpha V^\beta; \qquad \alpha, \beta = 1, \ldots, m,$$

на подпространстве, порожденном векторами $\{e_k\}$, и как единичный оператор — на ортогональном дополнении. Пусть

$$|\psi_{AB}\rangle = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |e_k\rangle \otimes |e_k\rangle.$$

Задача 42. Система векторов

$$\left(W_{\alpha\beta} \otimes I_B\right)|\psi_{AB}\rangle; \quad \alpha, \beta = 1, \ldots, m$$

является ортонормированной в $\mathcal{H}_A \otimes \mathcal{H}_B$; в частности, если $m = \dim \mathcal{H}_A$, то эта система — базис. Докажите также, что

$$\sum_{\alpha,\beta=1}^{m} \left(W_{\alpha\beta} \otimes I_B\right)|\psi_{AB}\rangle\langle\psi_{AB}|\left(W_{\alpha\beta} \otimes I_B\right)^* = P \otimes P. \tag{8.3}$$

Таким образом, операторы $\{W_{\alpha\beta};\ \alpha, \beta = 1, \ldots, m\}$ играют роль, аналогичную матрицам Паули в протоколе сверхплотного кодирования для q-бита. Пусть теперь классический сигнал принимает

значения $x=(\alpha,\beta)$ с равными вероятностями $1/m^2$. В качестве сцепленного состояния возьмем $|\psi_{AB}\rangle\langle\psi_{AB}|$, а в качестве кодирующих отображений — $\mathcal{E}_A^i[S]=W_{\alpha\beta}SW_{\alpha\beta}^*$. Тогда получим

$$C_{ea}^{(1)}(\Phi^{\otimes n})\geqslant H\Big(\frac{1}{m^2}\sum_{\alpha\beta}(\Phi\otimes \mathrm{Id}_B)[S_{AB}^{\alpha\beta}]\Big)-\frac{1}{m^2}\sum_{\alpha\beta}H\left((\Phi\otimes \mathrm{Id}_B)[S_{AB}^{\alpha\beta}]\right),$$

где $S_{AB}^{\alpha\beta}=\left(W_{\alpha\beta}\otimes I_B\right)|\psi_{AB}\rangle\langle\psi_{AB}|\left(W_{\alpha\beta}\otimes I_B\right)^*$. Тогда согласно (8.3) первое слагаемое в правой части равно

$$H\Big((\Phi\otimes \mathrm{Id}_B)\Big[\frac{P}{m}\otimes\frac{P}{m}\Big]\Big)=H\Big(\frac{P}{m}\Big)+H\Big(\Phi\Big[\frac{P}{m}\Big]\Big).$$

Поскольку состояние $S_{AB}^{\alpha\beta}$ очищает смешанное состояние $\frac{P}{m}$ в $\mathcal{H}_B$, 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 (8.2). 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 (7.5).

Пусть теперь $S_A=S$ произвольное состояние в $\mathcal{H}_A$, и пусть $P^{n,\delta}$ проектор на типичное подпространство оператора плотности $S^{\otimes n}$ в $\mathcal{H}_A^{\otimes n}$. В работе Беннета и др. [1] было высказано предположение, что для произвольного канала Ψ выполнено равенство

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}H\Big(\Psi^{\otimes n}\Big(\frac{P^{n,\delta}}{\dim P^{n,\delta}}\Big)\Big)=H(\Psi(S)),$$

откуда, согласно выражениям для взаимной информации и обменной энтропии, следовало бы

$$\lim_{\delta\to 0}\lim_{n\to\infty}\frac{1}{n}I\Big(\frac{P^{n,\delta}}{\dim P^{n,\delta}},\Phi^{\otimes n}\Big)=I(S,\Phi),\qquad(8.4)$$

а значит, согласно (8.2), и доказываемое неравенство (8.1). Мы докажем, что соотношение (8.4) действительно имеет место, если под $P^{n,\delta}$ понимать проектор на *сильно типичное подпространство состояния* $S^{\otimes n}$.

Зафиксируем положительное δ, и обозначим через λ_j собственные числа, $|e_j\rangle$ собственные векторы оператора плотности S. Тогда

[1] См. там же.

собственные числа и собственные векторы оператора плотности $S^{\otimes n}$ суть $\lambda_J = \lambda_{j_1} \cdot \ldots \cdot \lambda_{j_n}$, $|e_J\rangle = |e_{j_1}\rangle \otimes \ldots \otimes |e_{j_n}\rangle$, где $J = (j_1, \ldots, j_n)$. Последовательность J называется *сильно типичной* [9], если числа $N(j \mid J)$ появлений символов j в последовательности J удовлетворяют условию

$$\left| \frac{N(j|J)}{n} - \lambda_j \right| < \delta, \quad j = 1, \ldots, d,$$

причем $N(j \mid J) = 0$, если $\lambda_j = 0$. Обозначим совокупность всех сильно типичных последовательностей $B^{n,\delta}$, и пусть P^n — распределение вероятностей, задаваемое собственными числами λ_J. Согласно закону больших чисел, $\mathsf{P}^n\left(B^{n,\delta}\right) \to 1$ при $n \to \infty$. Известно [9], что размер множества $B^{n,\delta}$ оценивается как

$$2^{n[H(S) - \Delta_n(\delta)]} < |B^{n,\delta}| < 2^{n[H(S) + \Delta_n(\delta)]}, \tag{8.5}$$

где $H(S) = -\sum_{j=1}^{d} \lambda_j \log \lambda_j$, и $\lim_{\delta \to 0} \lim_{n \to \infty} \Delta_n(\delta) = 0$.

Для произвольной функции $f(j)$, $j = 1, \ldots, d$, и последовательности $J = (j_1, \ldots, j_n) \in B^{n,\delta}$ имеем

$$\left| \frac{f(j_1) + \ldots + f(j_n)}{n} - \sum_{j=1}^{d} \lambda_j f(j) \right| < \delta \max f. \tag{8.6}$$

В частности, любая сильно типичная последовательность является типичной в обычном (энтропийном) смысле: полагая $f(j) = -\log \lambda_j$, имеем

$$n[H(S) - \delta_1] < -\log \lambda_J < n[H(S) + \delta_1], \tag{8.7}$$

где $\delta_1 = \delta \max_{\lambda_j > 0}(-\log \lambda_j)$. Обратное неверно — не всякая типичная последовательность сильно типична.

Проектор на сильно типичное подпространство определим формулой

$$P^{n,\delta} = \sum_{J \in B^{n,\delta}} |e_J\rangle\langle e_J|.$$

Обозначим $d_{n,\delta} = \dim P^{n,\delta} = |B^{n,\delta}|$ и $\bar{S}^{n,\delta} = \frac{P^{n,\delta}}{d_{n,\delta}}$, и докажем, что

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{n} H(\Psi^{\otimes n}\left(\bar{S}^{n,\delta}\right)) = H(\Psi(S)) \tag{8.8}$$

для произвольного канала Ψ.

Имеем

$$nH(\Psi(S)) - H(\Psi^{\otimes n}(\bar{S}^{n,\delta})) = H(\Psi(S)^{\otimes n}) - H(\Psi^{\otimes n}(\bar{S}^{n,\delta})) =$$
$$= H(\Psi^{\otimes n}(\bar{S}^{n,\delta}); \Psi^{\otimes n}(S^{\otimes n})) + \operatorname{Tr}\log\Psi(S)^{\otimes n}(\Psi^{\otimes n}(\bar{S}^{n,\delta}) - \Psi(S)^{\otimes n}), \tag{8.9}$$

где $H(\cdot\,;\,\cdot)$ относительная энтропия. 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

$$H(\Psi^{\otimes n}(\bar{S}^{n,\delta}); \Psi^{\otimes n}(S^{\otimes n})) \leqslant H(\bar{S}^{n,\delta}; S^{\otimes n}),$$

где правая часть

$$H(\bar{S}^{n,\delta}; S^{\otimes n}) = \sum_{J\in B^{n,\delta}} \frac{1}{d_{n,\delta}} \log \frac{1}{d_{n,\delta}\lambda_J} = -\log d_{n,\delta} - \sum_{J\in B^{n,\delta}} \frac{1}{d_{n,\delta}} \log \lambda_J,$$

согласно (8.7), (8.5) оценивается величиной $n(\delta_1 + \Delta_n(\delta))$, стремящейся к нулю при $n\to\infty$, $\delta\to 0$.

