Дж. Прескилл

Квантовая информация и квантовые вычисления

Том 1

R&C
Dynamics


Институт
Компьютерных


R&C
Dynamics

Lecture Notes for Physics 229: Quantum Information and Computation

John Preskill
California Institute of Technology

September, 1998

Дж. Прескилл

Квантовая информация и квантовые вычисления

Том 1

Перевод с английского
Нечаевой Т. С.

Под научной редакцией
Ефифанова С. С. и Новокшонова С. Г.

Москва ♦ Ижевск

2008

УДК 22.314.1
ББК 517.958:530.145.6
П73

Издание осуществлено при финансовой поддержке Российского фонда фундаментальных исследований по проекту №01-02-30013.

Прескилл Дж.
Квантовая информация и квантовые вычисления. Том 1. — М.–Ижевск: НИЦ «Регулярная и хаотическая динамика»; Институт компьютерных исследований, 2008. — 464 с.

Книга Дж. Прескилла, известного специалиста в области квантовых вычислений и квантовой информации, написана на базе одноименного курса, читаемого автором в КАЛТЕХе, и представляет собой подробное и всестороннее введение в эту новую, быстро развивающуюся область науки.

На русском языке книга издается в двух томах. В первом из них излагаются основы теории квантовой информации и квантовых вычислений: свойства, отличающие квантовую информацию от классической; использование этих свойств для разработки систем безопасной передачи информации, квантовой телепортации, быстрых квантовых алгоритмов и другие вопросы.

Ясный физический характер изложения делает книгу доступной для новичка в этой области, находящейся на стыке физики, математики, информатики и технологии. Она содержит необходимые для понимания основного материала сведения из квантовой механики, классической теории информации и основные понятия из классических теорий вычислений и сложности. Каждая глава завершается небольшой подборкой задач разного уровня, способствующих более глубокому усвоению материала.

Для студентов, аспирантов, преподавателей и исследователей в области физики, математики и информатики, интересующихся квантовой теорией информации и вычислений.

ISBN 978-5-93972-651-1 **ББК 22.314.1**

http://shop.rcd.ru
http://ics.org.ru

Оглавление

Предисловие к русскому изданию

Физика квантовой информации и квантовых вычислений — новая, стремительно развивающаяся область науки, возникшая на стыке квантовой механики, современной математической физики и информатики. Огромный интерес к ней во многом стимулируется захватывающими перспективами, которые обещает открыть реализация ее идей практически во всех областях человеческой деятельности, связанных с передачей, хранением и обработкой информации.

За последние десять лет на русском языке было издано довольно много книг, посвященных данной теме (см., например, [1–7]). Несмотря на это, ощущается некоторый дефицит учебной литературы по теории квантовых вычислений и квантовой информации, адресованной в первую очередь читателю не математику в строгом смысле этого слова. Мы надеемся, что этот пробел может восполнить предлагаемый вашему вниманию курс лекций Дж. Прескилла, на протяжении более чем десяти лет читаемый автором в Калифорнийском Технологическом Институте (КАЛТЕХе), одном из крупнейших мировых исследовательских и образовательных центров в области квантовых информационных технологий. На русском языке книга выходит в двух томах; второй том в настоящее время готовится к изданию.

Большая часть этого курса лекций была написана в конце 90-х годов. Для столь бурно развивающейся отрасли науки это очень большой срок. Тем не менее книга Дж. Прескилла не утратила своего значения и по ряду причин до сих пор остается одним из базовых учебников по теории квантовых вычислений и квантовой информации. Во-первых, автор постоянно работал над совершенствованием содержания этого курса лекций. В частности, по сравнению с исходным вариантом была серьезно переработана и дополнена четвертая глава, посвященная квантовому запутыванию. Добавлена новая глава, посвященная квантовым топологическим вычислениям (войдет во второй том этой книги), теме, которой читатель не найдет ни в одной из вышедших до 2007 г. книг. Таким образом, курс Дж. Прескилла до сих пор остается одним из наиболее полных, отражающим практически все актуальные в настоящее время темы. Во-вторых, автор довольно подробно и математически строго рассматривает решение конкретных за-

дач, но при доказательстве математических теорем и утверждений он, как правило, переходит на качественный язык, справедливо полагая, что физику и инженеру гораздо важнее не уметь доказывать тонкие математические теоремы, а правильно пользоваться ими. Это делает книгу привлекательной для читателя, не имеющего специальной математической подготовки. Наконец, каждая глава данной книги сопровождается небольшим, но тщательно подобранным комплектом задач. Многие из них фактически представляют дальнейшее развитие теоретического материала.

С исследованиями в области квантовых вычислений и квантовой информации тесно связано возрождение интереса к старым принципиальным проблемам, таким как интерпретация квантовой механики, квантовая теория измерений, корреляции в запутанных квантовых состояниях и другие. В свое время по этим проблемам в физическом сообществе сформировалась FAPP-пригодная[1] точка зрения и, хотя проблемы остались, научный интерес к ним стал угасать. Как следствие, в современных учебниках по квантовой механике, преследующих более прагматические цели, этим вопросам уделяется незаслуженно мало внимания.

В этом отношении книга Дж. Прескилла представляет приятное исключение. Три ее главы (2-4) посвящены основаниям квантовой механики. После краткого введения в ее математический аппарат, автор детально обсуждает понятия и вопросы, имеющие непосредственное отношение к главной теме книги. При этом в большинстве случаев его подход оригинален и открывает перед читателем новые аспекты казалось бы хорошо знакомых проблем. Например, в стандартных курсах квантовой механики понятие матрицы плотности вводится аксиоматически. Дж. Прескилл выбирает физически более естественный путь. На простом примере он показывает, что к понятию матрицы плотности мы неизбежно приходим, пытаясь описать квантовомеханическое состояние доступной наблюдению части более широкой системы.

Большое внимание в книге уделено теории квантовых измерений, как ортогональных (измерения фон Неймана), так и обобщенных. При этом автор широко пользуется разложением единицы в гильбертовом пространстве физической системы (или ПОЗМ, положительной операторно-значной мерой), описывающим статистику результатов измерения. В настоящее время ПОЗМ является одним из эффективных инструментов теории квантовых измерений, хотя в русскоязычной учебной литературе это еще не нашло достаточного отражения.

Отдельная глава посвящена квантовому запутыванию несепарабельных двух- и многочастичных состояний, анализу возникающих в этих со-

[1] FAPP – *For All Practical Purposes*, то есть для всех практических целей (Дж. Белл).

стояниях квантовых корреляций, нарушающих неравенства Белла, и другим связанным с этим вопросам. Интерес к ним возобновился в последние годы, поскольку выяснилось, что именно использование запутывание как ресурса позволяет хранить и передавать квантовую информацию, а также оперировать ею и обеспечивать защиту от ошибок.

С этой точки зрения книга Дж. Прескилла может послужить прекрасным дополнением к любому стандартному курсу квантовой механики.

Все это позволяет сказать, что книга Дж. Прескилла представляет собой современный, оригинальный и достаточно полный курс лекций и будет полезна любому читателю, стремящемуся получить систематические знания в области квантовых вычислений и квантовой информации.

1) *Физика квантовой информации*, под ред. Д. Боумейстера, А. Экерта и А. Цайлингера. — М.: Постмаркет (2002).
2) К.А. Валиев, А.А. Кокин, *Квантовые компьютеры: надежда и реальность*. — Москва–Ижевск: РХД (2004).
3) А.А. Кокин, *Твердотельные квантовые компьютеры на ядерных спинах*. — Москва–Ижевск: ИКИ–РХД (2004).
4) Г.П. Берман, Г.Д. Дуплен, Р. Майнъери, В.И. Цифринович, *Введение в квантовые компьютеры*. — Москва–Ижевск: ИКИ–РХД (2004).
5) М. Нильсен, И. Чанг, *Квантовые вычисления и квантовая информация*. — М.: Мир (2006).
6) А. Китаев, А. Шень, М. Вялый, *Классические и квантовые вычисления*. — М.: МЦНМО-ЧеРо (1999).
7) А.С. Холево, *Введение в квантовую теорию информации*. — М.: МЦНМО (2002).

С. Г. Новокшонов

Часть I

Лекции

Глава 1

Введение и обзор

Курс имеет свою web-страницу:

`http://www.theory.caltech.edu/people/preskill/ph219/`

Здесь можно найти общую информацию, в том числе краткое содержание курса и важные ссылки.

К нашему предмету можно подойти с разных позиций, однако в этих лекциях будет принята точка зрения физика-теоретика (то есть моя точка зрения как физика-теоретика). Ввиду междисциплинарного характера предмета я осознаю, что студенты могут иметь самый разный уровень предварительной подготовки, и буду стараться учитывать это в лекциях. Пожалуйста, сообщайте мне, если я буду пользоваться неизвестными вам понятиями.

1.1. Физика информации

Почему физик преподает курс информации? Дело в том, что по меньшей мере несколько десятилетий *физика информации и вычислений* является общепризнанной дисциплиной. Это естественно. В конце концов, информация представляет собой нечто закодированное в состоянии физической системы; вычисление — нечто, что может быть выполнено реальным физически осуществимым устройством. Поэтому изучение информации и вычислений стоит связать с изучением лежащих в их основе физических процессов. Конечно, с технической точки зрения, владение принципами физики и материаловедения необходимо для совершенствования существующего вычислительного «железа» (Карвер Мид называет группой «физики вычислений» свою исследовательскую группу в КАЛТЕХе, которая занимается разработкой чипов).

С более абстрактной теоретической точки зрения, имеется несколько важных этапов в развитии нашего понимания того, как физика ограничивает возможность использовать информацию и оперировать ею. Например:

• **Принцип Ландауэра.** В 1961 году Рольф Ландауэр показал, что уничтожение информации — это непременно *диссипативный* процесс[1]. В его представлении уничтожение всегда влечет за собой сокращение фазового объема и, следовательно, необратимо.

Например, я могу хранить один бит информации, поместив единственную молекулу в ящик слева или справа от разделяющей его перегородки. Уничтожение означает, что мы перемещаем молекулу, скажем, в левую часть, независимо от того, где она находилась сначала – слева или справа. Я могу внезапно удалить перегородку, а затем с помощью поршня медленно сжимать состоящий из одной молекулы «газ» до тех пор, пока молекула действительно не окажется на левой стороне. Эта процедура уменьшает энтропию газа на $\Delta S = k \ln 2$ и сопровождается отводом соответствующего количества тепла из ящика в окружающее пространство. Если этот процесс является изотермическим при температуре T, то выполненная над ящиком работа $W = kT \ln 2$ — это работа, которую совершил я. Если я должен уничтожить информацию, то за это придется расплатиться энергией.

• **Обратимые вычисления.** Логические элементы (вентили), используемые для выполнения вычислений, обычно *необратимы*, например, логический элемент NAND (НЕ И)

$$(a, b) \rightarrow \neg(a \wedge b) \tag{1.1}$$

имеет два входящих бита и один выходящий бит, по которому мы не можем однозначно восстановить информацию на входе. Поскольку логическим элементом уничтожается около (в среднем по его возможным входам) одного бита информации, то, в соответствии с принципом Ландауэра, для его функционирования необходимо затратить работу, как минимум равную $W = kT \ln 2$. Если мы имеем ограниченный запас энергии, возникает теоретический предел продолжительности выполнения вычислений.

Однако в 1973 году Чарльз Беннет установил, что любые вычисления могут быть выполнены с помощью одних лишь обратимых операций, и, таким образом, в принципе нет необходимости ни в диссипации, ни в затратах энергии[2]. Фактически, мы можем создать обратимую версию логи-

[1] R. Landauer, Irreversibility and Heat Generation in the Computing Process, IBM J. Res. Develop., **3**, 183, (1961); русский перевод в сборнике статей *Квантовый компьютер и квантовые вычисления* под ред. В.А. Садовничего, Ижевск, РХД (1999). — *Прим. ред.*

[2] C.H. Bennett, Logical Reversibility of Computation, IBM J. Res. Develop., **17**, 525, (1973); русский перевод в сборнике статей *Квантовый компьютер и квантовые вычисления* под ред. В.А. Садовничего, Ижевск, РХД (1999). Популярное обсуждение физических ограничений, накладываемых на процессы вычислений, а также реализации обратимых вычислений см. в статье Шарль Г. Бенне (он же Чарльз Г. Беннет), Рольф Ландауэр, Физические пределы вычислений, В мире науки №9, 24 (1985), перевод журнала Scientific American. — *Прим. ред.*

ческого элемента NAND, который сохраняет всю входящую информацию: например, элемент Тоффоли

$$(a, b, c) \rightarrow (a, b, c \oplus a \wedge b) \tag{1.2}$$

является обратимым трехбитовым элементом, который «инвертирует» третий бит, если первые два принимают значение 1, и ничего не меняет в остальных случаях. Если $c = 1$, то третий выходящий бит принимает логическое значение НЕ a И b. Заменяя логические элементы NAND элементами Тоффоли, мы можем превратить необратимые вычисления в обратимые. В принципе эти вычисления могут быть выполнены с ничтожной диссипацией.

Однако в этом процессе мы производим много лишней информации. Возникает вопрос: может быть, мы всего лишь отложили энергетические затраты и нам придется расплатиться, когда потребуется уничтожить весь этот хлам. Обращаясь к этой проблеме, Беннет указал, что обратимый компьютер может выполнить вычисления до конца, распечатать ответ (логически обратимая операция), а затем совершить все шаги в обратном направлении, чтобы вернуться к начальному состоянию. Эта процедура удаляет избыточную информацию без каких-либо энергетических затрат.

Тогда в принципе нам не нужно тратить энергию для выполнения вычислений. На практике, используемые в настоящее время (необратимые) компьютеры рассеивают энергию, во всяком случае на порядки большую, чем $kT \ln 2$ на один элемент, поэтому с технической точки зрения предел Ландауэра не существенен. Но, поскольку вычислительное «железо» продолжает сокращаться в размерах, преодоление предела Ландауэра может оказаться важным для предотвращения плавления деталей, и тогда обратимые вычисления могут оказаться единственной альтернативой.

• **Демон Максвелла.** Идеи Ландауэра и Беннета привели последнего в 1982 году к примирению демона Максвелла со вторым началом термодинамики. Максвелл рассматривал газ в ящике, разделенном перегородкой на две части: A и B. В перегородке имеется заслонка, которой управляет демон. Наблюдая за молекулами, приближающимися к заслонке, он пропускает быстрые молекулы из A в B, а медленные – из B в A. Следовательно, A охлаждается, а B нагревается с пренебрежимо малыми затратами работы. Тепло без затрат переходит из холодной области в горячую, явно нарушая второе начало термодинамики.

Решение Беннета состоит в том, что демон должен собирать и хранить информацию о молекулах. Если объем памяти демона ограничен, то он не может бесконечно продолжать охлаждение газа; в конце концов эта

информация должна быть удалена. В этот момент мы и рассчитываемся энергией за достигнутое охлаждение. (Если же демон не уничтожает свою запись или мы хотим сделать термодинамический расчет до ее удаления, то с записанной информацией следует связывать некоторую энтропию.)[1]

Во многом эти идеи были предвосхищены еще в 1929 году Лео Сцилардом — настоящим пионером физики информации[2]. В своем анализе демона Максвелла Сцилард предложил понятие *бита* информации (само слово «бит» было введено позднее Тьюки) и связал энтропию $\Delta S = k \ln 2$ с приобретением одного бита (по-видимому, Сцилард не осознавал до конца принцип Ландауэра, согласно которому неизбежных затрат требует именно уничтожение бита информации).