Используя соотношение

$$\log\Psi(S)^{\otimes n} = \log\Psi(S)\otimes I\otimes\ldots\otimes I + \ldots + I\otimes\ldots\otimes I\otimes\log\Psi(S),$$

и вводя оператор $F = \Psi^*(\log\Psi(S))$, мы можем переписать второе слагаемое в виде

$$n\operatorname{Tr}\frac{(F\otimes I\otimes\ldots\otimes I + \ldots + I\otimes\ldots\otimes I\otimes F)}{n}(\bar{S}^{n,\delta} - S^{\otimes n}) =$$
$$= \frac{n}{d_{n,\delta}} \sum_{J\in B^{n,\delta}} \left[\frac{f(j_1)+\ldots+f(j_n)}{n} - \sum_{j=1}^{d} \lambda_j f(j)\right],$$

где $f(j) = \langle e_j|F|e_j\rangle$, что оценивается величиной $n\delta\max f$ в силу (8.6). Это доказывает соотношение (8.8) 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 $\Psi^{\otimes n}(\bar{S}^{n,\delta})$ 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 (8.9) следует понимать как

$$\operatorname{Tr} P_\Psi^{\otimes n}\log\left[P_\Psi^{\otimes n}\Psi(S)^{\otimes n}P_\Psi^{\otimes n}\right]P_\Psi^{\otimes n}(\Psi^{\otimes n}(\bar{S}^{n,\delta}) - \Psi(S)^{\otimes n}),$$

где логарифм берется от невырожденного оператора в подпространстве $P_\Psi^{\otimes n}\mathcal{H}_A^{\otimes n}$. Теперь можно повторить рассуждения, определив F как $\Psi^*(P_\Psi\left[\log P_\Psi\Psi(S)P_\Psi\right]P_\Psi)$. Это завершает доказательство соотношения (8.4), откуда вытекает (8.1).

Теперь докажем неравенство

$$C_{ea}(\Phi) \leqslant \max_S I(S,\Phi). \tag{8.10}$$

Докажем сначала, что

$$C_{ea}^{(1)}(\Phi) \leqslant \max_S I(S,\Phi). \tag{8.11}$$

Обозначим состояние системы AB (соотв. A) после кодирования

$$S_{AB}^x = (\Phi_x \otimes \mathrm{Id}_B)[S_{AB}], \quad \text{соответственно} \quad S_A^x = \Phi_x[S_A]. \tag{8.12}$$

Заметим, что частичное состояние B не изменяется после кодирования, $S_B^x = S_B$. Докажем, что

$$H\left(\sum_x \pi_x(\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^x]\right) - \sum_x \pi_x H\left((\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^x]\right) \leqslant$$
$$\leqslant I\left(\sum_x \pi_x S_A^x; \Phi_A\right). \tag{8.13}$$

Согласно квантовой теореме кодирования, максимум левой части по всем π_x, Φ_x есть $C_{ea}^{(1)}(\Phi)$, откуда будет следовать (8.11).

Используя субаддитивность квантовой энтропии, мы можем оценить первое слагаемое в левой части (8.13) как

$$H\left(\sum_x \pi_x \Phi_A[S_A^x]\right) + H(S_B) = H\left(\Phi_A\left[\sum_x \pi_x S_A^x\right]\right) + \sum_x \pi_x H(S_B).$$

Первый член уже дает выходную энтропию из $I\left(\sum_x \pi_x S_A^x; \Phi_A\right)$. Оценим остальные члены

$$\sum_x \pi_x \left[H(S_B) - H\left((\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^x]\right)\right].$$

Покажем сначала, что слагаемое в квадратных скобках не превосходит

$$H(S_A^x) - H\left((\Phi_A \otimes \mathrm{Id}_{R^x})[S_{AR^x}^x]\right),$$

где R^x эталонная система для очищения S_A^x, и $S_{AR^x}^x$ соответствующее чистое состояние. С этой целью рассмотрим унитарное

расширение кодирования Φ_x с окружением E_x, которое находится в начальном чистом состоянии, в соответствии со следствием 2 теоремы 12. Из (8.12) видно, что можно считать $R^x = BE_x$ (после взаимодействия, которое затрагивает только AE_x). Тогда, обозначая штрихом состояния после применения канала Φ_A, получаем

$$H(S_B) - H\left((\Phi_A \otimes \mathrm{Id}_B)[S_{AB}^x]\right) = \\ = H(S_B) - H(S_{A'B}^x) = -H_x(A' \mid B), \quad (8.14)$$

где нижний индекс x условной энтропии указывает на состояние $S_{A'B}^x$. Аналогично

$$H(S_A^x) - H\left((\Phi_A \otimes \mathrm{Id}_{R^x})[S_{AR^x}^x]\right) = H(S_{R^x}^x) - H\left(S_{A'R^x}^x\right) = \\ = -H_x(A'|R^x) = -H_x(A'|BE_x),$$

что больше или равно (8.14), 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

$$\sum_x \pi_x \left[H(S_A^x) - H\left((\Phi_A \otimes \mathrm{Id}_{R^x})[S_{AR^x}^x]\right)\right] \leqslant \\ \leqslant H\left(\sum_x \pi_x S_A^x\right) - H\left((\Phi_A \otimes \mathrm{Id}_R)[\widehat{S}_{AR}]\right),$$

где $\widehat{S}_{AR}$ — состояние, очищающее $\sum_x \pi_x S_A^x$ 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 $S_{A'E'}$, а отображение $S_A \to S_{A'E'}$ аффинно, поэтому $H(A' \mid E')$ — вогнутая функция состояния S_A.

Применяя это рассуждение к каналу $\Phi^{\otimes n}$, получаем

$$C_{ea}^{(n)}(\Phi) \leqslant \max_{S_A} I(S_A, \Phi^{\otimes n}).$$

Согласно субаддитивности квантовой взаимной информации (свойство 3)), 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,

$$C_{ea}(\Phi) \leqslant \max_{S_A} I(S_A, \Phi). \tag{8.15}$$

ГЛАВА 9

КВАНТОВАЯ ПРОПУСКНАЯ СПОСОБНОСТЬ И КОГЕРЕНТНАЯ ИНФОРМАЦИЯ

Пусть S — входное состояние, на которое действует фиксированный канал Φ: $S \longrightarrow \Phi[S]$. Теперь наша цель состоит в том, чтобы используя блочное кодирование, т. е. n-е степени канала $\Phi^{\otimes n}$, добиться (асимптотически при $n \to \infty$) точной передачи квантовых состояний с максимально высокой скоростью. Таким образом, речь идет об аналоге шенноновской теоремы кодирования для квантовой информации [1].

§ 9.1. Точность воспроизведения квантовой информации

В качестве меры точности воспроизведения состояния можно было бы использовать, например, величину $\max_S \| S - \Phi[S] \|_1$, где $\| A \|_1 = \mathrm{Tr} |A|$, и $|A| = \sqrt{A^* A}$, показывающую, насколько выход канала Φ отличается от выхода идеального канала (ср. с (1.16) в классическом случае). Однако в квантовой теории информации и компьютинге традиционно используются другие меры точности. Дадим их описание и рассмотрим взаимосвязи между ними. Докажем полезный вспомогательный результат.

ЛЕММА 15. *Пусть ψ единичный вектор, S произвольное состояние, тогда имеют место неравенства*

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

---

[1] В этой главе мы кратко излагаем некоторые результаты работ: *Barnum H., Nielsen M. A., Schumacher B.* Information transmission through noisy quantum channels // Phys. Rev. A. — 1998. — V. 57. — P. 4153–4175; LANL Report no. quant-ph/9702049, Feb. 1997; *Barnum H., Knill E., Nielsen M. A.* On quantum fidelities and channel capacities // IEEE Trans. Inform. Theory. — 2000. — V. 46. — P. 1317–1329; LANL Report no. quant-ph/9809010, Sep. 1998.

Если $S = |\varphi\rangle\langle\varphi|$ — чистое состояние, то второе неравенство обращается в равенство, т. е.

$$\| |\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi| \|_1 = 2\sqrt{1 - |\langle\psi|\varphi\rangle|^2}. \tag{9.1}$$

Величина

$$F(|\psi\rangle\langle\psi|, S) = \langle\psi| S |\psi\rangle \tag{9.2}$$

называется точностью воспроизведения (fidelity). Это понятие можно обобщить на случай, когда оба сравниваемые состояния — смешанные, при этом точность воспроизведения выражается через так называемое расстояние Бюреса, однако это нам здесь не понадобится.

Доказательство. 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.

Доказательство соотношения (9.1) предлагается в качестве упражнения. Для вычисления левой части достаточно найти собственные значения оператора $|\psi\rangle\langle\psi| - |\varphi\rangle\langle\varphi|$, который имеет ранг 2 (см. задачу 12).

Пусть теперь S произвольный оператор плотности, рассмотрим его спектральное разложение $S = \sum \lambda_i S_i$. Используя выпуклость нормы, вогнутость функции $\sqrt{}$, а также (9.1), получаем

$$\| |\psi\rangle\langle\psi| - S \|_1 \leqslant \sum_i \lambda_i \| |\psi\rangle\langle\psi| - S_i \|_1 =$$
$$= 2\sum_i \lambda_i \sqrt{1 - \langle\psi| S_i |\psi\rangle} \leqslant 2\sqrt{1 - \langle\psi| S |\psi\rangle}.$$

Заметим, что когда мы рассматривали вопрос о сжатии квантовой информации, мы фактически использовали меру точности (9.2). Если состояние $|\psi_i\rangle\langle\psi_i|$ возникает с вероятностью p_i, то средняя точность воспроизведения равна $F = \sum_i p_i \langle\psi_i| S_i |\psi_i\rangle$.

В связи с кодами, исправляющими ошибки (см. гл. 10), естественная мера точности определяется следующим образом. Пусть задано подпространство $\mathcal{L} \subset \mathcal{H}$. Точностью воспроизведения этого подпространства каналом Φ называется величина

$$F_s(\mathcal{L}, \Phi) = \min_{\psi \in \mathcal{L}, \|\psi\| = 1} \langle\psi| \Phi[\, |\psi\rangle\langle\psi| \,] |\psi\rangle.$$

В этом случае мы хотим как можно точнее передавать все чистые состояния из подпространства $\mathcal{L}$. В силу доказанной леммы, имеет место оценка

$$2[1 - F_s(\mathcal{L}, \Phi)] \leqslant \max_{\psi \in \mathcal{L}} \| |\psi\rangle\langle\psi| - \Phi[|\psi\rangle\langle\psi|] \|_1 \leqslant$$

$$\leqslant 2\sqrt{1 - F_s(\mathcal{L}, \Phi)}. \quad (9.3)$$

Эта оценка подобна неравенствам (1.16) в классическом случае.

Весьма важной является величина $F_e(S, \Phi)$, называемая точностью воспроизведения сцепленности (entanglement fidelity), которая определяется следующим образом. Расширим состояние $S = S_Q$ до чистого состояния $|\psi_{QR}\rangle\langle\psi_{QR}|$ в гильбертовом пространстве $\mathcal{H}_Q \otimes \mathcal{H}_R$ и рассмотрим точность воспроизведения этого чистого состояния тривиальным расширением канала:

$$F_e(S, \Phi) := \langle\psi_{QR}| (\Phi \otimes \mathrm{Id}_R)[|\psi_{QR}\rangle\langle\psi_{QR}|] |\psi_{QR}\rangle.$$

Имеется удобное выражение для величины $F_e(S, \Phi)$, которое также показывает независимость данного выше определения от способа очищения состояния. Пусть канал имеет представление:

$$\Phi[S] = \sum_i V_i S V_i^*, \quad (9.4)$$

тогда

$$F_e(S, \Phi) = \sum_i |\operatorname{Tr} SV_i|^2.$$

В самом деле,

$$\left\langle \psi_{QR} \middle| \sum_i (V_i \otimes I_R) |\psi_{Q,R}\rangle\langle\psi_{QR}| (V_i \otimes I_R)^* \middle| \psi_{QR} \right\rangle =$$

$$= \sum_i |\langle\psi_{QR}| (V_i \otimes I_R) |\psi_{QR}\rangle|^2 = \sum_i |\operatorname{Tr} SV_i|^2.$$

Задача 43. Доказать, что величина $F_e(S, \Phi)$ не зависит от способа представления канала Φ в виде (9.4).

Важность характеристики $F_e(S, \Phi)$ в значительной мере обусловлена следующим результатом.

Лемма 16 (квантовое неравенство Фано). *Для любого состояния S выполнено неравенство*

$$H(S, \Phi) \leqslant h(F_e(S, \Phi)) + (1 - F_e(S, \Phi)) \log(d^2 - 1) \leqslant$$

$$\leqslant 1 + 2(1 - F_e(S, \Phi)) \log d.$$

Доказательство. Пусть $S_{QR} = |\psi_{QR}\rangle\langle\psi_{QR}|$ чистое состояние в $\mathcal{H}_Q \otimes \mathcal{H}_R$, продолжающее S. Выберем ортонормированный базис $\{|e_j\rangle\}$ в $\mathcal{H}_Q \otimes \mathcal{H}_R$, так что $|e_1\rangle = |\psi_{QR}\rangle$. Обозначая $p_i = \langle e_i| S_{Q'R'} |e_i\rangle$, имеем

$$H(S, \Phi) = H(S_{Q'R'}) \leqslant -\sum_{i=1}^{d^2} p_i \log p_i.$$

В самом деле, правую часть этого неравенства можно записать как $H(\Psi[S_{Q'R'}])$, где $\Psi[S] = \sum |e_i\rangle\langle e_i| S |e_i\rangle\langle e_i|$, так что $H(\Psi[S]) \geqslant H(S)$ согласно следствию 2 теоремы 13. Заметим, что $p_1 = F_e(S, \Phi)$, поскольку $S_{Q'R'} = \Phi \otimes \mathrm{Id}_R[|\psi_{QR}\rangle\langle\psi_{QR}|]$. Далее,

$$-\sum_{i=1}^{d^2} p_i \log p_i = -p_1 \log p_1 - (1-p_1)\log(1-p_1) -$$

$$-(1-p_1)\sum_{i=2}^{d^2} \frac{p_i}{1-p_1} \log \frac{p_i}{1-p_1} \leqslant$$

$$\leqslant h(F_e(S, \Phi)) + (1 - F_e(S, \Phi)) \log(d^2 - 1),$$

так как максимум классической энтропии достигается на равномерном распределении.

Между различными мерами точности существует ряд соотношений. Докажем важное неравенство [1].

Лемма 17. *Для любого состояния S выполнено неравенство*

$$1 - F_e(S, \Phi) \leqslant 4\sqrt{1 - F_s(\mathrm{supp}\, S, \Phi)}.$$

Таким образом, F_s является более чувствительной мерой точности воспроизведения, нежели F_e.

Доказательство. Не ограничивая общности, будем считать, что $\mathrm{supp}\, S = \mathcal{H}$. Имеем

$$1 - F_e(S, \Phi) = 1 - \langle\psi_{QR}|(\Phi \otimes \mathrm{Id}_R)[|\psi_{QR}\rangle\langle\psi_{QR}|]|\psi_{QR}\rangle =$$

$$\langle\psi_{QR}|(\mathrm{Id}_Q - \Phi) \otimes \mathrm{Id}_R)[|\psi_{QR}\rangle\langle\psi_{QR}|]|\psi_{QR}\rangle.$$

Представляя

$$|\psi_{QR}\rangle = \sum_j |\psi_j\rangle \otimes |e_j\rangle,$$

[1] *Werner R. F.* Unpublished notes. 1999.