Эти примеры показывают, что работа на стыке физики и информации породила замечательные результаты, представляющие интерес как для физиков, так и для исследователей в области вычислений.

1.2. Квантовая информация

Итак, мы выяснили, что «информация материальна»[3], поэтому поучительно посмотреть, что говорит физика об информации. В своей основе Вселенная является квантово-механической. Как квантовая теория освещает природу информации?

Уже на заре квантовой теории должно было стать ясно, что в свете новой физики классические идеи об информации требуют пересмотра. Например, щелчки, регистрируемые детектором, который следит за радиоактивным источником, описываются *истинно случайным* пуассоновским процессом. Напротив, в детерминистской классической динамике нет места истинной случайности [хотя, конечно, поведение сложной

[1] Популярное изложение идей Беннета можно найти в статье Чарльз Г. Беннет, Демоны, двигатели и второе начало термодинамики, В мире науки №1, 52 (1988), перевод журнала Scientific American. Всестороннее обсуждение этих вопросов см. также в книге Б.Б. Кадомцев, *Динамика и информация*, редакция журнала УФН, М. (1999). — *Прим. ред.*

[2] Классическая работа Сцилларда опубликована на немецком языке: L. Szilard, *Über die Entropieverminderung in Einem Thermodynamischen System bei Eingriffen Intelligenter Wesen*, Zeitschrift für Physik, 53, 840–856 (1929); на английском языке статья: L. Szilard, *On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings* публиковалась по меньшей мере трижды (1964, 1983, 2003гг.) см., например, *Maxwell's Demon 2. Entropy, Classical and Quantum Information, Computing*, Ed. by H.S. Leff and A.F. Rex, IoP Publishing, Bristol, Philadelphia, (2003) pp. 110–119. — *Прим. ред.*

[3] Этот очень емкий тезис фактически представляет собой название статьи Р. Ландауэра: R. Landauer, Information is physical, Phys. Today, May, 23–29 (1991). — *Прим. ред.*

(хаотической) системы может быть практически неотличимо от случайного].

Более того, в квантовой теории некоммутирующие наблюдаемые не могут одновременно иметь точно определенные значения (принцип неопределенности). Фактически измерение одной наблюдаемой A неизбежно влияет на результат последующего измерения наблюдаемой B, если A и B не коммутируют. Следовательно, процесс получения информации о физической системе неизбежно возмущает ее состояние. В классической физике не существует подобного ограничения.

Компромисс между получением информации и возмущением состояния системы тесно связан с квантовой случайностью. Поскольку результат измерения несет в себе элемент случайности, мы не можем извлечь из него информацию о начальном состоянии системы.

То, что получение информации является причиной возмущения, также связано с другим существенным различием между квантовой и классической информацией: квантовую информацию нельзя воспроизвести с абсолютной точностью (принцип невозможности клонирования, анонсированный Вутерсом и Зуреком, а также Диксом в 1982 году). Если бы мы могли сделать точную копию квантового состояния, то мы могли бы измерить наблюдаемую данной копии, не нарушая состояние оригинала и отменяя тем самым принцип возмущения. С другой стороны, ничто не мешает нам точно копировать классическую информацию (приятное свойство, дающее возможность засорять жесткие диски).

Эти свойства квантовой информации существенны, но особенно серьезный аспект, отличающий квантовую информацию от классической, выяснен в работе Джона Белла в 1964 году. Он показал, что никакая локальная теория скрытых параметров не может воспроизвести предсказания квантовой механики. Согласно Беллу, квантовая информация может быть закодирована (и фактически закодирована) в нелокальных корреляциях между различными частями физической системы, в корреляциях, не имеющих классического аналога. Я еще вернусь к теореме Белла в этой лекции, но детально мы обсудим ее позднее.

Изучение квантовой информации как последовательной дисциплины началось в 1980-х и достигло расцвета в 1990-х гг. Многие из основных результатов теории классической информации имеют квантовые аналоги, которые были обнаружены и разработаны в последнее время. Некоторые из них мы обсудим в этом курсе, включая сжатие квантовой информации, пределы классической информации, закодированной в квантовых системах, пределы квантовой информации, надежно пересылаемой по квантовому каналу с помехами (шумом).

1.3. Эффективные квантовые алгоритмы

Учитывая то, что квантовая информация обладает множеством необычных свойств, можно было ожидать, что квантовая теория окажет глубокое влияние на наше понимание вычислений. Но то, что это действительно так, для многих из нас явилось как гром среди ясного неба, произведенный Питером Шором в апреле 1994 года [специалист по вычислительной технике AT&T (Американ Телефон энд Телеграф) и выпускник КАЛТЕХа]. Шор показал, что, по крайней мере в принципе, квантовый компьютер может эффективно факторизовать большое число [1].

Факторизация (поиск простых множителей составного числа) является примером *трудно разрешимой* задачи, обладающей следующими свойствами:

— Найденное решение можно *легко проверить*.

— Но найти это решение *сложно*.

То есть, если p и q — большие простые числа, произведение $n = pq$ может быть вычислено быстро (необходимое число элементарных операций примерно равно $\log_2 p \log_2 q$). Но при заданном n найти p и q очень *сложно*. Время, необходимое для поиска множителей, твердо считается (хотя это никогда не было доказано) суперполиномиальным по $\log n$. То есть с ростом n необходимое время растет, как минимум, быстрее любой степени $\log n$. Наиболее известный алгоритм факторизации («решето числового поля») требует

$$\text{time} \simeq \exp\left[c(\ln n)^{1/3}(\ln \ln n)^{2/3}\right], \tag{1.3}$$

где $c = (64/9)^{1/3} \sim 1{,}9$. Текущее состояние дел таково, что 65-разрядные множители 130-разрядного числа могут быть найдены в течение одного месяца сетью сотен процессоров. Используя это для оценки префактора в уравнении (1.3), мы найдем, что факторизация 400-разрядного числа потребовала бы 10^{10} лет, что равно возрасту Вселенной. Итак, даже с учетом существенного развития технологии, факторизация 400-разрядного числа в ближайшее время останется недоступной.

Проблема факторизации интересна с точки зрения теории сложности, как пример задачи, которая считается трудно разрешимой; то есть задачи,

[1] Полная версия статьи: P. Shor, Polinomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, SIAM J. on Computing, **26**, 1484 (1997); русский перевод в сборнике статей *Квантовый компьютер и квантовые вычисления* под ред. В.А. Садовничего, Ижевск, РХД (1999). — *Прим. ред.*

которая не может быть решена за время, полиномиально зависящее от длины входящего сигнала, в данном случае от $\log n$. Она также имеет и практическое значение, поскольку сложность факторизации лежит в основе систем шифрования с открытым ключом, например, широко используемой схемы RSA[1].

Новый волнующий результат, полученный Шором, состоит в том, что квантовый компьютер может выполнять факторизацию за полиномиальное время, например, за время $O[(\ln n)^3]$. Таким образом, если бы у нас был квантовый компьютер, который мог бы факторизовать 130-разрядное число за один месяц (конечно, у нас его нет, пока по крайней мере!), то, следуя алгоритму Шора, он смог бы факторизовать 400-разрядное число менее, чем за три года. Чем сложнее задача, тем большим преимуществом обладает квантовый компьютер.

Результат Шора пробудил мой собственный интерес к квантовой информации (если бы не Шор, не думаю, что я преподавал бы этот курс). Очень приятно размышлять о проблемах, требующих знания теории сложности, квантовой теории и прикладных наук.

1.4. Квантовая сложность

Конечно, работа Шора имела серьезную предысторию. На то, что квантовая система может выполнять вычисления, было впервые явно указано Полем Бениоффом и независимо Ричардом Фейнманом в 1982 году[2]. В известном смысле интерес к этой проблеме был понятен, принимая во внимание неуклонную тенденцию миниатюризации в микросхемотехнике. Если эта тенденция будет продолжаться, мы неизбежно приблизимся к режиму, в котором квантовая теория исключительно важна для функционирования вычислительных устройств. Возможно, это наблюдение обеспечило некоторую мотивацию после работы Бениоффа. Однако главная мотивация Фейнмана была совершенно иной и весьма интересной. Чтобы понять точку зрения Фейнмана, необходимо более точное математическое описание квантовой информации и квантовых вычислений.

Неделимой единицей классической информации является бит: объект, который может принимать любое из двух значений: 0 или 1. Соответству-

[1] R. Rivest, A. Shamir, L. Adleman. — *Прим. ред.*

[2] P. Benioff, Quantum-Mechanical Hamiltonian Models of Turing Machines, J. Stat. Phys., **29**, 515, (1982); R.P. Feynman, Simulation Physics with Computers Int. J. Theor. Phys., **21**, 467 (1982); русские переводы в сборнике статей *Квантовый компьютер и квантовые вычисления* под ред. В.А. Садовничего, Ижевск, РХД (1999). Впервые идея квантовых вычислений была выдвинута Ю.И. Маниным именно в связи с большей информационной емкостью квантовых систем. См. Ю.И. Манин *Вычислимое и невычислимое*, Сов. Радио, М., 1980. — *Прим. ред.*

ющая единица квантовой информации — квантовый бит или *кубит*. Кубит представляет собой вектор в двумерном комплексном векторном пространстве со скалярным (внутренним) произведением; из уважения к классическому биту будем называть элементы ортонормированного базиса в этом пространстве $|0\rangle$ и $|1\rangle$. Тогда нормированный вектор может быть представлен в виде:

$$|\psi\rangle = a|0\rangle + b|1\rangle, \qquad |a|^2 + |b|^2 = 1, \tag{1.4}$$

где $a, b \in \mathbb{C}$. Мы можем выполнить измерение, которое проецирует $|\psi\rangle$ на базис $|0\rangle$, $|1\rangle$. Результат такого измерения не детерминирован — вероятность того, что в итоге мы получим $|0\rangle$, равна $|a|^2$, а вероятность того, что мы получим $|1\rangle$, равна $|b|^2$.

Квантовое состояние N кубитов можно изобразить вектором в 2^N-мерном пространстве. В качестве ортонормированного базиса в этом пространстве можно выбрать состояния, в которых каждый кубит имеет определенное значение $|0\rangle$ или $|1\rangle$. Их можно обозначить двоичными последовательностями чисел, такими как:

$$|01110010\ldots1001\rangle. \tag{1.5}$$

Произвольный нормированный вектор разлагается в данном базисе как

$$\sum_{x=0}^{2^N-1} a_x|x\rangle, \tag{1.6}$$

где каждой двоичной последовательности сопоставляется номер, равный соответствующему ей числу в двоичной системе счисления и изменяющийся в пределах от 0 до $2^N - 1$. Здесь величины a_x — комплексные числа, удовлетворяющие условию $\sum\limits_x |a_x|^2 = 1$. Если мы измеряем все N кубитов, проецируя каждый из них на базис $\{|0\rangle,\ |1\rangle\}$, то вероятность получения результата $|x\rangle$ равна $|a_x|^2$.

Итак, квантовое вычисление можно описать следующим образом. Мы собираем N кубитов и готовим их в стандартном начальном состоянии, таком как $|0\rangle|0\rangle|0\rangle\cdots|0\rangle$ или $|x = 0\rangle$. Затем *применяем к ним* унитарное преобразование U. (Преобразование U сконструировано как произведение стандартных *квантовых логических вентилей*, унитарных преобразований, которые действуют лишь на несколько кубитов одновременно). После применения U мы измеряем все кубиты путем проецирования на базис $\{|0\rangle,\ |1\rangle\}$. Итогом измерения является результат вычисления. Таким образом, окончательным результатом является классическая информация, которая может быть распечатана на листе бумаги и опубликована в журнале «Физикл Ревью» (Physical Review).

Обратите внимание на то, что реализованный квантовым компьютером алгоритм является *вероятностным*. То есть мы можем выполнить одну и ту же операцию дважды и, вследствие случайности процесса квантового измерения, получить разные результаты. Фактически, квантовый алгоритм порождает распределение вероятностей возможных результатов. (В действительности алгоритм факторизации Шора не гарантирует успеха в получении простых множителей; он достигает цели лишь с определенной вероятностью. Однако этого достаточно, поскольку легко проверить, верны ли найденные множители).

Из данного описания должно быть ясно, что квантовый компьютер, в отличие от классического, должен будет работать в соответствии с другими физическими принципами. Тем не менее он не сможет сделать ничего сверх того, что может делать классический компьютер. Классические компьютеры могут хранить векторы, вращать их и моделировать процесс квантового измерения, проецируя векторы на взаимно ортогональные оси. То есть классический компьютер, несомненно, может сколь угодно точно имитировать (моделировать) квантовый компьютер. Наше представление о том, что является вычислимым, не зависит от того, пользуемся мы классическим или квантовым компьютером.

Но мы должны считаться с тем, как много времени занимает это моделирование. Предположим, у нас есть компьютер, который оперирует с умеренным количеством кубитов, например, $N = 100$. Тогда, чтобы представить типичное квантовое состояние компьютера, нам пришлось бы записать $2^N = 2^{100} \sim 10^{30}$ комплексных чисел! Ни один из существующих или будущих цифровых компьютеров не сможет сделать этого. А выполнение произвольного поворота вектора в пространстве размерности 10^{30} находится далеко за пределами вычислительных способностей любого мыслимого классического компьютера.

(Конечно, N классических битов тоже могут принимать 2^N возможных значений. Но полное описание конфигурации каждого из них очень просто — это двоичная последовательность длиной N. Квантовая информация отличается тем, что полное описание даже одной типичной конфигурации N кубитов очень сложно).

Итак, классический компьютер действительно может имитировать квантовый, но с ростом числа кубитов N имитация становится крайне неэффективной. Квантовая механика *сложна* (с точки зрения вычислений), потому что мы должны работать с огромными матрицами — гильбертово пространство слишком велико. Это наблюдение привело Фейнмана к предположению, что квантовый компьютер может оказаться способным выполнять определенные задания, недосягаемые для любого возможного

классического компьютера (квантовому компьютеру не нужно моделировать *самого себя*!). Похоже, что результат Шора поддерживает эту точку зрения.

Так ли неизбежен этот вывод? В конце концов, моделирование должно предоставить способ определения вероятностей всех возможных результатов окончательного измерения. При классическом моделировании совсем не обязательно следовать полному описанию квантового состояния N кубитов. Нас вполне устроил бы *классический вероятностный алгоритм*, результаты которого не определяются однозначно входом, а возникают в соответствии с тем распределением вероятностей, которое генерируется квантовым вычислением. Мы могли бы рассчитывать на выполнение *локального* моделирования, при котором каждый кубит в каждый момент времени имеет определенное значение, а каждый квантовый вентиль может действовать на кубиты различными возможными способами, которые определяются генератором (псевдо-) случайных чисел. Такое моделирование было бы гораздо проще, чем описание эволюции вектора в экспоненциально огромном пространстве.

Однако вывод теоремы Джона Белла однозначно говорит о том, что такое моделирование осуществить невозможно: не существует *локального вероятностного алгоритма*, способного воспроизводить результаты квантовой механики. Таким образом, хотя доказательство этого отсутствует, кажется весьма правдоподобным, что моделирование квантового компьютера является очень сложной задачей для любого классического компьютера.