где $|e_j\rangle$ — ортонормированный базис в $\mathcal{H}_R$, $\sum_j \|\psi_j\|^2 = 1$, получаем, что это равно

$$\sum_{jk} \langle\psi_j|(\mathrm{Id}_Q - \Phi)[|\psi_j\rangle\langle\psi_k|]|\psi_k\rangle \leqslant \|\mathrm{Id}_Q - \Phi\| \equiv \max \frac{\|T - \Phi[T]\|_1}{\|T\|_1},$$

где максимум берется по всевозможным ненулевым операторам T. Разлагая $T = T_1 + iT_2$, где $T_1^* = T_1$, $T_2^* = T_2$, и учитывая, что $\|T_{1,2}\|_1 \leqslant \leqslant \|T\|_1$, а также неравенство треугольника, получаем, что это не превосходит

$$2 \max_{T^* = T} \frac{\|T - \Phi[T]\|_1}{\|T\|_1} = 2 \max_{S \in S(\mathcal{H})} \|S - \Phi[S]\|_1 = \\ = 2 \max_{\|\psi\| = 1} \|\,|\psi\rangle\langle\psi| - \Phi[|\psi\rangle\langle\psi|]\|_1 .$$

Последнее равенство следует из выпуклости нормы и того факта, что максимум выпуклой функции достигается на крайних точках выпуклого множества. Используя второе неравенство в (9.3), получаем, что это выражение не превосходит $4\sqrt{1 - F_s}$.

Следствие. *Следующие условия эквивалентны*:

1) $F_e(S, \Phi) = 1$;

2) $F_s(\mathrm{supp}\, S, \Phi) = 1$;

3) $\Phi[\widetilde{S}] = \widetilde{S}$ *для любого состояния* $\widetilde{S}$ *с носителем* $\mathrm{supp}\, \widetilde{S} \subset \subset \mathrm{supp}\, S$.

Доказательство. Согласно определению F_s, условие 2 эквивалентно условию 3 для чистых, а значит, и для смешанных состояний $\widetilde{S}$. Из леммы 15 вытекает, что 2)⇒3). Покажем, что 1)⇒2). Пусть $|\psi\rangle \in \mathrm{supp}\, S$, тогда

$$S = p|\psi\rangle\langle\psi| + (1 - p)S',$$

где $p > 0$ и S' некоторый оператор плотности. Функция $S \to F_e(S, \Phi)$ положительная и квадратичная, а значит, выпуклая. Поэтому

$$pF_e(|\psi\rangle\langle\psi|, \Phi) + (1 - p)F_e(S', \Phi) \geqslant F_e(S, \Phi) = 1,$$

следовательно $F_e(|\psi\rangle\langle\psi|, \Phi) = 1$. Но это означает, что

$$\langle\psi|\,\Phi[\,|\psi\rangle\langle\psi|\,]\,|\psi\rangle = 1$$

для всех $|\psi\rangle \in \mathrm{supp}\, S$, и поэтому $F_s(\mathrm{supp}\, S, \Phi) = 1$.

§ 9.2. Когерентная информация и обратимость канала

Важной составной частью квантовой взаимной информации $I(S_Q, \Phi)$ является *когерентная информация*

$$I_c(S,\Phi) := H(\Phi[S]) - H(S;\Phi) = H(S_{Q'}) - H(S_{E'}) = H(S_{Q'}) - H(S_{R'Q'}).$$

Как мы увидим в следующем параграфе, она связана с верхней границей скорости передачи чисто квантовой информации по каналу Ф. Когерентная информация не обладает рядом «естественных» свойств, таких как: вогнутость по S_Q; субаддитивность; цепное свойство 2).

Более того, она может принимать отрицательные значения. Заметим, что ее классический аналог вообще неположителен, $H(Y) - H(XY) = -H(X|Y) \leqslant 0$. Однако, $I_c(S,\Phi)$ вогнута по Ф и обладает цепным свойством 1):

$$I_c(S, \Phi_2 \circ \Phi_1) \leqslant I_c(S, \Phi_1), \tag{9.5}$$

которое следует из соотношения $I_c(S,\Phi) = I(S,\Phi) - H(S)$, и соответствующего свойства квантовой взаимной информации $I(S,\Phi)$.

Следующий идеальный случай позволяет прояснить соотношение между возможностью безошибочной передачи квантовой информации, исправлением ошибок и когерентной информацией. Канал Ф называется *обратимым на состоянии* S, если найдется такой канал $\mathcal{D}$, что $F_e(S, \mathcal{D}\circ\Phi) = 1$. Согласно следствию из леммы 17, в этом случае $\mathcal{D}\circ\Phi[\widetilde{S}] = \widetilde{S}$ для всех $\widetilde{S}$ с носителем $\operatorname{supp}\widetilde{S} \subset \operatorname{supp} S$. Как мы увидим в следующей главе, это означает, что подпространство $\mathcal{L} = \operatorname{supp} S$ «исправляет ошибку Ф».

ТЕОРЕМА 17. *Канал* Ф *обратим на состоянии* S *тогда и только тогда, когда*

$$I_c(S,\Phi) = H(S). \tag{9.6}$$

Доказательство. Используем стандартные обозначения Q, R, E для основной системы, эталонной системы и окружения, так что $S = S_Q$. Имеем

$$H(S) - I_c(S,\Phi) = H(S_Q) + H(S_{E'}) - H(S_{Q'}) =$$
$$= H(S_{R'}) + H(S_{E'}) - H(S_{R'E'}) = I(R';E') \geqslant 0,$$

где равенство достигается тогда и только тогда, когда $I(R';E')=0$, то есть $S_{R'E'}=S_{R'}\otimes S_{E'}$. Отсюда можно получить достаточность условия (9.6), см. [13, п. 12.4.2].

Для доказательства необходимости используем цепное свойство когерентной информации. Имеем

$$H(S)\geqslant I_c(S,\Phi)\geqslant I_c(S,\mathcal{D}\circ\Phi)=H(\mathcal{D}\circ\Phi[S])-H(S,\mathcal{D}\circ\Phi). \quad (9.7)$$

Если канал Φ обратим на S, то $\mathcal{D}\circ\Phi[S]=S$. Более того, $F_e(S,\mathcal{D}\circ\Phi)=1$. Используя квантовое неравенство Фано, получаем $H(S,\mathcal{D}\circ\Phi)=0$. Поэтому правая часть (9.7) равна $H(S)$, что и доказывает необходимость условия (9.6).

§ 9.3. Квантовая пропускная способность

Перейдем к вопросу о квантовой пропускной способности. В принципе, эта величина может зависеть от выбора меры точности воспроизведения. Замечательно однако, что, как можно показать, квантовая пропускная способность канала нечувствительна к этому выбору, именно, она одинакова для F_e и F_s. Для того, чтобы сформулировать определение квантовой пропускной способности, рассмотрим канал $\Phi^{\otimes n}$ в пространстве $\mathcal{H}^{\otimes n}=\mathcal{H}\otimes\ldots\otimes\mathcal{H}$.

ОПРЕДЕЛЕНИЕ. Скорость R является *достижимой* при передаче подпространств через канал Φ, если существует такая последовательность подпространств $\mathcal{H}^{(n)}$, что

$$\overline{\lim_{n\to\infty}}\ \log\dim\mathcal{H}^{(n)}/n=R,$$

и последовательности каналов $\mathcal{E}^{(n)}$ из $\mathcal{H}^{(n)}$ в $\mathcal{H}^{\otimes n}$ (кодирования) и $\mathcal{D}^{(n)}$ из $\mathcal{H}^{\otimes n}$ в $\mathcal{H}^{(n)}$ (декодирования), такие что

$$F_s(\mathcal{H}^{(n)},\mathcal{D}^{(n)}\circ\Phi^{\otimes n}\circ\mathcal{E}^{(n)})\to 1,$$

при $n\to\infty$. Точную верхнюю грань таких скоростей передачи обозначим $Q_s(\Phi)$.