Чтобы лучше понять, почему математическое описание квантовой информации с необходимостью такое сложное, представим квантовую систему $3N$ кубитов ($N \gg 1$), состоящую из трех подсистем по N кубитов каждая (называемых подсистемами (1), (2) и (3)). Мы случайным образом выбираем квантовое состояние $3N$ кубитов, а затем разделяем три подсистемы, отправляя (1) в Санта-Барбару, (3) — в Сан-Диего, в то время как (2) остается в Пасадене. Теперь мы бы хотели произвести некоторые измерения, чтобы как можно больше узнать о квантовом состоянии. Чтобы облегчить себе задачу, представим, что мы имеем огромное количество копий состояния системы, поэтому мы можем измерить любую из них, а также какие угодно наблюдаемые[1]. Но при одном условии: нам разрешено выполнять измерения лишь внутри одной из подсистем — коллективные измерения, снимающие границы между подсистемами, запрещены. Тогда для *типичного* состояния системы $3N$ кубитов наши измерения почти ни-

[1]Мы не можем сделать копии неизвестного квантового состояния сами, но можем попросить приятеля приготовить множество идентичных копий состояния (он это может, потому что знает, что делать) и не сообщать нам о том, что он сделал.

чего о нем не скажут. Почти вся информация, отличающая одно состояние от другого, содержится в *нелокальных корреляциях* между результатами измерений в подсистемах (1), (2) и (3). Это и есть те самые нелокальные корреляции, которые Белл считал важнейшей частью физического описания.

Мы увидим, что объем информации можно определить количественно с помощью энтропии (большая энтропия подразумевает незначительную информацию). Если мы выбираем состояние $3N$ кубитов случайно, то почти всегда находим, что энтропия каждой подсистемы очень близка к

$$S \cong N - 2^{-(N+1)}, \tag{1.7}$$

результат, полученный Доном Пейджем. Здесь N — максимально возможное значение энтропии, соответствующее случаю, в котором подсистема вообще не несет доступной информации. Таким образом, рассматривая каждую подсистему независимо, при большом N мы можем иметь доступ лишь к экспоненциально малому количеству информации.

То есть измерения дают очень мало информации, если мы не учитываем корреляции результатов, полученных в Сан-Диего, Пасадене и Санта-Барбаре. В терминологии, которой я пользуюсь, измерение корреляции считается «коллективным» измерением (даже если бы фактически оно было выполнено экспериментаторами, которые наблюдали разные части одной и той же копии состояния, а затем созвонились, чтобы сравнить свои результаты). Измеряя корреляции, мы можем узнать намного больше; в принципе мы можем полностью реконструировать состояние.

Любое удовлетворительное описание состояния $3N$ кубитов должно характеризовать эти исключительно сложные нелокальные корреляции. Вот почему классическое моделирование большой квантовой системы требует огромных ресурсов. (Когда подобные нелокальные корреляции существуют между частями системы, мы говорим, что части «запутаны», имея в виду то, что мы не можем полностью расшифровать состояние системы путем ее деления и изучения отдельных частей).

1.5. Квантовый параллелизм

В 1985 году Дэвид Дойч придал идее Фейнмана более конкретную форму. Дойч подчеркнул, что квантовый компьютер может лучше всего реализовать свой вычислительный потенциал, осуществляя то, что он назвал «квантовым параллелизмом»[1]. Чтобы понять, что это означает, лучше всего рассмотреть пример.

[1] D. Deutsch, Quantum Theory, the Church–Turing Principle and the Universal Quantum Computer, Proc. Roy. Soc., London, **A400**, 97 (1985); русский перевод в сборнике статей *Кван-*

Следуя Дойчу, представим черный ящик, вычисляющий функцию, которая преобразует один бит x в один бит $f(x)$. Мы не знаем, что происходит внутри ящика, но это должно быть нечто сложное, потому что вычисление занимает 24 часа. Существует четыре возможные функции $f(x)$ [поскольку каждая из $f(0)$ и $f(1)$ может принять одно из двух возможных значений], и мы бы хотели знать, что вычисляет ящик. Вычисление обеих функций $f(0)$ и $f(1)$ заняло бы 48 часов.

Но мы не располагаем таким временем; нам нужен ответ через 24 часа, а не через 48. И пусть нас даже устроило бы знание того, является $f(x)$ постоянной $[f(0) = f(1)]$ или сбалансированной $[f(0) \neq f(1)]$[1]. Но даже в этом случае получение ответа займет 48 часов.

Теперь представим квантовый черный ящик, вычисляющий $f(x)$. Конечно, $f(x)$ может быть необратимой, в то время как действие квантового компьютера унитарно и должно быть обратимым, поэтому нам понадобится преобразование U_f, трансформирующее два кубита в два:

$$U_f : |x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle. \tag{1.8}$$

(Этот механизм инвертирует второй кубит, если действие f на первый кубит дает 1, и ничего не делает, если действие f на первый кубит дает 0.) Мы можем определить, является ли $f(x)$ постоянной или сбалансированной, использовав квантовый черный ящик дважды. Но получение одного результата по-прежнему занимает сутки, поэтому такой способ не годится. Можем ли мы получить ответ (за 24 часа), воспользовавшись квантовым черным ящиком *лишь раз*? (Это так называемая «Задача Дойча»).

Так как черный ящик является квантовым компьютером, мы можем выбрать в качестве входящего состояния *суперпозицию* $|0\rangle$ и $|1\rangle$. Если второй кубит приготовлен в начальном состоянии $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, то

$$U_f : |x\rangle\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \rightarrow |x\rangle\frac{1}{\sqrt{2}}(|f(x)\rangle - |1 \oplus f(x)\rangle) =$$

$$= |x\rangle(-1)^{f(x)}\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{1.9}$$

то есть функция f оказывается локализованной в зависящей от x фазе.

товый компьютер и квантовые вычисления под ред. В.А. Садовничего, Ижевск, РХД (1999). *Прим. ред.*

[1]Этот несколько необычный термин обозначает, что два различных значения функции $f(x)$ сбалансированы, то есть каждое из них соответствует ровно половине значений аргумента x. . – *Прим. ред.*

Теперь предположим, что первый кубит приготовлен в начальном состоянии $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Тогда черный ящик действует следующим образом:

$$U_f: \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \rightarrow$$

$$\rightarrow \frac{1}{\sqrt{2}}\left[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\right]\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (1.10)$$

Наконец, мы можем выполнить измерение, проецирующее первый кубит на базис

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle). \quad (1.11)$$

Очевидно, что мы всегда будем получать $|-\rangle$, если функция $f(x)$ сбалансированная, и $-|+\rangle$, если она постоянна[1].

Итак, мы решили задачу Дойча, а также нашли разницу между тем, что доступно классическому компьютеру, а что квантовому. Классическому компьютеру придется воспользоваться черным ящиком дважды, чтобы отличить сбалансированную функцию от постоянной, а квантовый компьютер выполняет это задание за один раз!

Это возможно, потому что квантовый компьютер не ограничивается вычислением только $f(0)$ или $f(1)$. Он может действовать на суперпозицию $|0\rangle$ и $|1\rangle$, извлекая таким образом «глобальную» информацию о функции, информацию, которая зависит и от $f(0)$, и от $f(1)$. Это и есть квантовый параллелизм.

Теперь предположим, что нас интересуют глобальные свойства функции, которая действует на N битов, функции, зависящей от 2^N возможных аргументов. Чтобы вычислить полную таблицу значений $f(x)$, нам пришлось бы считать f 2^N раз, что совершенно невозможно при $N >> 1$ (например, 10^{30} раз для $N = 100$). Но с квантовым компьютером, действующим в соответствии с

$$U_f : |x\rangle|0\rangle \rightarrow |x\rangle|f(x)\rangle, \quad (1.12)$$

мы могли бы выбрать следующее состояние входного регистра:

$$\left[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right]^N = \frac{1}{2^{N/2}}\sum_{x=0}^{2^N-1}|x\rangle, \quad (1.13)$$

[1] В предыдущем описании квантовых вычислений мы говорили, что окончательное измерение проецирует каждый кубит на базис $\{|0\rangle, |1\rangle\}$, но здесь мы допускаем возможность измерения в другом базисе. Чтобы описать эту процедуру в старом базисе, необходимо перед окончательным измерением применить к каждому кубиту подходящее унитарное преобразование.

и, вычислив $f(x)$ только один раз, генерировать состояние:

$$\frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle|f(x)\rangle. \tag{1.14}$$

В этом состоянии закодированы глобальные свойства f, и мы могли бы извлечь некоторые из них, если бы только смогли придумать для этого эффективный способ.

Это квантовое вычисление демонстрирует «массовый квантовый параллелизм»; имитация подготовки такого состояния на классическом компьютере потребовала бы от нас вычислять f невообразимо огромное количество раз (для $N >> 1$). Тем не менее с помощью квантового компьютера мы сделали это в один заход. Это именно тот тип массового параллелизма, который был осуществлен Шором в его алгоритме факторизации.

Как было отмечено ранее, характерной чертой квантовой информации является то, что она может быть закодирована в нелокальных корреляциях между разными частями физической системы. Действительно, этот случай отображен в уравнении (1.14); свойства функции f хранятся в корреляциях между «входным регистром» и «выходным регистром» квантового компьютера. Однако не так просто расшифровать эту нелокальную информацию.

Например, если бы я измерил входной регистр, то получил бы результат $|x_0\rangle$, где x_0 совершенно случайным образом выбрано из 2^N возможных значений. Такая процедура приготовила бы состояние

$$|x_0\rangle|f(x_0)\rangle. \tag{1.15}$$

Мы могли бы перейти к измерению выходного регистра, чтобы получить значение $f(x_0)$. Но у нас нет возможности определить $f(y_0)$ для любого $y_0 \neq x_0$ посредством дальнейших измерений, поскольку уравнение (1.14) разрушено предыдущим измерением и сложные корреляции между регистрами утеряны. Следовательно, в этом случае квантовое вычисление не дает преимущества перед классическим.

Урок, полученный при решении задачи Дойча, состоит в том, что использование закодированных в уравнении (1.14) корреляций иногда требует изобретательности. Искусство создания квантовых алгоритмов во многом заключается в поиске способов эффективного использования нелокальных корреляций.

1.6. Новая классификация сложности

Компьютер на вашем рабочем столе не является квантовым, но все же это замечательное устройство: в принципе, он способен выполнить любое мыслимое вычисление. Но в действительности существуют вычисления,

которые вы не можете выполнить: вам не хватает либо времени, либо памяти. Но если объем вашей памяти неограничен и вы согласны ждать столько, сколько потребуется, тогда все, что достойно называться вычислением, может быть выполнено вашим маленьким ПК. Поэтому мы называем его «универсальным компьютером».

Классическая теория сложности изучает, какие задачи являются сложными, а какие простыми. Обычно понятия «сложная» и «простая» определяются количеством необходимых времени и/или памяти. Но как придать смысл различию между простым и сложным, не описав аппаратные средства, которые мы будем использовать? Задача может быть сложной для ПК, но, наверное, я смог бы разработать устройство специального назначения, которое смогло бы решить ее гораздо быстрее. А может быть в будущем появится более совершенный универсальный компьютер, способный решить эту задачу гораздо эффективнее. Действительно имеющие смысл различия между сложным и простым должны быть универсальными — они не должны зависеть от того, какое устройство мы используем.

Теория сложности главным образом обращает внимание на различие между алгоритмами, выполняемыми за «полиномиальное время» и за «экспоненциальное время». С любым алгоритмом A, действующим на вход переменной длины, можно сопоставить функцию сложности $T_A(N)$, где N — длина входа в битах. $T_A(N)$ представляет собой максимальное «время» (то есть количество элементарных операций), необходимое для полного выполнения алгоритма для любого входа из N битов. [Например, если A — алгоритм факторизации, то $T_A(N)$ — время, необходимое для факторизации N-битового числа в худшем случае.] Мы говорим, что A выполняется за полиномиальное время, если

$$T_A(N) \leqslant \mathrm{Poly}(N), \tag{1.16}$$

где $\mathrm{Poly}(N)$ обозначает полином от переменной N. Следовательно, полиномиальное время означает, что необходимое для решения задачи время растет не быстрее степени количества входящих битов.

Если задача выполняется не за полиномиальное время, то мы говорим, что ее решение требует экспоненциального времени (хотя на самом деле терминологически это не совсем верно, поскольку, конечно, существуют суперполиномиальные функции типа $N^{\log N}$, которые на самом деле возрастают гораздо медленнее экспоненциальных). Это рациональный критерий определения грани между простым и сложным. Но действительно решающим доводом в пользу этого различия служит его независимость от того, какой компьютер мы используем. Универсальность различия между полиномиальным и экспоненциальным следует из одного из главных результатов теории вычислительных систем: универсальный (классический)

компьютер может моделировать другой в худшем случае с «полиномиальным превышением» (времени). Это значит, что если на вашем компьютере алгоритм выполняется за полиномиальное время, то я всегда могу выполнить его на своем компьютере за полиномиальное время. Если я не могу придумать лучший способ сделать это, то я всегда могу эмулировать работу вашего компьютера на своем; ценой эмуляции является лишь полиномиальное время. Так же ваш компьютер может эмулировать мой, то есть мы всегда будем единодушны в том, какие алгоритмы выполняются за полиномиальное время[1].

Итак, истина в том, что информация и вычисления в физическом мире в основе своей квантово-механические. Но это мнение, каким бы дорогим оно ни было для физиков, не представляло бы особого интереса (по крайней мере, с точки зрения теории сложности), если бы было возможно моделирование квантового компьютера классическим с полиномиальным превышением времени. Тогда квантовые алгоритмы представляли бы лишь технический (специальный) интерес, возможно, не больший, чем будущие успехи классических алгоритмов, способных ускорить решение некоторых задач.

Но если, как показывает (но не доказывает!) алгоритм Шора, никакое моделирование квантового компьютера за полиномиальное время невозможно, то это меняет все. Результаты тридцатилетних исследований в теории сложности по-прежнему останутся математическими истинами, как, например, теоремы, характеризующие возможности классических компьютеров. Но они не устоят как физические истины, поскольку классическая машина Тьюринга — неподходящая модель вычислений, которые действительно могут быть выполнены в физическом мире.

Если квантовая классификация сложности действительно отличается от классической (как подозревается, но не доказано), тогда этот результат перевернет основы информатики. В долгосрочном плане это также может сильно повлиять на технологию. Однако какое значение это имеет для физики? Я не уверен. Но, наверное, это говорит о том, что ни одно мыслимое классическое вычисление не может точно предсказать поведение даже скромного числа кубитов (порядка 100). Это наводит на мысль, что сравнительно небольшие квантовые системы имеют больший потенциал, чем мы можем себе представить даже в самых смелых мечтах.

1.7. Как насчет ошибок?

Существует другое, недавно обнаруженное, свойство квантовой информации, такое же важное, каким может оказаться алгоритм факториза-

[1] Чтобы сделать это утверждение точным, мы должны соблюдать некоторую осторожность. Например, следует исключить некоторые виды «неуместных» машин, как, например, параллельная сеть компьютеров с неограниченным числом станций.

ции Шора: открытие коррекции квантовых ошибок. Действительно, если бы не этот результат, перспективы квантовых вычислительных технологий не казались бы такими радужными.

Как мы уже отмечали, основным свойством квантовой информации, которую использует квантовый компьютер, является наличие нелокальных корреляций между разными частями физической системы. Если я наблюдаю лишь часть системы за раз, то я могу извлечь только очень малую долю закодированной в ней информации.