ОПРЕДЕЛЕНИЕ. Скорость R *достижима* при передаче сцепленности через канал Φ, если существует источник с энтропией R, т. е. такая последовательность состояний $S^{(n)}$ в гильбертовых пространствах $\mathcal{H}^{(n)}$, что

$$R=\overline{\lim_{n\to\infty}}\frac{H(S^{(n)})}{n},$$

и последовательности кодирований $\mathcal{D}^{(n)}$ и декодирований $\mathcal{E}^{(n)}$, такие, что

$$F_e(S^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) \to 1$$

при $n \to \infty$. Точную верхнюю грань таких скоростей обозначим $Q_e(\Phi)$.

Из соотношения между мерами точности (лемма 17) вытекает, что $Q_s(\Phi) \leqslant Q_e(\Phi)$; в самом деле, предположим, что скорость R достижима при передаче подпространств и рассмотрим оператор плотности $S^{(n)} = I_{\mathcal{H}^{(n)}}/\dim \mathcal{H}^{(n)}$. Тогда согласно лемме 17 R удовлетворяет и второму определению, и значит, достижима при передаче сцепленности. Труднее доказывается обратное неравенство $Q_e(\Phi) \geqslant Q_s(\Phi)$. Доказательство использует квантовый аналог леммы 2 об эквивалентности минимальной и средней ошибок при определении классической пропускной способности [1]. Величина $Q_s(\Phi) = Q_e(\Phi) = Q(\Phi)$ и называется *квантовой пропускной способностью* канала Φ.

Получим принципиальную верхнюю границу для квантовой пропускной способности в терминах когерентной информации.

ЛЕММА 18. *Для любого входного состояния S канала Φ и любого канала $\mathcal{D}$ справедливо неравенство*

$$H(S) \leqslant I_c(S, \Phi) + 2 + 4(1 - F_e(S, \mathcal{D} \circ \Phi)) \log d.$$

Доказательство. Используя цепное неравенство для когерентной информации и неравенство треугольника (7.8), получаем

$$H(S) - I_c(S, \Phi) \leqslant H(S) - I_c(S, \mathcal{D} \circ \Phi) =$$
$$= H(S) - H(\mathcal{D} \circ \Phi[S]) + H(S, \mathcal{D} \circ \Phi) \leqslant 2H(S, \mathcal{D} \circ \Phi).$$

Согласно квантовому неравенству Фано имеем

$$H(\mathcal{D} \circ \Phi[S]) \leqslant 1 + 2(1 - F_e(S, \mathcal{D} \circ \Phi)) \log d.$$

Комбинируя эти неравенства, получаем утверждение леммы.

С помощью этой леммы можно доказать следующее утверждение.

ТЕОРЕМА 18.

$$Q(\Phi) \leqslant \overline{Q}(\Phi) = \lim_{n \to \infty} \frac{1}{n} \max_S I_c(S, \Phi^{\otimes n}).$$

---

[1] *Barnum H., Knill E., Nielsen M. A.* On quantum fidelities and channel capacities // IEEE Trans. Inform. Theory. — 2000. — V. 46. — P. 1317–1329; LANL Report no. quant-ph/9809010, Sep. 1998.

Доказательство. Мы ограничимся доказательством более слабого утверждения:

$$Q(\Phi) \leqslant \lim_{n\to\infty} \frac{1}{n} \max_{S,\mathcal{E}} I_c(S, \Phi^{\otimes n} \circ \mathcal{E}). \tag{9.8}$$

Заметим, что поскольку когерентная информация не удовлетворяет второму цепному неравенству, теорема не вытекает непосредственно из (9.8). Применяя лемму к n-кратному каналу, получим

$$H(S^{(n)}) \leqslant I_c(S^{(n)}, \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) + 2 + 4(1 - F_e(S^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{(n)} \circ \mathcal{E}^{(n)})) \log d^n.$$

Пусть найдутся кодирования и декодирования $\mathcal{E}^{(n)}$, $\mathcal{D}^{(n)}$, такие, что

$$\lim_{n\to\infty} F_e(S^{(n)}, \mathcal{D}^{(n)} \circ \Phi^{\otimes n} \circ \mathcal{E}^{(n)}) = 1,$$

тогда поделив это неравенство на n и взяв предел, с учетом того, что $F_e \to 1$, получим (9.8).

Существует предположение, что квантовая пропускная способность на самом деле равна верхней границе $\overline{Q}(\Phi)$. Это еще одна из интересных и важных открытых проблем квантовой теории информации.

ГЛАВА 10

КВАНТОВЫЕ КОДЫ, ИСПРАВЛЯЮЩИЕ ОШИБКИ

§ 10.1. Постановка вопроса

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

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

Прямолинейное обобщение этого рецепта на квантовый случай наталкивается на трудность — квантовую информацию невозможно размножить. Кроме того, по самому существу квантовой информации, при передаче через канал с шумом безошибочно должны приниматься не только базисные состояния, но и всевозможные их суперпозиции. На первый взгляд, такая задача кажется неразрешимой. Однако квантовый код, исправляющий ошибки, был построен независимо в работах Шора и Стина (см. [7, 16]). Многие авторы

сделали вклад в последующее развитие теории, фрагменты которой представлены в этой главе, см. обзор в [13].

Следуя классической аналогии, рассмотрим сначала код

$$|0\rangle \to |000\rangle, |1\rangle \to |111\rangle. \tag{10.1}$$

Такой код исправляет «переворот бита», т. е. переход $|0\rangle \leftrightarrow |1\rangle$ в любом одном q-бите. Нас интересует произвольное состояние $a|0\rangle + b|1\rangle$, которое при выбранном способе кодирования переходит в $a|000\rangle + b|111\rangle$. Пусть, например, произошла ошибка в первом q-бите:

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

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

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

Теперь заметим, что преобразование Адамара

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

отображает переворот фазы в переворот бита и наоборот. Преобразуя соответствующим образом код (10.1), получим код

$$\begin{aligned} |0\rangle &\to \frac{1}{2^{3/2}}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)(|0\rangle + |1\rangle), \\ |1\rangle &\to \frac{1}{2^{3/2}}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle)(|0\rangle - |1\rangle), \end{aligned} \tag{10.2}$$

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

Код Шора, который исправляет как переворот бита, так и переворот фазы в одном q-бите, получается комбинированием кодов (10.1), (10.2) и требует 9 q-битов

$$\begin{aligned} |0\rangle &\to \frac{1}{2^{3/2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ |1\rangle &\to \frac{1}{2^{3/2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \end{aligned} \tag{10.3}$$

Оказывается, что этот код исправляет не только битовую и фазовую, но любую ошибку, возникающую в результате применения *произвольного* канала в одном (любом) из q-битов. Прямая проверка этого факта является трудоемким упражнением, см. [13], § 10.2.

§ 10.2. Общая формулировка

Пусть S — произвольное состояние в гильбертовом пространстве $\mathcal{M}$. *Кодом* называется изометрическое отображение $V\colon \mathcal{M} \to \mathcal{N}$, переводящее состояния S в кодирующие состояния VSV^* в гильбертовом пространстве $\mathcal{N}$. Система $\mathcal{N}$ может быть подвержена ошибкам, которые описываются вполне положительными отображениями Φ в $\mathcal{N}$, удовлетворяющими условию $\Phi(I) \leqslant I$, образующими в каждой конкретной задаче некий класс $\mathcal{E}$.

Итак, преобразования квантовой информации описываются диаграммой

$$S \longrightarrow VSV^* \longrightarrow \Phi(VSV^*), \quad \Phi \in \mathcal{E}.$$

Определение. Код V называется *кодом, исправляющим ошибки* из $\mathcal{E}$, если существует такой *восстанавливающий канал* Ψ, что

$$\Psi[\Phi(VSV^*)] = c(\Phi)S, \tag{10.4}$$

для любого состояния S и любого $\Phi \in \mathcal{E}$, где $c(\Phi)$ — некоторая постоянная, зависящая только от Φ.

Замечания. 1. На самом деле код можно задавать подпространством $\mathcal{L} = V\mathcal{M} \subset \mathcal{N}$, не вводя явно $\mathcal{M}, V$.

2. Множество ошибок обычно имеет следующую структуру: $\Phi(S) = \sum\limits_j V_j S V_j^*$, где $V_j \in \mathrm{Lin}(B_1, \ldots, B_p)$, а B_j — операторы *элементарных ошибок*.