К сожалению, эти нелокальные корреляции крайне хрупкие и на практике склонны очень быстро распадаться. Проблема в том, что наша квантовая система неизбежно контактирует с намного большей системой — с окружающей ее средой. Полностью изолировать большую квантовую систему от ее окружения практически невозможно, даже если для этого мы предпримем героические усилия. Взаимодействия между квантовым устройством и окружающей средой устанавливают нелокальные корреляции между ними. В итоге квантовая информация, изначально закодированная нами в устройстве, вместо этого оказывается закодированной в корреляциях между устройством и окружающей средой. На этой стадии мы уже не можем получить доступ к информации, наблюдая только за устройством. На практике информация безвозвратно потеряна. С макроскопическим устройством это происходит очень быстро, даже если его связь с окружением достаточно слабая.

Эрвин Шредингер критиковал сторонников основного направления интерпретации квантовой механики, обращая внимание на то, что теория допускает квантовое состояние кота в форме

$$|\mathrm{cat}\rangle = \frac{1}{\sqrt{2}}\big(|\mathrm{dead}\rangle + |\mathrm{alive}\rangle\big). \tag{1.17}$$

В возможности таких состояний Шредингер видел дефект теории, потому что каждый встречавшийся ему кот был либо живым, либо мертвым, а не полуживым-полумертвым.

Одно из наиболее важных достижений квантовой теории за последние 15 лет состоит в том, что мы узнали, как аргументированно ответить Шредингеру. Состояние $|\mathrm{cat}\rangle$ в принципе возможно, но редко наблюдается вследствие его *крайней нестабильности*. Встречавшиеся Шредингеру коты никогда не были достаточно изолированы от окружающей среды. Если бы кто-нибудь приготовил состояние $|\mathrm{cat}\rangle$, то квантовая информация, закодированная в суперпозиции $|\mathrm{dead}\rangle$ и $|\mathrm{alive}\rangle$, *немедленно* переместилась бы в корреляции между котом и окружающей средой, став тем самым полностью недоступной. В действительности окружающая среда постоянно измеряет кота, проецируя его на состояние $|\mathrm{alive}\rangle$ или $|\mathrm{dead}\rangle$. Этот процесс

называется *декогерентизацией*. Мы еще вернемся к изучению декогерентизации в этом курсе.

Итак, чтобы выполнить сложное квантовое вычисление, мы должны приготовить хрупкую суперпозицию состояний относительно большой квантовой системы (хотя, возможно, и не такой большой, как кот). К сожалению, эту систему нельзя полностью изолировать от окружающей среды, поэтому эта суперпозиция, как состояние $|\text{cat}\rangle$, очень быстро распадается. Закодированная квантовая информация быстро теряется и наш квантовый компьютер терпит банкротство.

Другими словами, контакт компьютера с окружающей средой (декогерентизация) служит причиной *ошибок*, разрушающих квантовую информацию. Для того чтобы обеспечить надежную работу квантового компьютера, необходимо найти какой-нибудь способ предотвращения или исправления этих ошибок.

На самом деле декогерентизация — не единственная проблема. Мы не могли бы ожидать безупречно точной работы квантового компьютера, даже если бы добились его полной изоляции от окружающей среды. Реализованные в машине квантовые вентили представляют собой унитарные преобразования, которые одновременно оперируют несколькими кубитами, скажем, унитарные 4×4-матрицы, действующие на два кубита. Конечно, эти унитарные матрицы образуют континуум. Мы можем иметь протокол применения U_0 к двум кубитам, но его выполнение не будет безупречным, поэтому фактическое преобразование

$$U = U_0(1 + O(\varepsilon)) \tag{1.18}$$

будет отличаться от запланированного U_0 на некоторую величину порядка ε. Накопление этих ошибок после применения около $1/\varepsilon$ вентилей приведет к серьезной неудаче. Подобные трудности испытывают и классические аналоговые приборы, тогда как для устройств, работающих на основе дискретной логики, малые ошибки создают гораздо меньше проблем.

В действительности современные цифровые цепи удивительно надежны. Столь высокой точности они достигают, благодаря диссипации. Мы можем представить классический вентиль, действующий на бит, который изображается мячом, находящимся в одном из минимумов двухъямного потенциала. Вентиль может перебросить мяч через промежуточный барьер на другую сторону потенциала. Конечно, действие вентиля не будет идеальным; он может ударить по мячу чуть сильнее. Со временем эти несовершенства могут накопиться и явиться причиной ошибки.

Чтобы исправить положение, мы охлаждаем бит (в буквальном смысле этого слова) после каждого вентиля. Это диссипативный процесс, при кото-

ром высвобождается тепло в окружающую среду и сокращается доступный мячу фазовый объем, приближая его к локальному минимуму потенциала. Поэтому малые ошибки, которые мы можем совершить, скорее ликвидируются путем выброса тепла в окружающую среду, нежели подвергнут риску работу устройства.

Но мы не можем подобным образом охлаждать квантовый компьютер. Контакт с окружающей средой может повысить достоверность классической информации, но закодированную квантовую информацию он разрушит. В более широком смысле аккумуляция ошибок будет проблемой и для обратимых классических вычислений. Чтобы предотвратить накопление ошибок, нужно избавляться от несущей их информации, а удаление информации — всегда диссипативный процесс.

И все же, не будем сдаваться раньше времени. Для борьбы с ошибками в классической информации был разработан изощренный аппарат — теория кодов, исправляющих ошибки. В какой степени можно воспользоваться этим опытом, чтобы также защитить и квантовую информацию?

Как работает классическая коррекция ошибок? Простейшим примером классического кода, исправляющего ошибки, является код повторения: мы заменяем бит, который хотим защитить, его тремя копиями:

$$\begin{aligned} 0 &\to (000), \\ 1 &\to (111). \end{aligned} \tag{1.19}$$

Теперь пусть может возникнуть ошибка, которая инвертирует один из трех битов; если это, скажем, первый бит, то

$$\begin{aligned} (000) &\to (100), \\ (111) &\to (011). \end{aligned} \tag{1.20}$$

Несмотря на эту ошибку, мы по-прежнему можем правильно декодировать бит путем подсчета простого большинства голосов.

Конечно, если вероятность ошибки в каждом бите равна p, то возможно инвертирование двух битов из трех или даже всех трех битов. Двойное инвертирование может произойти в трех разных случаях, таким образом, его вероятность равна $3p^2(1-p)$, в то время как вероятность тройного инвертирования равна p^3. Тогда в целом вероятность того, что подсчет простого большинства потерпит неудачу, равна $3p^2(1-p)+p^3 = 3p^2 - 2p^3$. Но при

$$3p^2 - 2p^3 < p, \quad \text{или} \quad p < \frac{1}{2}, \tag{1.21}$$

код увеличивает достоверность информации.

Мы можем еще больше увеличить достоверность, используя более длинный код. Один такой код (хотя и далеко не самый эффективный) —

код повторения N-битов. Согласно центральной предельной теореме, при $N \to \infty$ распределение вероятностей для среднего значения бита стремится к гауссовскому с шириной $1/\sqrt{N}$. Если $P = \frac{1}{2} + \varepsilon$ — вероятность того, что каждый бит имеет истинное значение, тогда вероятность того, что подсчет простого большинства потерпит неудачу (для большого N), определяется хвостом распределения Гаусса и равна

$$P_{errtot} \sim \mathrm{e}^{-N\varepsilon^2}. \tag{1.22}$$

Таким образом, для любого $\varepsilon > 0$, вводя достаточное количество вспомогательных битов, можно достичь сколь угодно высокой надежности. Все в порядке будет даже при $\varepsilon < 0$, если учитывать, что в этом случае большинство голосов отдается ошибочному результату. Лишь при $P = \frac{1}{2}$ эта схема не обоснована, ибо тогда блок из N битов будет случайным и не будет содержать никакой информации.

В 1950-х гг. Джон Фон Нейман показал, что классический компьютер с шумящими элементами может надежно работать, используя достаточное количество вспомогательных битов. Он обратил внимание на то, что при необходимости каждую логическую операцию можно выполнить многократно и получить мажоритарный результат. (Фон Нейману было особенно интересно, почему его мозг так хорошо функционировал, несмотря на ненадежность нейронов. Думаю, что ему было приятно найти объяснение своей сообразительности).

Но теперь мы хотим использовать коррекцию ошибок для того, чтобы сохранить работоспособность *квантового компьютера*. Нетрудно видеть связанные с этим трудности:

1. Фазовые ошибки. Квантовая информация более подвержена ошибкам. В дополнение к ошибкам инвертирования битов

$$\begin{aligned} |0\rangle &\to |1\rangle, \\ |1\rangle &\to |0\rangle \end{aligned} \tag{1.23}$$

в ней также могут появляться и фазовые ошибки:

$$\begin{aligned} |0\rangle &\to |0\rangle, \\ |1\rangle &\to -|1\rangle. \end{aligned} \tag{1.24}$$

Фазовая ошибка влечет за собой серьезные последствия, потому что она превращает состояние $\frac{1}{\sqrt{2}}\big[|0\rangle + |1\rangle\big]$ в ортогональное ему состояние $\frac{1}{\sqrt{2}}\big[|0\rangle - |1\rangle\big]$. Однако классическое кодирование не обеспечивает защиты от фазовых ошибок.

2. Малые ошибки. Как уже отмечалось, квантовая информация непрерывна. Если кубит находится в состоянии

$$a|0\rangle + b|1\rangle, \tag{1.25}$$

то ошибка может изменить a и b на величину порядка ε. Со временем эти малые ошибки могут накапливаться. Классические же методы предназначены для исправления больших ошибок (ошибок инвертирования битов).

3. Измерение — причина возмущения. Согласно схеме подсчета простого большинства голосов, для обнаружения и исправления ошибок нужно было измерять биты в коде. Однако нельзя измерять кубиты, не возмущая закодированную в них информацию.

4. Невозможность клонирования. При классическом кодировании информация защищалась путем создания ее дополнительных копий. Однако известно, что квантовую информацию нельзя воспроизвести с абсолютной точностью.

1.8. Квантовые коды, корректирующие ошибки

Несмотря на эти трудности, оказывается, что квантовая коррекция ошибок действительно возможна. Первый пример квантового кода, исправляющего ошибки, был построен около двух лет назад (угадайте, кем!) Питером Шором. Это открытие привело к возникновению новой, удивительно быстро развившейся, дисциплины — теории квантовых кодов коррекции ошибок. Мы рассмотрим ее позже в этом курсе.

По-видимому, проще всего понять принцип работы квантовой коррекции ошибок, рассматривая оригинальный код Шора. Это наиболее простое квантовое обобщение классического трехбитового кода повторения.

Давайте еще раз рассмотрим этот трехбитовый код, но на этот раз учитывая требование, что в случае с квантовым кодом мы должны быть в состоянии исправлять ошибки без измерения какой бы то ни было закодированной информации.

Предположим, что мы кодируем один кубит тремя кубитами:

$$\begin{aligned} |0\rangle &\to |\bar{0}\rangle = |000\rangle, \\ |1\rangle &\to |\bar{1}\rangle = |111\rangle, \end{aligned} \tag{1.26}$$

другими словами, мы кодируем суперпозицию

$$a|0\rangle + b|1\rangle \to a|\bar{0}\rangle + b|\bar{1}\rangle = a|000\rangle + b|111\rangle. \tag{1.27}$$

Мы хотели бы суметь исправить ошибку инвертирования бита, не разрушая эту суперпозицию.

Конечно, нельзя измерять значение одного кубита. Если я измерил первый кубит и получил результат $|0\rangle$, то это значит, что я приготовил состояние $|\bar{0}\rangle$ для всех трех кубитов, и мы потеряли квантовую информацию, закодированную в коэффициентах a и b.

Но нет никакой необходимости ограничиваться измерением одного кубита. Я мог бы также выполнить коллективное измерение на двух кубитах сразу, и этого достаточно для диагностики ошибки инвертирования бита. Для трехкубитового состояния $|x, y, z\rangle$ я мог бы измерить, скажем, двухкубитовые наблюдаемые $y \oplus z$ или $x \oplus z$ (где $\oplus$ обозначает сложение по модулю два). Как для $|x, y, z\rangle = |000\rangle$, так и для $|x, y, z\rangle = |111\rangle$ результат был бы равен нулю, но если какой-нибудь бит инвертируется, тогда по крайней мере одна из этих величин будет равна единице. Фактически, если инвертируется один бит, то два бита

$$(y \oplus z, x \oplus z) \tag{1.28}$$

непосредственно определяют в двоичной записи позицию (1, 2 или 3) инвертированного бита. Эти два бита составляют *синдром*, диагностирующий появившуюся ошибку.

Например, если инвертировался первый бит

$$a|000\rangle + b|111\rangle \rightarrow a|100\rangle + b|011\rangle, \tag{1.29}$$

тогда измерение $(y \oplus z, x \oplus z)$ дает результат $(0, 1)$, информирующий нас о необходимости инвертировать первый бит; это действительно исправляет ошибку.

Конечно, вместо (большой ошибки) инвертирования бита, возможна малая ошибка:

$$\begin{aligned} |000\rangle &\rightarrow |000\rangle + \varepsilon|100\rangle, \\ |111\rangle &\rightarrow |111\rangle - \varepsilon|011\rangle. \end{aligned} \tag{1.30}$$

Но даже в этом случае вышеописанная процедура будет прекрасно работать. Измеряя $(y \oplus z, x \oplus z)$, мы восстанавливаем собственное состояние этой наблюдаемой. В большинстве случаев (с вероятностью $1 - |\varepsilon|^2$) мы получаем результат $(0, 0)$ и проецируем поврежденное состояние на исходное, исправляя таким образом ошибку. Иногда (с вероятностью $|\varepsilon|^2$) мы получаем результат $(0, 1)$ и проецируем состояние на уравнение (1.29). Но тогда синдром приказывает нам инвертировать первый бит, что восстанавливает исходное состояние. Аналогично, если амплитуда вероятности инвертирования каждого из трех кубитов имеет порядок ε, тогда измерение синдрома с вероятностью порядка $|\varepsilon|^2$ спроецирует состояние на то,

в котором один из трех битов инвертирован, а синдром укажет — какой из них.

Итак, мы преодолели три из четырех ранее упомянутых трудностей. Мы видим, что без ущерба для информации можно выполнить диагностирующее ошибку измерение [ответ на пункт (3)]. Квантовое измерение может проецировать состояние с малой ошибкой на состояние без ошибки или на состояние с большой дискретной ошибкой, способ исправления которой нам известен [ответ на пункт (2)]. Что касается пункта (4), то эта проблема вообще не возникла, поскольку состояние $a|\bar{0}\rangle + b|\bar{1}\rangle$ получено не клонированием — оно не совпадает с $(a|0\rangle + b|1\rangle)^3$; то есть не образовано тремя копиями незакодированного состояния.

Остается только одна проблема: (1) фазовые ошибки. Наш код пока не обеспечивает никакой защиты от них. Если в любом одном из трех кубитов возникнет фазовая ошибка, тогда наше закодированное состояние $a|\bar{0}\rangle + b|\bar{1}\rangle$ преобразуется в $a|\bar{0}\rangle - b|\bar{1}\rangle$ и закодированная квантовая информация разрушится. В действительности *использование кода троекратного повторения* втрое увеличивает частоту возникновения фазовых ошибок. Но, располагая методами, преодолевшими проблемы (2)–(4), мы можем с уверенностью подойти и к первой проблеме. Введя вспомогательные (дополнительные) биты, мы защитились от ошибок инвертирования битов. Это подсказывает как защититься от фазовых ошибок с помощью кодирования вспомогательных (дополнительных) фаз.

Следуя Шору, закодируем один кубит девятью

$$
\begin{aligned}
|0\rangle \to |\bar{0}\rangle &= \frac{1}{2^{3/2}}\big(|000\rangle + |111\rangle\big)\big(|000\rangle + |111\rangle\big)\big(|000\rangle + |111\rangle\big),\\
|1\rangle \to |\bar{1}\rangle &= \frac{1}{2^{3/2}}\big(|000\rangle - |111\rangle\big)\big(|000\rangle - |111\rangle\big)\big(|000\rangle - |111\rangle\big).
\end{aligned}
\tag{1.31}
$$

$|\bar{0}\rangle$ и $|\bar{1}\rangle$ состоят из трех кластеров, в каждом из которых по три кубита, причем каждый кластер приготовлен в одном и том же квантовом состоянии. Каждый из кластеров имеет три вспомогательных бита, поэтому мы можем исправить инвертирование одного бита в любом кластере описанным выше методом.

Предположим, что в одном из кластеров происходит обращение фазы. Ошибка изменяет относительный знак $|000\rangle$ и $|111\rangle$ в этом кластере так, что

$$
\begin{aligned}
|000\rangle + |111\rangle &\to |000\rangle - |111\rangle,\\
|000\rangle - |111\rangle &\to |000\rangle + |111\rangle.
\end{aligned}
\tag{1.32}
$$

Это значит, что относительная фаза поврежденного кластера отличается от фаз двух других кластеров. Таким образом, как и при исправлении инвертированного бита, мы можем определить поврежденный кластер, не измеряя относительные фазы в каждом из них (что вызвало бы возмущение закодированной информации), а сравнивая фазы пар кластеров. В этом случае для проведения сравнения нам необходимо измерить шестикубитовую наблюдаемую, например, наблюдаемую, которая инвертирует от одного до шести кубитов. Поскольку двукратное инвертирование является тождественным преобразованием, квадрат этой наблюдаемой равен единице, а ее собственные значения ± 1. Пара кластеров с одинаковым знаком представляет собой собственное состояние с собственным значением $+1$, а пара кластеров с противоположными знаками — собственное состояние с собственным значением -1. Измерив шестикубитовую наблюдаемую для второй пары кластеров, мы можем определить, какой из кластеров имеет отличный от других знак. Тогда мы применяем унитарное фазовое преобразование к одному из кубитов в этом кластере, чтобы обратить знак и исправить ошибку.

Предположим, что унитарная ошибка $U = 1 + O(\varepsilon)$ возникает в каждом из девяти кубитов. Наиболее общее однокубитовое унитарное преобразование (с точностью до физически несущественной общей фазы) может быть разложено в первом порядке по ε:

$$U = 1 + i\varepsilon_x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + i\varepsilon_y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + i\varepsilon_z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (1.33)$$

Три слагаемых порядка ε в этом разложении можно интерпретировать как оператор инвертирования бита, оператор обращения фазы и оператор, совершающий обе эти операции. Если мы приготовим закодированное состояние $a|\bar{0}\rangle + b|\bar{1}\rangle$, в котором появление унитарных ошибок возможно в каждом кубите, а затем измерим синдромы инвертирования бита и обращения фазы, тогда в большинстве случаев мы вернем состоянию его первоначальный вид. Но с вероятностью порядка $|\varepsilon|^2$ в одном из кубитов возникнет большая ошибка: инвертирование бита, обращение фазы или и то и другое. Благодаря синдрому, мы узнаем, какой из битов инвертировался и в каком из кластеров появилась фазовая ошибка, тогда для исправления ошибки можно применить соответствующее однокубитовое унитарное преобразование.

Исправление ошибки не удастся, если, согласно измерению синдрома, в каждом из двух кластеров имеется две ошибки инвертирования бита (что влечет за собой фазовую ошибку в закодированной информации) или если фазовые ошибки возникнут в двух разных кластерах (что сводится к ошибке инвертирования бита в закодированной информации). Но вероятность

такой двойной фазовой ошибки имеет порядок $|\varepsilon|^4$. Поэтому при достаточно малой $|\varepsilon|$ кодирование увеличивает надежность квантовой информации.

Код также защищает от декогерентизации. Восстанавливая квантовое состояние, независимо от природы ошибки, наша процедура устраняет любое запутывание между квантовым состоянием и его окружением.

Как обычно, коррекция ошибок представляет диссипативный процесс, поскольку информация о природе ошибок изгоняется из квантовой системы. В этом случае, когда информация является неотъемлемой частью записанных результатов наших измерений, при удалении этой записи будет выделяться тепло.

Позднее в этом курсе мы обсудим дальнейшие результаты в области коррекции квантовых ошибок, включая такие как:

- Как и в случае классического кодирования, оказывается, что существуют «хорошие» квантовые коды, позволяющие добиться сколь угодно высокой надежности, пока доля ошибок на один кубит достаточно мала.

- Мы предположили, что процедура исправления ошибок сама по себе выполняется безупречно. Но измерение синдрома оказалось сложным — для выявления ошибок нам было необходимо измерить двухкубитовую и шестикубитовую коллективные наблюдаемые — так что, пытаясь исправить информацию, фактически мы могли нанести ей еще больший ущерб. Однако мы покажем, что коррекция ошибок может быть выполнена так, что она будет эффективно работать даже при появлении случайных ошибок в самом процессе восстановления.

- Чтобы оперировать с квантовым компьютером, мы хотим не только надежно хранить информацию, но также и обрабатывать ее. Мы покажем, что возможно применение квантовых вентилей к закодированной информации.

Подытожим важнейшие идеи, которые лежат в основе квантовой схемы коррекции ошибок:

1) Мы перевели ошибки в цифровую форму. Даже если ошибки в квантовой информации были малыми, мы выполнили измерение, которое проецировало наше состояние на состояние без ошибок либо на состояние с одной из дискретного множества ошибок, способ исправления которой известен.

2) Мы измеряли ошибки, не измеряя информацию. Наши измерения вскрывали природу ошибки, не вскрывая при этом (и, следовательно, не возмущая) закодированную информацию.

3) Ошибки локальны, а закодированная информация нелокальна. Необходимо подчеркнуть основное предположение, лежащее в основе строения кода — ошибки, повреждающие разные кубиты, в хорошем приближении некоррелированы. Мы неявно предположили, что событие, вызывающее ошибку в двух кубитах, гораздо менее вероятно, чем событие, вызывающее ошибку в одном кубите. Конечно, это вопрос физики, обоснованно это предположение или нет — нетрудно представить себе процесс, который станет причиной возникновения ошибок в двух кубитах сразу. Если подобные коррелированные ошибки окажутся достаточно распространенными, то кодирование потерпит неудачу в повышении надежности.

Код пользуется предполагаемой локальной природой ошибок, кодируя информацию нелокальным способом, то есть информация хранится в корреляциях, охватывающих несколько кубитов. Невозможно отличить $|\bar{0}\rangle$ от $|\bar{1}\rangle$, измеряя один кубит из девяти. Если мы измерим один кубит, то с вероятностью $1/2$ получим $|0\rangle$ или $|1\rangle$, независимо от значения закодированного кубита. Чтобы получить доступ к закодированной информации, нам необходимо измерить трехкубитовую наблюдаемую (оператор, инвертирующий все три кубита в кластере, может отличить $|000\rangle + |111\rangle$ от $|000\rangle - |111\rangle$).

Окружающая среда может случайно повлиять на один из кубитов, на самом деле «измеряя» его. Но закодированная информация не может быть повреждена возмущением одного кубита, потому что сам по себе одиночный кубит фактически вообще не несет информации. Нелокально закодированная информация невосприимчива к локальным воздействиям — это основной принцип, на котором основаны квантовые коды коррекции ошибок.

1.9. Квантовое «железо»

Теоретические разработки в области квантовой сложности и квантовой коррекции ошибок сопровождались первыми экспериментальными попытками обработки когерентной квантовой информации. Здесь я кратко опишу некоторые из этих опытов[1].

Чтобы создать аппаратное обеспечение для квантового компьютера, необходима технология, позволяющая управлять кубитами. «Железу» придется столкнуться с некоторыми жесткими техническими требованиями:

[1]Это единственный раздел, в котором очень кратко говорится об аппаратном обеспечении и физических реализациях квантовых вычислений. Читателям, интересующимся этими вопросами, можно порекомендовать книги [1–5] из списка литературы к предисловию, а также цитированную в них литературу. — *Прим. ред.*

1. **Хранение:** Необходимо хранить кубиты в течение длительного времени, достаточного для выполнения интересующих нас вычислений.

2. **Изоляция:** Необходима надежная изоляция кубитов от окружения, чтобы минимизировать ошибки, возникающие вследствие декогерентизации.

3. **Считывание:** Необходимо эффективно и достоверно измерять кубиты.

4. **Вентили:** Необходимо управлять квантовыми состояниями отдельных кубитов и индуцировать контролируемые взаимодействия между ними так, чтобы можно было выполнять квантовые операции.

5. **Точность:** Для надежности работы устройства необходима высокая точность выполнения квантовых операций.

1.9.1. Ионная ловушка

Один из возможных путей достижения этих целей был предложен Игнасио Цираком и Питером Цоллером; его дальнейшей разработкой занялась группа Дэвида Вайнленда из национального института стандартов и технологий (NIST), а также и другие группы. В этой схеме каждый кубит переносится одним ионом, удерживаемым в линейной ловушке Пауля. Квантовое состояние каждого иона – это линейная комбинация основного состояния $|g\rangle$ (интерпретируемого как $|0\rangle$) и особого долгоживущего метастабильного возбужденного состояния $|e\rangle$ (интерпретируемого как $|1\rangle$). Когерентная линейная комбинация двух уровней,

$$a|g\rangle + be^{i\omega t}|e\rangle, \tag{1.34}$$

может существовать в течение времени, сравнимого со временем жизни возбужденного состояния (хотя, конечно, относительная фаза осциллирует вследствие расщепления энергии $\hbar\omega$ между уровнями). Ионы настолько хорошо изолированы, что доминирующей формой декогерентизации может быть спонтанный распад.

Нетрудно считать информацию о состоянии ионов с помощью измерения, проецирующего на базис $\{|g\rangle, |e\rangle\}$. Пусть лазер настроен на переход из состояния $|g\rangle$ в короткоживущее возбужденное состояние $|e'\rangle$. Когда лазер освещает ионы, каждый кубит со значением $|0\rangle$ многократно поглощает и снова излучает свет лазера и таким образом становится видимым (флуоресценция). Кубиты со значением $|1\rangle$ остаются невидимыми.

Вследствие их взаимного кулоновского отталкивания, ионы достаточно хорошо изолированы и на каждый из них можно индивидуально направить импульсы лазеров. Если лазер настроен на частоту перехода ω и сфокусирован на n-ом ионе, то между состояниями $|0\rangle$ и $|1\rangle$ возбуждаются осцилляции Раби. При подходящем выборе продолжительности и фазы лазерного импульса мы сможем реализовать любое однокубитовое унитарное преобразование. В частности, действуя на $|0\rangle$, лазерный импульс может приготовить любую желаемую линейную комбинацию $|0\rangle$ и $|1\rangle$.

Но наиболее сложной частью разработки и создания аппаратного обеспечения квантовых вычислений является организация взаимодействия двух кубитов между собой. В ионной ловушке взаимодействия обусловлены кулоновским отталкиванием между ионами, вследствие чего возникает спектр связанных нормальных мод колебаний захваченных ионов. Когда ион поглощает или излучает лазерный фотон, его центр масс смещается. Но если лазер настроен подходящим образом, то при поглощении или излучении одним ионом произойдет когерентное смещение множества вовлеченных в нормальную моду ионов (эффект Мёссбауэра).

Наиболее низкочастотной колебательной модой (частота ν) является мода центра масс (cm), в которой ионы синхронно колеблются в гармонических ямах ловушек. Ионы можно охладить лазером до температур гораздо меньших ν, так что каждая колебательная мода с большой вероятностью находится в своем основном квантово-механическом состоянии. Теперь представим, что настроенный на частоту $\omega - \nu$ лазер светит на n-ый ион. При должной длительности импульса состояние $|e\rangle_n$ перейдет в $|g\rangle_n$, в то время как cm-осциллятор совершит переход из его основного состояния $|0\rangle_{cm}$ в первое возбужденное $|1\rangle_{cm}$ (рождение cm-«фонона»). В то же время состояние $|g\rangle_n|0\rangle_{cm}$ не находится в резонансе для любого перехода и поэтому не меняется под влиянием лазерного импульса. Таким образом, лазерный импульс совершает унитарное преобразование, действующее как

$$\begin{aligned} |g\rangle_n|0\rangle_{cm} &\to |g\rangle_n|0\rangle_{cm}, \\ |e\rangle_n|0\rangle_{cm} &\to -i|g\rangle_n|1\rangle_{cm}. \end{aligned} \tag{1.35}$$

Эта операция удаляет бит информации, первоначально хранившийся во внутреннем состоянии n-го иона, и помещает его в коллективное состояние движения всех ионов.

Это означает, что внутреннее состояние n-го иона оказало влияние на состояние движения m-го иона ($m \neq n$). В этом смысле нам удалось индуцировать взаимодействие между ионами. Для завершения квантовой операции мы должны переместить квантовую информацию от cm-фоно-

на обратно во внутреннее состояние одного из ионов. Процедура должна быть построена таким образом, чтобы после выполнения операции cm-мода возвращалась в ее основное состояние $|0\rangle_{cm}$. Например, Цирак и Цоллер показали, что квантовый XOR-вентиль (исключающее ИЛИ, или контролируемое НЕ)

$$|x, y\rangle \rightarrow |x, y \oplus x\rangle \tag{1.36}$$

может быть выполнен в ионной ловушке всего пятью лазерными импульсами. Обусловленное этим возбуждение фонона (1.35) для одного захваченного в ловушку иона было продемонстрировано экспериментально группой из NIST.

Серьезным недостатком компьютера на ионных ловушках является то, что по своей природе это медленно работающее устройство. Очевидно, его быстродействие ограничено соотношением неопределенности энергия–время. Поскольку неопределенность энергии лазерного фотона должна быть мала по сравнению с характерной колебательной энергией ν, продолжительность каждого лазерного импульса должна быть велика по сравнению с ν^{-1}. На практике ν, как правило, имеет порядок 100 кГц.

1.9.2. КЭД-резонатор

Другое направление разработки аппаратного обеспечения (предложенное Пеллицари, Гардинером, Цираком и Цоллером) поддерживает группа Джефа Кимбла здесь, в КАЛТЕХе. Идея состоит в том, чтобы захватить несколько нейтральных атомов в маленький с высокой точностью изготовленный оптический резонатор[1]. Вновь квантовая информация может храниться во внутренних состояниях атомов. Но здесь атомы взаимодействуют, благодаря связи с нормальными модами электромагнитного поля в резонаторе (вместо колебательных мод ионной ловушки). Снова, возбуждая переходы импульсами лазеров, мы можем вызвать в одном атоме переход, обусловленный внутренним состоянием другого атома.