Пример. Хранение квантовой информации в памяти квантового компьютера. Пусть $\mathcal{N} = \mathcal{H}_2^{\otimes n}$ — квантовый регистр, в котором предполагается хранить информацию из $\mathcal{M}$. Рассмотрим ошибки, при которых изменению может подвергнуться не более m q-битов регистра. Соответствующее множество $\mathcal{E}(n, m)$ состоит из отображений $\Phi = \Phi_1 \otimes \ldots \otimes \Phi_n$, 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.

Пример Шора демонстрирует возможность исправления ошибок из $\mathcal{E}(n, 1)$, если n достаточно велико (можно доказать, что наименьшее значение n для кода, исправляющего одну ошибку, равно 5). Возможность исправления только одной ошибки является, конечно, серьезным ограничением. Однако удалось показать, что существуют коды, исправляющие ошибки из $\mathcal{E}(n, m)$, где m может быть сколь угодно большим для достаточно больших размеров регистра n. Более того, была предложена принципиальная схема квантового компьютера, исправляющего ошибки не только в квантовой памяти, но и в самой схеме, исправляющей ошибки, при условии, что вероятность ошибки не превосходит некоторого порогового значения (fault-tolerant quantum computing) [2; 16]. Грубые оценки показывают, что квантовое устройство, способное решать задачи, недоступные классическому компьютеру, должно иметь регистр, насчитывающий 2000–3000 q-бит и порядка 10^{10} элементарных логических операторов, которые должны иметь уровень надежности порядка 10^{-13}. Более реалистический уровень надежности 10^{-4} достаточен при использовании исправления ошибок, влекущем увеличение вычислительных ресурсов в ≈ 100 раз [16]. В настоящее время имеются экспериментальные установки, позволяющие оперировать с состояниями 2–3 q-битов. Это красноречиво свидетельствует о том, какой сложный и длинный путь предстоит пройти до практической реализации квантового компьютера. Однако не следует при этом забывать, что все многообразие современных средств обработки информации, в корне преобразившее мир, в котором мы живем, возникло в течение последнего столетия, а лежащие в его основе фундаментальные открытия были сделаны не ранее XIX-го века.

§ 10.3. Необходимые и достаточные условия исправления ошибок

Пусть $\mathcal{E}$ класс ошибок, имеющий структуру, описанную в замечании 2.

ТЕОРЕМА 19[1]. *Следующие утверждения эквивалентны*:

1) *код $\mathcal{L}$ исправляет ошибки класса $\mathcal{E}$;*

[1] Ср. *Knill E., Laflamme R.* Theory of quantum error-correcting codes // Phys. Rev. A. — 1997. — V. 55. — P. 900–911.

2) *код* $\mathcal{L}$ *исправляет ошибку* $\Phi[S]=\sum_{j=1}^{p} B_j S B_j^*$;

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

4) *для любого ортонормированного базиса* $\{|k\rangle\}$ *в* $\mathcal{L}$ *выполняются равенства*

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

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

5) $P_{\mathcal{L}} B_i^* B_j P_{\mathcal{L}} = b_{ij} P_{\mathcal{L}}$, *где* $P_{\mathcal{L}}$ — *проектор на* $\mathcal{L}$.

Доказательство. 1) $\Rightarrow$ 2) очевидно.

2) $\Rightarrow$ 3). Пусть существует восстанавливающий канал $\Psi[S] = \sum_r R_r S R_r^*$ 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 $R_r B_j|\psi\rangle = c_{jr}|\psi\rangle$ ($c = \sum_r \sum_j |c_{jr}|^2$).

Пусть $|\varphi\rangle$, $|\psi\rangle \in \mathcal{L}$ два ортогональных вектора. Поскольку восстанавливающий канал сохраняет след, то $I = \sum R_r^* R_r$, и значит

$$\langle\varphi|B_i^* B_j|\psi\rangle = \sum_r \langle\varphi|B_i^* R_r^* R_r B_j|\psi\rangle = \left(\sum_r \overline{c_{ir}} c_{jr}\right)\langle\varphi|\psi\rangle = 0.$$

3) $\Rightarrow$ 4). Второе равенство очевидно, а первое получается из рассмотрения ортогональных векторов $|\varphi\rangle = |k+l\rangle$, $|\psi\rangle = |k-l\rangle$.

4) $\Rightarrow$ 5). Обозначая $b_{ij} = \langle k|B_i^* B_j|k\rangle$ мы можем записать условие 4 в виде

$$\langle k|B_i^* B_j|l\rangle = \delta_{kl} b_{ij},$$

что эквивалентно условию 5.

5) $\Rightarrow$ 1). Матрица $[b_{ij}]$ эрмитова положительна, поэтому найдется такая унитарная матрица $[u_{ij}]$, что

$$\sum_{ij} \bar{u}_{ir} b_{ij} u_{js} = \delta_{rs} \lambda_r,$$

где $\lambda_r \geqslant 0$. Полагая $\widetilde{B}_s = \sum_j u_{js} B_j$, имеем $\Phi[S] = \sum_{s=1}^{p} \widetilde{B}_s S \widetilde{B}_s^*$ и

$$P_{\mathcal{L}} \widetilde{B}_r^* \widetilde{B}_s P_{\mathcal{L}} = \delta_{rs} \lambda_r P_{\mathcal{L}}.$$

Таким образом, при $\lambda_s > 0$ имеем

$$\lambda_s^{-1/2} \widetilde{B}_s P_{\mathcal{L}} = U_s P_{\mathcal{L}},$$

где U_s частично изометрические операторы, отображающие $\mathcal{L}$ на взаимно ортогональные подпространства $\mathcal{L}_s \subset \mathcal{H}$.

Пусть P_s — проектор на подпространство $\mathcal{L}_s$, и обозначим

$$P_0 = I - \sum_s P_s.$$

Определим канал

$$\Psi[S] = \sum_s U_s^* P_s S P_s U_s + P_0 S P_0$$

и покажем, что он является восстанавливающим для всех ошибок из $\mathcal{E}$. Поскольку линейная комбинация элементарных ошибок B_j является также линейной комбинацией операторов $\widetilde{B}_r$, то, принимая во внимание, что $U_s^* P_s \widetilde{B}_r P_{\mathcal{L}} = \lambda_r^{1/2} \delta_{sr} P_{\mathcal{L}}$, мы получаем (10.4) для произвольного $\Phi \in \mathcal{E}$.

Задача 44. Проверьте выполнение условий 4) для кода Шора (10.3) и операторов элементарных ошибок, задаваемыми матрицами Паули в произвольном q-бите.

§ 10.4. Аддитивные (симплектические) коды

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

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

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

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

Именно, вводя кососимметричную функцию Δ с значениями

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

имеем

$$\sigma_\gamma \sigma_{\gamma'} = (-i)^{\Delta(\gamma,\gamma')}\sigma_{\gamma+\gamma'} = (-1)^{\Delta(\gamma,\gamma')}\sigma_{\gamma'}\sigma_\gamma, \quad \gamma, \gamma' \in B^2. \qquad (10.5)$$

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

Задача 45. Покажите, что форма $\Delta(\gamma, \gamma') \pmod 2$ является билинейной и невырожденной на B^2.

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

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

где $\Delta(f, g) = \Delta(\gamma_1, \gamma_1') + \ldots + \Delta(\gamma_n, \gamma_n')$ если $g = (\gamma_1', \ldots, \gamma_n')$. Таким образом, B^{2n} становится $2n$-мерным векторным пространством над полем B, снабженным невырожденной симплектической формой $\Delta(f, g) \pmod 2$. Заметим, что операторы $\sigma(g), \sigma(f)$ коммутируют (антикоммутируют) тогда и только тогда, когда $\Delta(f, g) = 0 \pmod 2$, (соответственно $\Delta(f, g) \neq 0 \pmod 2$).