Другая возможность хранения кубита – использование поляризации фотона вместо внутреннего состояния иона. Тогда захваченный атом может использоваться как посредник, обеспечивающий взаимодействие одного фотона с другим (вместо фотона, использовавшегося для связи одного атома с другим). Эти эксперименты с «летающим кубитом» продолжаются уже два года. Группа Кимбла продемонстрировала действие двухфотонного квантового вентиля, в котором циркулярная поляризация одного фотона

[1] Квантово-электродинамический (КЭД) резонатор. – *Прим. ред.*

влияет на фазу другого фотона:

$$\begin{aligned} |L\rangle_1|L\rangle_2 &\to |L\rangle_1|L\rangle_2, \\ |L\rangle_1|R\rangle_2 &\to |L\rangle_1|R\rangle_2, \\ |R\rangle_1|L\rangle_2 &\to |R\rangle_1|L\rangle_2, \\ |R\rangle_1|R\rangle_2 &\to e^{i\Delta}|R\rangle_1|R\rangle_2, \end{aligned} \tag{1.37}$$

где $|L\rangle$, $|R\rangle$ обозначают состояния фотонов с левой и правой циркулярной поляризацией. Чтобы добиться этого взаимодействия, один фотон хранится в резонаторе, где он, находясь в состоянии с поляризацией $|L\rangle$, не взаимодействует с атомом и, напротив, сильно связан с ним, будучи в состоянии с поляризацией $|R\rangle$. Второй фотон пересекает резонатор, и он также преимущественно взаимодействует с атомом, находясь в состоянии с одной определенной поляризацией. Волновой пакет второго фотона приобретает некоторый фазовый сдвиг $e^{i\Delta}$, если только оба фотона имеют $|R\rangle$ поляризацию. Так как фазовый сдвиг обусловлен поляризацией обоих фотонов, это нетривиальный двухкубитовый квантовый вентиль.

1.9.3. ЯМР

Третья схема аппаратного обеспечения (темная лошадка) появилась в прошлом году и обошла ионную ловушку и КЭД-резонатор, захватив лидерство в когерентной квантовой обработке. Новая схема использует технику ядерного магнитного резонанса (ЯМР). Здесь носителями кубитов служат ядерные спины определенного типа молекул. Каждый спин может быть ориентирован по ($|\uparrow\rangle = |0\rangle$) или против ($|\downarrow\rangle = |1\rangle$) направления приложенного постоянного магнитного поля. Спины имеют большое время релаксации или декогерентизации, поэтому кубиты могут сохраняться в течение приемлемого времени.

Мы можем также включить импульсное вращающееся магнитное поле с частотой ω (где ω – расщепление энергии между состояниями спин вверх и спин вниз) и возбудить осцилляции Раби между двумя спиновыми состояниями. При подходящей длительности импульса, мы можем выполнить желаемое унитарное преобразование на отдельном спине (точно так же, как в случае ионной ловушки). Все спины в молекуле испытывают воздействие вращающегося магнитного поля, но отзываются на него лишь те, которые находятся в резонансе.

Более того, между спинами существуют диполь-дипольные взаимодействия, и эта связь может использоваться для выполнения операций. Расщепление между $|\uparrow\rangle$ и $|\downarrow\rangle$ для одного спина фактически зависит от состояния

соседних спинов. Следовательно, находится или нет управляющий импульс в резонансе, чтобы опрокинуть спин, обусловлено состоянием другого спина.

Все это уже давно было известно химикам. Тем не менее лишь в прошлом году Гершенфельд и Чанг и независимо Кори, Фами и Гавел указали, что ЯМР предоставляет полезную реализацию квантовых вычислений. Это не было очевидным по ряду причин. Наиболее важная: ЯМР-системы очень *горячие*. Типичная спиновая температура (скажем, комнатная температура) по порядку может быть в миллионы раз больше энергии расщепления между $|0\rangle$ и $|1\rangle$. Это означает, что квантовое состояние нашего компьютера (спинов в отдельной молекуле) существует на фоне очень интенсивного шума – оно испытывает очень сильные термические флуктуации. Этот шум будет искажать квантовую информацию. Более того, мы фактически выполняем нашу процедуру не на одной молекуле, а на макроскопическом образце, содержащем порядка 10^{23} «компьютеров», а считываемый нами сигнал в действительности усреднен по этому ансамблю. Но *вероятностный* характер квантовых алгоритмов обусловлен случайностью квантовых измерений. Следовательно, усреднение по ансамблю не эквивалентно выполнению вычислений на одном приборе; усреднение может скрыть результат.

Гершенфельд и Чанг, а также Кори, Фами и Гавел объяснили, как преодолеть эти трудности. Они описали, как можно готовить, управлять и контролировать «эффективно чистые состояния», выполняя соответствующие операции на термическом ансамбле. Идея состоит в том, чтобы обеспечить усреднение флуктуирующих свойств молекулы во время детектирования сигнала, так чтобы измерялись только интересующие нас когерентные свойства. Они также отметили, что некоторые квантовые алгоритмы (в том числе алгоритм факторизации Шора) можно разработать в детерминистской форме (так что, по крайней мере большая часть компьютеров будет давать один и тот же результат); тогда усреднение по многим вычислениям не нанесет ущерба результату.

Совсем недавно методы ЯМР были использованы для приготовления максимально запутанного состояния трех кубитов, чего не удавалось добиться раньше.

Очевидно, разработки квантового вычислительного «железа» находятся в младенческом состоянии. Необходимо на много порядков величины (как по количеству хранящихся кубитов, так и по количеству операций, которые могут быть применены) увеличить возможности существующего аппаратного обеспечения, прежде чем можно будет пытаться реализовать вычислительные амбиции. В случае метода ЯМР существует особенно се-

рьезное ограничение, которое возникает как принципиальный вопрос, поскольку отношение когерентного сигнала к фону экспоненциально спадает с ростом количества спинов, приходящихся на одну молекулу. На практике было бы очень заманчиво выполнить квантовое ЯМР-вычисление с более чем десятью кубитами. Возможно, для того чтобы квантовые компьютеры в конце концов стали реальными приборами, потребуются новые идеи в разработке квантового аппаратного обеспечения.

1.10. Резюме

Этим завершается наш вводный обзор квантовых вычислений. Мы увидели, что здесь соединились три фактора, сделавшие захватывающим этот предмет.

1) Квантовые компьютеры могут решать сложные задачи. Представим, что построена новая классификация сложности, лучше опирающаяся на фундаментальные законы физики, чем традиционная теория сложности. (Тогда остается более строго охарактеризовать класс задач, в которых квантовые компьютеры имеют большое преимущество перед классическими компьютерами.)

2) Квантовые ошибки можно корректировать. С помощью подходящих методов кодирования мы можем защитить сложную квантовую систему от разрушительного действия декогерентизации. Мы никогда не сможем увидеть настоящего полуживого-полумертвого кота, но, вероятно, сможем приготовить и сохранить в таком состоянии закодированного кота.

3) Квантовое аппаратное обеспечение можно сконструировать. У нас есть привилегия быть свидетелями начала эпохи когерентных манипуляций с квантовой информацией в лаборатории.

Целью этого курса будет углубление нашего понимания пунктов (1), (2) и (3).

Глава 2

Основы I: Состояния и ансамбли

2.1. Аксиомы квантовой механики

В предыдущих лекциях я говорил то одно, то другое о квантовом, хотя нигде не давал определения, что такое квантовая теория. Пришло время восполнить этот пробел.

Квантовая теория — это математическая модель физического мира. Чтобы охарактеризовать модель, необходимо определить, как она будет представлять состояния, наблюдаемые, измерения и динамику.

1. Состояния. Состояния представляют полное описание физической системы. В квантовой механике состояниями являются *лучи* в *гильбертовом пространстве*.

Что такое гильбертово пространство?

1) Это векторное пространство над полем комплексных чисел $\mathbb{C}$. В дальнейшем векторы будут обозначаться как $|\psi\rangle$ (кет-вектор Дирака).
2) В нем определено внутреннее произведение $\langle\varphi|\psi\rangle$. Оно представляет собой отображение упорядоченной пары векторов на $\mathbb{C}$, определяемое свойствами:
   а) положительность: $\langle\psi|\psi\rangle > 0$ для любого $|\psi\rangle \neq 0$;
   б) линейность: $\langle\varphi|(a|\psi_1\rangle + b|\psi_2\rangle) = a\langle\varphi|\psi_1\rangle + b\langle\varphi|\psi_2\rangle$;
   в) эрмитовская симметрия: $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$.
3) Оно полно по норме $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$.
   (В бесконечномерных функциональных пространствах полнота является важным условием, обеспечивающим сходимость разложений по определенным базисам собственных функций, например, рядов Фурье. Однако, как правило, нас вполне удовлетворит работа с внутренними произведениями в конечномерных пространствах.)

Что такое луч? Это класс эквивалентности векторов, отличающихся друг от друга ненулевым комплексным скалярным множителем. В качестве представителя такого класса (для любого ненулевого вектора) можно выбрать вектор, нормированный на единицу

$$\langle\psi|\psi\rangle = 1. \tag{2.1}$$

Будем также говорить, что $|\psi\rangle$ и $e^{i\alpha}|\psi\rangle$, где $|e^{i\alpha}| = 1$, описывают одно и то же физическое состояние.

[Заметим, что каждый луч соответствует возможному состоянию, то есть из двух данных состояний $|\varphi\rangle$ и $|\psi\rangle$ можно сформировать другое состояние $a|\varphi\rangle + b|\psi\rangle$ («принцип суперпозиции»). Физический смысл в этой суперпозиции имеет *относительная фаза*; мы отождествляем $a|\varphi\rangle + b|\psi\rangle$ с $e^{i\alpha}(a|\varphi\rangle + b|\psi\rangle)$, но отличаем его от $a|\varphi\rangle + e^{i\alpha}b|\psi\rangle$.]

2. Наблюдаемые. Наблюдаемой является свойство физической системы, которое в принципе может быть измерено. В квантовой механике наблюдаемая представляется *самосопряженным оператором*. Оператор определяет линейное отображение одного вектора в другой

$$\mathbf{A} : |\psi\rangle \to \mathbf{A}|\psi\rangle, \qquad \mathbf{A}(a|\psi\rangle + b|\varphi\rangle) = a\mathbf{A}|\psi\rangle + b\mathbf{B}|\varphi\rangle. \tag{2.2}$$

Оператор, сопряженный к $\mathbf{A}$, определяется соотношением

$$\langle\varphi|\mathbf{A}\psi\rangle = \langle\mathbf{A}^\dagger\varphi|\psi\rangle \tag{2.3}$$

для любой пары векторов $|\varphi\rangle$ и $|\psi\rangle$ (здесь я обозначил $\mathbf{A}|\psi\rangle$ символом $|\mathbf{A}\psi\rangle$). Оператор $\mathbf{A}$ самосопряжен, если $\mathbf{A} = \mathbf{A}^\dagger$.

Если $\mathbf{A}$ и $\mathbf{B}$ — самосопряженные операторы, то $\mathbf{A}+\mathbf{B}$ — тоже самосопряжен [поскольку $(\mathbf{A}+\mathbf{B})^\dagger = \mathbf{A}^\dagger+\mathbf{B}^\dagger$], но $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger\mathbf{A}^\dagger$ и поэтому $\mathbf{AB}$ — самосопряженный оператор только в том случае, когда $\mathbf{A}$ и $\mathbf{B}$ коммутируют. Заметим, что $\mathbf{AB}+\mathbf{BA}$ и $i(\mathbf{AB}-\mathbf{BA})$ всегда самосопряженные, если таковыми являются $\mathbf{A}$ и $\mathbf{B}$.

Для самосопряженного оператора в гильбертовом пространстве $\mathcal{H}$ существует спектральное представление — его собственные состояния образуют полный, ортонормированный базис в $\mathcal{H}$. Мы можем представить самосопряженный оператор $\mathbf{A}$ в виде

$$\mathbf{A} = \sum_n a_n \mathbf{P}_n. \tag{2.4}$$

Здесь каждое a_n — собственное значение оператора $\mathbf{A}$, а $\mathbf{P}_n$ — соответствующий ортогональный проектор (проекционный оператор) на пространство собственных векторов, отвечающих собственному значению a_n. (Если a_n не вырождено, то $\mathbf{P}_n = |n\rangle\langle n|$ — проектор на соответствующий собственный вектор.) Проекторы $\mathbf{P}_n$ обладают свойствами

$$\mathbf{P}_n\mathbf{P}_m = \delta_{n,m}\mathbf{P}_n, \qquad \mathbf{P}_n^\dagger = \mathbf{P}_n. \tag{2.5}$$

(Определение самосопряженности и формулировка спектральной теоремы для неограниченных операторов в бесконечномерном пространстве более тонкие, но здесь это не должно нас беспокоить.)

3. Измерение. В квантовой механике численным результатом измерения наблюдаемой $\mathbf{A}$ является одно из ее собственных значений; сразу после измерения квантовым состоянием является собственное состояние $\mathbf{A}$, соответствующее измеренному собственному значению a_n. Если непосредственно перед измерением квантовое состояние описывалось вектором $|\psi\rangle$, то результат a_n получается с *вероятностью*

$$\mathrm{Prob}(a_n) = \|\mathbf{P}_n|\psi\rangle\|^2 = \langle\psi|\mathbf{P}_n|\psi\rangle. \tag{2.6}$$

Если полученным результатом является a_n, то (нормированным) квантовым состоянием становится

$$\frac{\mathbf{P}_n|\psi\rangle}{\sqrt{\langle\psi|\mathbf{P}_n|\psi\rangle}}. \tag{2.7}$$

(Заметим, что если тотчас повторить измерение, тогда согласно этому правилу вновь, с вероятностью единица, будет получен тот же самый результат.)

4. Динамика. Эволюция во времени квантового состояния унитарна; она порождается самосопряженным оператором, называемым *гамильтонианом* системы. В *шредингеровской картине* динамики вектор, описывающий систему, изменяется во времени согласно *уравнению Шредингера*

$$\frac{d}{dt}|\psi(t)\rangle = -i\mathbf{H}|\psi(t)\rangle, \tag{2.8}$$

где $\mathbf{H}$ — гамильтониан. Мы можем переписать это уравнение в первом порядке по бесконечно малой величине dt:

$$|\psi(t+dt)\rangle = (\mathbf{1} - i\mathbf{H}dt)|\psi(t)\rangle. \tag{2.9}$$

Оператор $\mathbf{U}(dt) = \mathbf{1} - i\mathbf{H}dt$ унитарен, поскольку $\mathbf{H}$ самосопряжен; в линейном порядке по dt он удовлетворяет соотношению $\mathbf{U}^\dagger\mathbf{U} = 1$. Поскольку произведение унитарных операторов унитарно, эволюция в конечном интервале времени также унитарна:

$$|\psi(t)\rangle = \mathbf{U}(t)|\psi(0)\rangle. \tag{2.10}$$

Если гамильтониан $\mathbf{H}$ не зависит от времени, то можно записать $\mathbf{U}(t) = e^{-i\mathbf{H}t}$.

Этим завершается математическая формулировка квантовой механики. Мы непосредственно замечаем ее некоторые необычные черты. Одна ее странность состоит в том, что уравнение Шредингера линейно, тогда как мы привыкли к нелинейным динамическим уравнениям классической физики. Очевидно, это свойство требует объяснения. Однако гораздо более странным представляется таинственный дуализм; два совершенно разных способа изменения квантовых состояний. С одной стороны, существует детерминистская унитарная эволюция. Если мы точно определили $|\psi(0)\rangle$, то теория предсказывает состояние $|\psi(t)\rangle$ в любой более поздний момент времени.

С другой стороны, имеется вероятностное измерение. Теория не дает определенных предсказаний относительно результатов измерения; она лишь приписывает вероятности различным альтернативам. Это вызывает беспокойство, поскольку неясно, почему, в отличие от других процессов, измерение должно управляться иными физическими законами.

Начинающие изучать квантовую механику, впервые столкнувшись с этими правилами, редко спрашивают «почему?». В этом есть определенная мудрость. Но я надеюсь, что может быть полезно спросить: почему? В будущих лекциях мы вернемся к этому вызывающему замешательство дуализму между унитарной эволюцией и измерением и найдем его разрешение.

2.2. Кубит

Неделимой единицей классической информации является *бит*, принимающий одно из двух возможных значений $\{0, 1\}$. Соответствующую единицу квантовой информации называют «квантовый бит» или *кубит*. Он описывает состояние простейшей квантовой системы.

Минимальное нетривиальное гильбертово пространство двумерно. Будем обозначить ортогональный базис в двумерном векторном пространстве

как $\{|0\rangle, |1\rangle\}$. Тогда наиболее общее нормированное состояние может быть представлено в виде

$$a|0\rangle + b|1\rangle, \tag{2.11}$$

где a, b — комплексные числа, удовлетворяющие условию $|a|^2 + |b|^2 = 1$, а общая фаза физически несущественна. *Кубитом* является состояние в двумерном гильбертовом пространстве, которое может принимать любое значение, описываемое уравнением(2.11). Мы можем выполнить измерение, проецирующее кубит на базис $\{|0\rangle, |1\rangle\}$. Тогда с вероятностью $|a|^2$ мы получим результат $|0\rangle$, а с вероятностью $|b|^2$ — результат $|1\rangle$. Более того, за исключением случаев $a = 0$ и $b = 0$ измерение неизбежно ведет к возмущению состояния. Если начальное значение кубита неизвестно, тогда нет способа определить a и b с помощью одного такого или любого другого мыслимого измерения. Однако *после* измерения кубит оказывается в *известном* состоянии — $|0\rangle$ или $|1\rangle$ — отличающемся (вообще говоря) от его предыдущего состояния.

В этом отношении кубит отличается от классического бита; мы можем измерить классический бит, не возмущая его, и расшифровать всю закодированную в нем информацию. Допустим, мы имеем классический бит, который в действительности имеет определенное, но неизвестное нам значение (0 или 1). Опираясь на доступную информацию, мы можем только сказать, что с *вероятностью* p_0 бит имеет значение 0, а с вероятностью p_1 — значение 1, причем $p_0 + p_1 = 1$. Измеряя бит, мы получаем дополнительную информацию, позволяющую узнать его значение со 100% уверенностью.

Важный вопрос: в чем суть различия между кубитом и *вероятностным* классическим битом? По разным причинам, которые мы с вами изучим, это действительно не одно и то же.

2.2.1. Спин-1/2

Прежде всего заметим, что коэффициенты a и b в уравнении (2.11) содержат нечто большее, чем просто вероятности результатов измерения в базисе $\{|0\rangle, |1\rangle\}$. В частности, *относительная фаза* a и b также имеет физическое значение.

Для физика естественно интерпретировать уравнение (2.11) как спиновое состояние объекта со спином-1/2 (типа электрона). Тогда состояния $|0\rangle$ и $|1\rangle$ представляют собой состояния спин вверх $|\uparrow\rangle$ и спин вниз $|\downarrow\rangle$ вдоль некоторой оси, например, оси z. Два вещественных числа, характеризующих кубит (комплексные числа a и b, без учета их общей фазы и нормы), описывают ориентацию спина в трехмерном пространстве (полярный угол θ и азимутальный угол φ).

Мы не имеем возможности углубляться здесь в теорию симметрии в квантовой механике, напомним лишь кратко ее некоторые элементы, которые окажутся полезными в дальнейшем. Симметрия представляет собой преобразование, действие которого на состояние системы оставляет неизменными все наблюдаемые свойства системы. В квантовой механике наблюдениями являются измерения самосопряженных операторов. Если $\mathbf{A}$ измеряется в состоянии $|\psi\rangle$, то с вероятностью $|\langle a|\psi\rangle|^2$ будет получен результат $|a\rangle$ (собственный вектор оператора $\mathbf{A}$). Симметрия должна оставлять неизменными эти вероятности (если мы «поворачиваем» систему *вместе* с приборами).

Операция симметрии представляет собой отображение векторов в гильбертовом пространстве

$$|\psi\rangle \rightarrow |\psi'\rangle, \tag{2.12}$$

сохраняющее абсолютные значения внутренних произведений

$$|\langle\varphi|\psi\rangle| = |\langle\varphi'|\psi'\rangle| \tag{2.13}$$

для любых $|\varphi\rangle$ и $|\psi\rangle$. Согласно знаменитой теореме Вигнера, отображение с таким свойством всегда может быть выбрано (принимая соответствующее соглашение относительно фазы) унитарным или антиунитарным. Важная для дискретных симметрий антиунитарная альтернатива может быть исключена в случае непрерывных симметрий. Тогда преобразование симметрии действует как

$$|\psi\rangle \rightarrow |\psi'\rangle = \mathbf{U}|\psi\rangle, \tag{2.14}$$

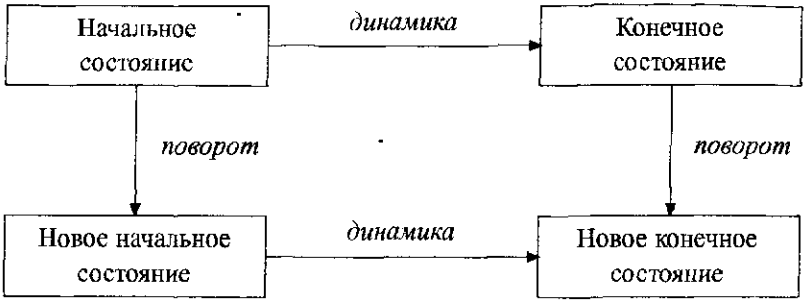
где $\mathbf{U}$ — унитарный оператор (и, в частности, линейный).

Симметрии образуют группу: преобразование симметрии можно обратить, а произведение двух симметрий, в свою очередь, является симметрией. Каждой операции симметрии R, действующей на нашу систему, соответствует унитарное преобразование $\mathbf{U}(R)$. Перемножение этих унитарных операторов должно соответствовать групповому закону перемножения симметрий — применение $R_1 \circ R_2$ должно быть эквивалентно последовательному применению сначала R_2, а затем R_1. Таким образом, мы требуем

$$\mathbf{U}(R_1)\mathbf{U}(R_2) = Phase(R_1, R_2)\mathbf{U}(R_1 \circ R_2). \tag{2.15}$$

В уравнении (2.15) допускается фазовый множитель, поскольку квантовыми состояниями являются *лучи*; нам нужно требовать лишь того, чтобы $\mathbf{U}(R_1 \circ R_2)$ действовал так же, как и $\mathbf{U}(R_1)\mathbf{U}(R_2)$, на лучи, а не на векторы. $\mathbf{U}(R)$ обеспечивает унитарное (с точностью до фазы) представление группы симметрии.

До сих пор наше понятие симметрии не имело связи с динамикой. Обычно мы требуем от симметрии, чтобы она сохраняла динамическую эволюцию системы. Это означает, что не должно иметь значения, преобразуем ли мы сначала систему, а затем она эволюционирует, или наоборот, сначала происходит эволюция системы, а затем мы преобразуем ее. Другими словами, диаграмма



коммутативна. Это означает, что оператор эволюции во времени должен коммутировать с преобразованиями симметрии $\mathbf{U}(R)$

$$\mathbf{U}(R)e^{-it\mathbf{H}} = e^{-it\mathbf{H}}\mathbf{U}(R). \tag{2.16}$$

Разлагая это уравнение до линейного порядка по t, получаем

$$\mathbf{U}(R)\mathbf{H} = \mathbf{H}\mathbf{U}(R). \tag{2.17}$$

В случае непрерывной симметрии операция R может быть выбрана сколь угодно близкой к единице $R = 1 + \epsilon T$, тогда $\mathbf{U}$ также близок к единичному оператору $\mathbf{1}$:

$$\mathbf{U}(R) = \mathbf{1} - i\epsilon\mathbf{Q} + O(\epsilon^2). \tag{2.18}$$

Из унитарности (в линейном порядке по ϵ) $\mathbf{U}$ следует, что $\mathbf{Q}$ является наблюдаемой $\mathbf{Q} = \mathbf{Q}^\dagger$. Разлагая уравнение (2.17) до линейных по ϵ слагаемых, находим

$$[\mathbf{Q}, \mathbf{H}] = 0\,; \tag{2.19}$$

наблюдаемая $\mathbf{Q}$ коммутирует с гамильтонианом.

Уравнение (2.19) представляет собой *закон сохранения*. Он говорит, например, что если мы приготовили собственное состояние оператора $\mathbf{Q}$, то управляемая уравнением Шредингера эволюция во времени будет сохранять это собственное состояние. Таким образом, симметрии влекут за

собой законы сохранения. И наоборот, по заданной сохраняющейся величине $\mathbf{Q}$, удовлетворяющей уравнению (2.19), можно построить соответствующее преобразование симметрии. Конечное преобразование может быть построено как произведение множества ифинитезимальных преобразований

$$R = \left(1 + \frac{\theta}{N} T\right)^N \Rightarrow \mathbf{U}(R) = \left(\mathbf{1} + i\frac{\theta}{N}\mathbf{Q}\right)^N \to e^{i\theta\mathbf{Q}} \tag{2.20}$$

(в пределе $N \to \infty$). Выяснив, как выглядит унитарное представление инфинитезимальных преобразований, мы тем самым определили представление конечных преобразований; они могут быть построены как произведения инфинитезимальных преобразований. Мы говорим, что $\mathbf{Q}$ – *генератор* симметрии.

Кратко напомним, как эта общая теория применяется к пространственным поворотам и моменту количества движения (импульса). Бесконечно малый поворот на угол $d\theta$ вокруг оси, определяемой единичным вектором $\vec{n} = (n_1, n_2, n_3)$, может быть представлен в виде

$$\mathbf{U}(\vec{n}, d\theta) = 1 - id\theta\, \vec{n} \cdot \vec{\mathbf{J}}, \tag{2.21}$$

где $(\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3)$ – компоненты момента импульса. Конечный поворот выражается как

$$\mathbf{U}(\vec{n}, \theta) = \exp(-i\theta\, \vec{n} \cdot \vec{\mathbf{J}}). \tag{2.22}$$

Повороты вокруг разных осей не коммутируют между собой. Из их элементарных свойств вытекают коммутационные соотношения

$$\left[\mathbf{J}_k, \mathbf{J}_l\right] = i\varepsilon_{klm}\mathbf{J}_m, \tag{2.23}$$

где ε_{klm} – полностью антисимметричный единичный псевдотензор ($\varepsilon_{123} = 1$), а по повторяющимся индексам предполагается суммирование. Чтобы совершать повороты квантовой системы, найдем удовлетворяющие коммутационным соотношениям (2.23) самосопряженные операторы $\mathbf{J}_1$, $\mathbf{J}_2, \mathbf{J}_3$ в гильбертовом пространстве.

«Определяющее» представление группы поворотов трехмерно, однако простейшее нетривиальное неприводимое представление является двумерным и задается генераторами

$$\mathbf{J}_k = \frac{1}{2}\boldsymbol{\sigma}_k, \tag{2.24}$$

где

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.25}$$

— матрицы Паули. С точностью до унитарного преобразования базиса это единственное двумерное неприводимое представление. Поскольку собственные числа операторов $\mathbf{J}_k$ равны $\pm 1/2$, мы называем его представлением спина-1/2. (Отождествляя $\mathbf{J}$ с моментом импульса, мы неявно выбрали единицы, в которых $\hbar = 1$). Матрицы Паули также обладают свойствами взаимной антикоммутации и идемпотентности

$$\boldsymbol{\sigma}_k \boldsymbol{\sigma}_l + \boldsymbol{\sigma}_l \boldsymbol{\sigma}_k = 2\delta_{kl}\mathbf{1}. \tag{2.26}$$

Таким образом, $(\vec{n}\cdot\vec{\boldsymbol{\sigma}})^2 = n_k n_l \boldsymbol{\sigma}_k \boldsymbol{\sigma}_l = n_k n_k \mathbf{1} = \mathbf{1}$. Разлагая экспоненту в ряд, мы видим, что конечный поворот представляется как

$$\mathbf{U}(\vec{n},\theta) = \exp\Big(-i\frac{\theta}{2}\vec{n}\cdot\vec{\boldsymbol{\sigma}}\Big) = \mathbf{1}\cos\frac{\theta}{2} - i(\vec{n}\cdot\vec{\boldsymbol{\sigma}})\sin\frac{\theta}{2}. \tag{2.27}$$

В такой форме можно представить наиболее общую унитарную 2×2-матрицу с единичным определителем. Это позволяет думать о кубите как о состоянии объекта со спином-1/2, а о произвольном унитарном преобразовании (кроме возможного поворота общей фазы), действующем на это состояние (кубит), — как о *повороте* спина.

Необычным свойством представления $\mathbf{U}(\vec{n},\theta)$ является его *двузначность*. В частности, нетривиально представляется поворот на угол 2π вокруг любой оси:

$$\mathbf{U}(\vec{n},\theta = 2\pi) = -\mathbf{1}. \tag{2.28}$$

Наше представление группы вращений в действительности является представлением «с точностью до знака»

$$\mathbf{U}(R_1)\mathbf{U}(R_2) = \pm\mathbf{U}(R_1 \circ R_2). \tag{2.29}$$

Но, как уже отмечалось, это вполне приемлемо, *поскольку групповое умножение относится к лучам*, а не к векторам. Эти двузначные представления группы вращений называются *спинорными* представлениями. (Существование спиноров следует из топологического свойства группы — она не односвязна.)

Несмотря на то, что поворот на угол 2π действительно не ведет к наблюдаемому изменению состояния объекта со спином-1/2, было бы ошибкой считать, что спинорное свойство не имеет наблюдаемых следствий.

Допустим, у меня есть машина, которая действует на пару спинов. Если первый спин направлен вверх, то она ничего не меняет, но если первый спин направлен вниз, она поворачивает второй спин на угол 2π. Пусть теперь эта машина действует на состояние, в котором первый спин находится в суперпозиции состояний «вверх» и «вниз». Тогда

$$\frac{1}{\sqrt{2}}(|\uparrow\rangle_1 + |\downarrow\rangle_1)|\uparrow\rangle_2 \rightarrow \frac{1}{\sqrt{2}}(|\uparrow\rangle_1 - |\downarrow\rangle_1)|\uparrow\rangle_2. \tag{2.30}$$

Несмотря на отсутствие наблюдаемого влияния на состояние второго спина, состояние первого спина стало ортогональным его исходному состоянию, что очень даже наблюдаемо.