Пусть $g_1, \ldots, g_{n-k}$ — такие линейно независимые векторы из B^{2n}, что $\Delta(g_i, g_j) = 0 \pmod 2$ для всех i, j. Тогда операторы $\sigma(g_1)$, $\sigma(g_2), \ldots, \sigma(g_{n-k})$ коммутируют между собой. *Аддитивным кодом с проверочным операторами* $\sigma(g)_1, \ldots, \sigma(g)_{n-k}$ называется линейное подпространство

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

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

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

$$f_1 - f_2 \in G^{\perp} = \{f \in B^{2n} \mid \Delta(f, g) \pmod 2 = 0, \quad g \in G\}.$$

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

ТЕОРЕМА 20[1]. *Аддитивный код $\mathcal{L}$ исправляет ошибки класса $\mathcal{E}$, если для любых двух ошибок $\sigma(f_1)$, $\sigma(f_2)$ либо*

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

либо

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

Заметим, что условие теоремы может быть компактно записано как

$$(E - E) \cap (G^{\perp} \setminus G) = \varnothing.$$

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

Доказательство. Проверим выполнение условия 3) теоремы 19. Пусть ψ, φ — ортогональные векторы из $\mathcal{L}$, и $f_1, f_2 \in E$. Если ошибки различимы, то векторы $\sigma(f_1)\psi$, $\sigma(f_2)\varphi$ являются собственными векторами проверочных операторов с различными наборами

[1] Ср. *Calderbank A. R., Rains E. M., Shor P. W., Sloane N. J. A.* Quantum error correction and orthogonal geometry // Phys. Rev. Lett. — 1997. — V. 78. — P. 404–408.

собственных значений и, следовательно, они ортогональны. Если они неразличимы, то по условию, должно выполняться $f_1 - f_2 \in G$, и в этом случае, используя канонические коммутационные соотношения, получаем

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

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

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

Задача 46. Убедитесь, что код Шора является аддитивным кодом с проверочными операторами

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

Первые шесть операторов обнаруживают переворот бита, а последние два — переворот фазы в любом из блоков. Проверьте выполнение условий теоремы.

ПРИЛОЖЕНИЕ

ДОКАЗАТЕЛЬСТВО МОНОТОННОСТИ ОТНОСИТЕЛЬНОЙ ЭНТРОПИИ

Существует несколько доказательств монотонности относительной энтропии, и все они достаточно сложны. В серии работ [1)] Линдблад получил ряд последовательных обобщений свойства монотонности и в заключение показал, что это свойство равносильно сильной субаддитивности квантовой энтропии, установленной ранее Либом и Рускаи. В основе этого подхода лежат фундаментальные «теоремы вогнутости» Либа. Позднее Ульман дал другое доказательство, основанное на методе интерполяции (см. обзоры в [17; 14]). Ниже приводится новое доказательство, данное Лесьневским и Рускаи [2)], которое на наш взгляд, является наиболее прямым, и в конечном итоге, основано на некотором обобщении неравенства Коши — Буняковского. Это доказательство позволяет установить монотонность целого класса инвариантов пары состояний в квантовой геометростатистике, изучение которой было начато Морозовой и Ченцовым [3)].

Пусть S_1 и S_2 невырожденные операторы плотности. *Относительный модулярный оператор*, ассоциированный с S_1, S_2, определяется соотношением

$$\Delta_{S_2, S_1} = L_{S_2} R_{S_1}^{-1},$$

где L_{S_2} и R_{S_1} операторы левого и правого умножения, соответственно, так что $\Delta_{S_2, S_1}(A) = S_2 A S_1^{-1}$. Легко проверяется, что Δ_{S_2, S_1} является положительным эрмитовым оператором в пространстве операторов Гильберта — Шмидта.

ОПРЕДЕЛЕНИЕ. Пусть g операторно-выпуклая функция (см. [8]) на $(0, \infty)$, такая что $g(1) = 0$. *Относительная g-энтропия* опера-

---

[1)] *Lindblad G.* Entropy, information and quantum measurements // Commun. Math. Phys. — 1973. — V. 33. — P. 305–322; Expectations and entropy inequalities for finite quantum systems // Commun. Math. Phys. — 1974. — V. 39. — P. 111–119; Completely positive maps and entropy inequalities // Commun. Math. Phys. — 1975. — V. 40. — P. 147–151.

[2)] *Lesniewski A., Ruskai M. B.* Monotone Riemannian metrics and relative entropy on non-commutative probability spaces // J. Math. Phys. — 1999. — V. 40. — P. 5702–5724.

[3)] *Морозова Е. А., Ченцов Н. Н.* Марковская инвариантная геометрия на многообразиях состояний // Итоги науки и техники. Современные проблемы математики. Новейшие достижения. — М.: ВИНИТИ, 1990. Т. 36. С. 69–102.

торов S_1, S_2 определяется соотношением

$$H_g(S_1; S_2) = \mathrm{Tr}(S_1^{1/2} g(\Delta_{S_2, S_1}) S_1^{1/2}).$$

Обозначим $\mathcal{G}$ множество операторно-выпуклых функций, удовлетворяющих этим условиям. Используя стандартные результаты из теории операторно-выпуклых и операторно-монотонных функций [8], можно показать, что класс $\mathcal{G}$ состоит из функций вида

$$g(w) = a(w-1) + b(w-1)^2 + \int_0^\infty \frac{(w-1)^2}{w+s} d\nu(s), \tag{11.1}$$

где $b \geqslant 0$ и ν положительная мера на $[0, \infty)$, такая что

$$\int_0^\infty \frac{1}{s+1} d\nu(s) < \infty.$$

На самом деле, для наших целей важно только то, что рассматриваемая функция $g(w)$ допускает такое представление, и что функция $g(w) = -\log w$, которой соответствует обычная относительная энтропия $H_{\log}(S_1; S_2) = H(S_1; S_2)$, входит в этот класс. Легко проверяется, что

$$-\log w = -(w-1) + \int_0^\infty \frac{(w-1)^2}{(w+s)(s+1)^2} ds.$$

ТЕОРЕМА 21. *Для любой функции $g \in \mathcal{G}$ имеет место соотношение*

$$H_g(S_1; S_2) = \mathrm{Tr}(S_2 - S_1)\left[bS_1^{-1}\right](S_2 - S_1) + \\ + \int_0^\infty \mathrm{Tr}\left((S_2 - S_1)[L_{S_2} + sR_{S_1}]^{-1}(S_2 - S_1)\right) d\nu(s). \tag{11.2}$$

Доказательство. Заметим, что

$$(\Delta_{S_2, S_1} - I)(S_1^{1/2}) = (S_2 - S_1) S_1^{-1/2} = R_{S_1^{-1/2}}(S_2 - S_1), \tag{11.3}$$

так что

$$H_{w-1}(S_1; S_2) = \mathrm{Tr}\left[S_1^{1/2}(S_2 - S_1) S_1^{-1/2}\right] = 0,$$

так что линейный член в (11.1) не дает вклада. Для $g = (w-1)^2/(w+s)$, используя (11.3), находим

$$H_g(S_1; S_2) = \langle (\Delta_{S_2, S_1} - I)(S_1^{1/2}), (\Delta_{S_2, S_1} + sI)^{-1}(\Delta_{S_2, S_1} - I)(S_1^{1/2}) \rangle =$$
$$= \operatorname{Tr}\left[(S_2 - S_1)(\Delta_{S_2, S_1} + sI)^{-1} R_{S_1^{-1}}(S_2 - S_1)\right] =$$
$$= \operatorname{Tr}(S_2 - S_1)[L_{S_2} + sR_{S_1}]^{-1}(S_2 - S_1).$$

Полагая $s = 0$, получаем

$$H_{(w-1)^2/w}(S_1; S_2) = \operatorname{Tr}(S_2 - S_1) S_2^{-1}(S_2 - S_1) = H_{(w-1)^2}(S_2; S_1).$$

Подставляя эти соотношения в (11.1), получаем (11.2).

ТЕОРЕМА 22. *Для любой функции $g \in \mathcal{G}$ относительная g-энтропия $H_g(S_1; S_2)$ монотонна, т. е.*

$$H_g(S_1; S_2) \leqslant H_g(\Phi(S_1); \Phi(S_2)),$$

для любого канала Φ.

Доказательство вытекает из интегрального представления (11.2) и следующей теоремы

ТЕОРЕМА 23. *Для любого канала Φ и $s \geqslant 0$ выполнено неравенство*

$$\operatorname{Tr} A^*[R_{S_1} + sL_{S_2}]^{-1} A = \operatorname{Tr} \Phi\left(A^*[R_{S_1} + sL_{S_2}]^{-1} A\right) \geqslant$$
$$\geqslant \operatorname{Tr} \Phi(A^*)[R_{\Phi(S_1)} + sL_{\Phi(S_2)}]^{-1}\Phi(A). \quad (11.4)$$

Доказательство. Если $S_1 \geqslant 0$, то $\operatorname{Tr} A^* S_1 A \geqslant 0$ и $\operatorname{Tr} A^* A S_1 \geqslant 0$, так что L_{S_1} и R_{S_1} являются положительными операторами в пространстве операторов Гильберта — Шмидта. Поэтому, поскольку $S_2 \geqslant 0$, оператор $R_{S_1} + sL_{S_2}$ также является положительным. Положим

$$X = [R_{S_1} + sL_{S_2}]^{-1/2}(A) - [R_{S_1} + sL_{S_2}]^{1/2}\Phi^*(B),$$

где $B = [R_{\Phi(S_1)} + sL_{\Phi(S_2)}]^{-1}\Phi(A)$. Тогда $\operatorname{Tr} X^* X \geqslant 0$, так что

$$\operatorname{Tr} A^*[R_{S_1} + sL_{S_2}]^{-1} A - \operatorname{Tr} A^*\Phi^*(B) - \operatorname{Tr} \Phi^*(B^*)A +$$
$$+ \operatorname{Tr} \Phi^*(B^*)[R_{S_1} + sL_{S_2}]\Phi^*(B) \geqslant 0. \quad (11.5)$$

Нетрудно проверить, что

$$-\operatorname{Tr} A^*\Phi^*(B) - \operatorname{Tr}\Phi^*(B^*)A = -2\operatorname{Tr}\Phi(A^*)[R_{\Phi(S_1)} + sL_{\Phi(S_2)}]^{-1}\Phi(A),$$

поэтому утверждение будет доказано, если мы покажем, что последний член в (11.5) меньше или равен, чем правая часть (11.4). Имеем

$$\begin{aligned}\operatorname{Tr}\Phi^*(B^*)[R_{S_1}+sL_{S_2}]\Phi^*(B) &= \operatorname{Tr}[\Phi^*(B^*)\Phi^*(B)S_1+\Phi^*(B^*)sS_2\Phi^*(B)] = \\ &= \operatorname{Tr}[\Phi^*(B^*)\Phi^*(B)S_1 + \Phi^*(B)\Phi^*(B^*)sS_2] \leqslant \\ &\leqslant \operatorname{Tr}[\Phi^*(B^*B)S_1 + \Phi^*(BB^*)sS_2],\end{aligned}$$

где неравенство следует из положительности S_1 и S_2 и операторного неравенства

$$\Phi^*(B^*)\Phi^*(B) \leqslant \Phi^*(B^*B),$$

которое имеет место для любого B, поскольку сохранение следа отображением Φ влечет $\Phi^*(I_2) = I_1$. Тогда, используя, например, соотношение $\operatorname{Tr}\widehat{\Phi}(B^*B)S_1 = \operatorname{Tr} B^*B\Phi(S_1)$, мы находим

$$\begin{aligned}\operatorname{Tr}\Phi^*(B^*)[R_{S_1} + sL_{S_2}]\Phi^*(B) &\leqslant \operatorname{Tr}[B^*B\Phi(S_1) + BB^*s\Phi(S_2)] = \\ &= \operatorname{Tr} B^*[B\Phi(S_1) + s\Phi(S_2)B] = \operatorname{Tr} B^*[R_{\Phi(S_1)} + sL_{\Phi(S_2)}]B = \\ &= \operatorname{Tr} B^*\Phi(A) = \operatorname{Tr}\Phi(A^*)[R_{\Phi(S_1)} + sL_{\Phi(S_2)}]^{-1}\Phi(A).\end{aligned}$$

ЛИТЕРАТУРА

1. Галлагер Р. Теория информации и надежная связь. — М.: Сов. Радио, 1974.

2. Китаев А. Ю. Квантовые вычисления: алгоритмы и исправление ошибок // УМН. — 1997. — Т. 52, № 6. — С. 53–112.

3. Хелстром К. Квантовая теория проверки гипотез и оценивания. — М.: Мир, 1978.

4. Холево А. С. Вероятностные и статистические аспекты квантовой теории. — М.: Наука, 1980.

5. Холево А. С. Квантовая вероятность и квантовая статистика // Итоги науки и техники. Современные проблемы математики. Фундаментальные направления. — М.: ВИНИТИ 1991.

6. Холево А. С. Квантовые теоремы кодирования // УМН. — 1998. — Т. 53, № 6. — С. 193–230.

7. Bennett C. H., Shor P. W. Quantum information theory // IEEE Trans. Inform. Theory. — 1998. — V. 44, № 6. — P. 2724–2742.

8. Bhatia R. Matrix analysis. — New York: Springer-Verlag, 1997.

9. Cover T. M., Thomas J. A. Elements of Information Theory. — New York: J. Wiley & Sons, 1991.

10. Davies E. B. Quantum theory of open systems. — London: Academic Press, 1976.

11. Holevo A. S. Statistical Structure of Quantum Theory. — Berlin: Springer-Verlag, 2001.

12. Kraus K. States, Effects and Operations. — Berlin: Springer 1983.

13. Nielsen M. A., Chuang I. I. Quantum Computation and Quantum Information. — Cambridge University Press, 2000.

14. Ohya M., Petz D. Quantum Entropy and Its Use. — Berlin: Springer-Verlag, 1993.

14. Shannon C. E., Weaver W. The mathematical theory of communication. — Univ. Illinois Press, Urbana Ill., 1949;

Имеется сокращенный перевод: Шеннон К. Статистическая теория передачи электрических сигналов // В кн.: Теория передачи электрических сигналов при наличии помех. — М.: ИЛ, 1953. С. 7–87.

15. Steane A. Quantum computing // Rept. Progr. Phys. — 1978. — V. 61. — P. 117–173; LANL Report no. quant-ph/9708022, Sep. 1997.

16. Wehrl A. General properties of entropy // Rev. Mod. Phys. — 1978. — V. 50, № . 2. — P. 221–260.

НЕЗАВИСИМЫЙ МОСКОВСКИЙ УНИВЕРСИТЕТ
ВЫСШИЙ КОЛЛЕДЖ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

СОВРЕМЕННАЯ МАТЕМАТИЧЕСКАЯ ФИЗИКА

ПРОБЛЕМЫ И МЕТОДЫ

Под редакцией А. И. Кириллова

ВЫПУСК 5

НЕЗАВИСИМЫЙ МОСКОВСКИЙ УНИВЕРСИТЕТ
ВЫСШИЙ КОЛЛЕДЖ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

А. С. ХОЛЕВО

ВВЕДЕНИЕ В КВАНТОВУЮ ТЕОРИЮ ИНФОРМАЦИИ

МЦНМО
МОСКВА • 2002

Холево А. С.

Введение в квантовую теорию информации. — М.: МЦНМО, 2002. — 128 с.

ISBN 5-94057-017-8.

Пятый выпуск серии «Современная математическая физика. Проблемы и методы» посвящен изложению основных понятий и строгих результатов новой научной дисциплины — квантовой теории информации. Возможности квантовых систем передачи и преобразования информации проиллюстрированы на примерах сверхплотного кодирования, квантовой телепортации и квантовых алгоритмов. Рассматриваются энтропийные и информационные характеристики квантовых систем. Подробно обсуждается понятие квантового канала связи, его классическая и квантовая пропускные способности, а также передача классической информации с помощью сцепленного состояния. Сформулировано несколько принципиальных открытых проблем, решение которых явилось бы существенным вкладом в квантовую теорию информации.

В лекциях приведены необходимые сведения из классической теории информации и дано подробное введение в статистическую структуру квантовой теории, поэтому для понимания лекций достаточно владения основными общематематическими дисциплинами.

ОГЛАВЛЕНИЕ