В повернутой системе отсчета поворот $R(\vec{n},\theta)$ превращается в поворот на тот же самый угол, но вокруг повернутой оси. Отсюда следует, что три компоненты момента количества движения при поворотах преобразуются как вектор

$$\mathbf{U}(R)\mathbf{J}_k\mathbf{U}^\dagger(R) = R_{kl}\mathbf{J}_l. \tag{2.31}$$

Таким образом, если состояние $|m\rangle$ является собственным состоянием оператора $\mathbf{J}_3$

$$\mathbf{J}_3|m\rangle = m|m\rangle, \tag{2.32}$$

тогда $\mathbf{U}(R)|m\rangle$ является собственным состоянием оператора $R\mathbf{J}_3$ с тем же самым собственным значением:

$$\begin{aligned} R\mathbf{J}_3\mathbf{U}(R)|m\rangle &= \mathbf{U}(R)\mathbf{J}_3\mathbf{U}^\dagger(R)\mathbf{U}(R)|m\rangle \\ &= \mathbf{U}(R)\mathbf{J}_3|m\rangle = m(\mathbf{U}(R)|m\rangle). \end{aligned} \tag{2.33}$$

Следовательно, мы можем построить собственные состояния оператора проекции момента импульса на ось $\vec{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, применяя поворот на угол θ вокруг оси $\vec{n}' = (-\sin\varphi, \cos\varphi, 0)$ к собственному состоянию оператора $\mathbf{J}_3$. Для нашего представления спина-1/2 таким поворотом является

$$\begin{aligned} \exp\left[-i\frac{\theta}{2}\vec{n}'\cdot\vec{\boldsymbol{\sigma}}\right] &= \exp\left[\frac{\theta}{2}\begin{pmatrix} 0 & -e^{-i\varphi} \\ e^{i\varphi} & 0 \end{pmatrix}\right] = \\ &= \begin{pmatrix} \cos\frac{\theta}{2} & -e^{-i\varphi}\sin\frac{\theta}{2} \\ +e^{+i\varphi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}. \end{aligned} \tag{2.34}$$

Применяя его к $\binom{1}{0}$ — собственному состоянию оператора $\mathbf{J}_3$ с собственным значением $+1$, получим (с точностью до общей фазы)

$$|\psi(\theta,\varphi)\rangle = \begin{pmatrix} e^{-i\varphi/2}\cos\dfrac{\theta}{2} \\ e^{+i\varphi/2}\sin\dfrac{\theta}{2} \end{pmatrix}. \tag{2.35}$$

Мы можем непосредственно проверить, что этот вектор является собственным состоянием оператора

$$\vec{n}\cdot\vec{\boldsymbol{\sigma}} = \begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta \end{pmatrix} \tag{2.36}$$

с собственным значением $+1$. Итак, мы видим, что уравнение (2.11) с $a = e^{-i\varphi/2}\cos\dfrac{\theta}{2}$, $b = e^{+i\varphi/2}\sin\dfrac{\theta}{2}$ может интерпретироваться как состояние спина, ориентированного вдоль направления (θ,φ).

Мы уже отмечали, что невозможно определить a и b с помощью одного измерения. Более того, даже имея множество идентичных копий данного состояния, мы не можем полностью его определить, измеряя каждую копию только вдоль оси z. Это могло бы позволить нам оценить $|a|$ и $|b|$, но не дало бы возможности получить информацию об относительной фазе a и b. Эквивалентно мы могли бы найти значение проекции спина на ось z

$$\langle\psi(\theta,\varphi)|\sigma_3|\psi(\theta,\varphi)\rangle = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} = \cos\theta, \tag{2.37}$$

но ничего не смогли бы узнать о его компонентах в плоскости xy. Проблема определения $|\psi\rangle$ путем измерения спина аналогична проблеме определения единичного вектора $\vec{n}$ путем измерения его компонент вдоль различных осей. В общем необходимы измерения вдоль трех различных осей. Например, из $\langle\sigma_3\rangle$ и $\langle\sigma_1\rangle$ мы можем определить n_3 и n_1, но знак n_2 останется неопределенным. Измерение $\langle\sigma_2\rangle$ могло бы устранить эту двусмысленность.

Конечно, если мы позволили поворачивать спин, тогда будет достаточно измерений только вдоль оси z. То есть измерение спина вдоль оси $\vec{n}$ эквивалентно предварительному совершению поворота, совмещающего ось $\vec{n}$ с осью z, а затем — измерению вдоль оси z.

В частном случае $\theta = \pi/2$, $\varphi = 0$ (ось x) наше спиновое состояние имеет вид

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}}\bigl(|\uparrow_z\rangle + |\downarrow_z\rangle\bigr), \tag{2.38}$$

(«спин вверх вдоль оси x»). Ортогональным ему состоянием («спин вниз вдоль оси x») является

$$|\downarrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle - |\downarrow_z\rangle). \tag{2.39}$$

Для этих обоих состояний, измеряя спин вдоль оси z, мы получим $|\uparrow_z\rangle$ с вероятностью 1/2 и $|\downarrow_z\rangle$ с вероятностью 1/2.

Рассмотрим теперь комбинацию

$$\frac{1}{\sqrt{2}}(|\uparrow_x\rangle + |\downarrow_x\rangle). \tag{2.40}$$

Это состояние обладает тем свойством, что если мы измеряем спин вдоль оси x, то с вероятностями, равными 1/2, получаем $|\uparrow_x\rangle$ или $|\downarrow_x\rangle$. Можно спросить, что мы получим, измеряя состояние (2.40) вдоль оси z?

Если бы это были классические вероятностные биты, то ответ был бы очевиден. Состояние (2.40) представляет собой одно из двух состояний, и для *каждого* из них вероятность направления вверх или вниз вдоль оси z равна 1/2. Таким образом, измеряя состояние (2.40) вдоль оси z, мы, конечно, должны с вероятностью 1/2 обнаружить спин, направленным вверх.

Но для кубитов это не так! Складывая уравнения (2.38) и (2.39), мы обнаруживаем, что в состоянии, описываемом уравнением (2.40), в действительности замаскировано состояние $|\uparrow_z\rangle$. Измеряя его вдоль оси z, мы всегда будем получать $|\uparrow_z\rangle$ и никогда — $|\downarrow_z\rangle$.

Таким образом, для кубитов, в противоположность классическим вероятностным битам, вероятности могут складываться довольно неожиданным образом. В этом и состоит (в его простейшей форме) явление, называемое «квантовой интерференцией», важная особенность квантовой информации.

Следует подчеркнуть, что, хотя *формальная* эквивалентность объекту со спином-1/2 применима к любой двухуровневой квантовой системе, конечно, не каждая двухуровневая система преобразуется при поворотах как спинор.

2.2.2. Поляризации фотона

Другую важную систему, имеющую два состояния, представляет *фотон*, который может иметь одну из двух независимых поляризаций. Состояния поляризации фотона тоже преобразуются при поворотах, однако фотоны отличаются от объектов со спином-1/2 в двух важных отношениях:

(1) Фотоны являются безмассовыми частицами. (2) Фотоны имеют спин-1 (это не спинорные частицы).

Мы не располагаем временем для детального обсуждения унитарных представлений группы Пуанкаре. Достаточно сказать, что спин частицы определяет, как преобразуется ее состояние под действием преобразований *малой* группы — сохраняющей импульс частицы подгруппы группы Лоренца. В случае массивной частицы мы всегда можем перейти в ее систему покоя и тогда малой группой является группа вращений.

Для безмассовых частиц не существует системы покоя. Конечномерные унитарные представления малой группы превращаются в представления группы вращений в *двумерном* пространстве, вращений вокруг направления импульса. Конечно, в случае с фотоном это соответствует знакомому свойству классического света – волны поляризованы перпендикулярно направлению распространения.

При повороте вокруг направления распространения два состояния линейной поляризации ($|x\rangle$ и $|y\rangle$ для горизонтальной и вертикальной поляризации) преобразуются как

$$\begin{aligned} |x\rangle &\to +\cos\theta\,|x\rangle + \sin\theta\,|y\rangle, \\ |y\rangle &\to -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle. \end{aligned} \tag{2.41}$$

Это двумерное представление в действительности приводимо. Матрица

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{2.42}$$

имеет собственные состояния

$$|R\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}, \qquad |L\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}, \tag{2.43}$$

отвечающие собственным значениям $e^{+i\theta}$ и $e^{-i\theta}$, состояния правой и левой циркулярной поляризации. То есть они являются собственными состояниями генератора поворотов

$$\mathrm{J} = \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix} = \boldsymbol{\sigma}_y \tag{2.44}$$

с собственными значениями ± 1. Поскольку собственные значения равны ± 1 (а не $\pm 1/2$), мы говорим, что фотон имеет спин-1.

В этом контексте явление квантовой интерференции может быть описано следующим образом. Предположим, что мы имеем анализатор поляризации, который позволяет пройти через него фотону только с одной из двух линейных поляризаций. Тогда вероятность прохождения x- или y-поляризованного фотона через повернутый на 45° анализатор равна 1/2; 1/2-ой равна и вероятность прохождения поляризованного под углом 45° фотона через x- или y-анализатор. Однако x-фотон *никогда* не пройдет через y-анализатор. Если мы поместим повернутый на 45° анализатор между x- и y-анализаторами, тогда через каждый анализатор пройдет половина падающих на него фотонов. Но если мы удалим промежуточный анализатор, то *ни один* фотон не пройдет через y-анализатор.

Очевидно, можно сконструировать прибор, который поворачивает плоскость поляризации фотона и, следовательно, применяет преобразование (2.41) к нашему кубиту. Как уже отмечалось, это не самое общее унитарное преобразование. Но если мы имеем также и прибор, который изменяет относительную фазу двух ортогональных, линейно поляризованных состояний

$$\begin{aligned} |x\rangle &\to e^{+i\omega/2}|x\rangle, \\ |y\rangle &\to e^{-i\omega/2}|y\rangle, \end{aligned} \tag{2.45}$$

то эти два прибора можно использовать вместе, чтобы совершить произвольное 2×2 унитарное преобразование (с определителем, равным единице) состояния поляризации фотона.

2.3. Матрица плотности

2.3.1. Бинарная квантовая система

Последняя лекция была об одном кубите. Эта лекция — о *двух* кубитах. (Догадайтесь, о чем будет следующая лекция!) Переход от одного кубита к двум — более серьезный шаг, чем вы могли бы ожидать. Многое из того, что есть странного и чудесного в квантовой механике, можно понять, рассматривая свойства квантовых состояний двух кубитов.

Аксиомы § 2.1 дают вполне приемлемую общую формулировку квантовой теории. Тем не менее при многих обстоятельствах мы обнаруживаем, что они кажутся нарушенными. Беда в том, что наши аксиомы нацелены на то, чтобы характеризовать поведение всей Вселенной. Но, как правило, у нас нет таких амбиций, как пытаться понять физику всей Вселенной; мы довольствуемся изучением только нашего маленького уголка. На практи-

ке наши исследования всегда ограничены малой частью гораздо большей квантовой системы.

В следующих нескольких лекциях мы увидим, что если мы ограничиваем наше внимание только на части большой системы, то (в противоположность аксиомам § 2.1):

1) состояния *не* являются лучами.

2) измерения *не* являются ортогональными проекторами.

3) эволюция *не* унитарна.

Мы сможем лучше понять эти моменты, рассматривая простейший пример: мир двух кубитов, в котором мы наблюдаем только один из них.

Итак, рассмотрим систему двух кубитов. Кубит A находится здесь, в комнате вместе с нами, и мы вольны наблюдать его или манипулировать им по своему усмотрению. Но кубит B заперт в подвале, где мы не можем до него добраться. Имея некоторое состояние двух кубитов, мы хотели бы найти простой способ описания наблюдений, которые мы можем делать только на кубите A.

Будем использовать $\{|0\rangle_A, |1\rangle_A\}$ и $\{|0\rangle_B, |1\rangle_B\}$ для обозначения ортонормированных базисов для кубитов A и B соответственно. Рассмотрим следующее квантовое состояние двухкубитовой Вселенной:

$$|\psi\rangle_{AB} = a|0\rangle_A \otimes |0\rangle_B + b|1\rangle_A \otimes |1\rangle_B. \tag{2.46}$$

В этом состоянии кубиты A и B *коррелированы*. Допустим, мы измеряем кубит A, проецируя его на базис $\{|0\rangle_A, |1\rangle_A\}$. Тогда с вероятностью $|a|^2$ мы получим результат $|0\rangle_A$, а измерение приготовит состояние

$$|0\rangle_A \otimes |0\rangle_B; \tag{2.47}$$

с вероятностью $|b|^2$ мы получим результат $|1\rangle_A$, а измерение приготовит состояние

$$|1\rangle_A \otimes |1\rangle_B. \tag{2.48}$$

В каждом случае, в результате измерения выбирается определенное состояние кубита B. Если мы сразу за этим измерим кубит B, то наверняка (с вероятностью единица) обнаружим его в состоянии $|0\rangle_B$, если перед этим было получено $|0\rangle_A$, и — в состоянии $|1\rangle_B$, если предыдущим результатом было $|1\rangle_A$. В этом смысле результаты измерений $\{|0\rangle_A, |1\rangle_A\}$ и $\{|0\rangle_B, |1\rangle_B\}$ полностью скоррелированы в состоянии $|\psi\rangle_{AB}$.

Теперь я хотел бы рассмотреть более общие наблюдаемые, действующие на кубит A, и я хотел бы характеризовать результаты измерения только A (независимо от результатов любых измерений недоступного кубита B). Наблюдаемую, действующую только на кубит A, можно представить в виде

$$\mathbf{M}_A \otimes \mathbf{1}_B, \tag{2.49}$$

где $\mathbf{M}_A$ – действующий на A самосопряженный оператор, а $\mathbf{1}_B$ – действующий на B единичный оператор. Ожидаемое значение наблюдаемой в состоянии $|\psi\rangle_{AB}$ равно

$$\begin{aligned} {}_{AB}\langle\psi|\mathbf{M}_A \otimes \mathbf{1}_B|\psi\rangle_{AB} &= (a^*{}_A\langle 0| \otimes {}_B\langle 0| + b^*{}_A\langle 1| \otimes {}_B\langle 1|)\mathbf{M}_A \otimes \mathbf{1}_B \\ &\quad (a|0\rangle_A \otimes |0\rangle_B + b|1\rangle_A \otimes |1\rangle_B) = \\ &= |a|^2{}_A\langle 0|\mathbf{M}_A|0\rangle_A + |b|^2{}_A\langle 1|\mathbf{M}_A|1\rangle_A \end{aligned} \tag{2.50}$$

(где мы воспользовались ортогональностью $|0\rangle_B$ и $|1\rangle_B$). Это выражение можно переписать в форме

$$\langle \mathbf{M}_A \rangle = \mathrm{tr}(\mathbf{M}_A \boldsymbol{\rho}_A), \tag{2.51}$$

$$\boldsymbol{\rho}_A = |a|^2|0\rangle_A\,{}_A\langle 0| + |b|^2|1\rangle_A\,{}_A\langle 1|, \tag{2.52}$$

а tr обозначает *след*. Оператор $\boldsymbol{\rho}_A$ называется *оператором плотности* (или *матрицей плотности*) кубита A. Он самосопряжен, положителен (его собственные значения неотрицательны) и имеет единичный след (поскольку $|\psi\rangle$ являются нормированными состояниями).

Поскольку $\langle \mathbf{M}_A \rangle$ имеет вид (2.51) для любой наблюдаемой $\mathbf{M}_A$, действующей на кубит A, то логично интерпретировать $\boldsymbol{\rho}_A$ как *ансамбль* возможных квантовых состояний, каждое из которых возникает с определенной вероятностью. Другими словами, мы получили бы для $\langle \mathbf{M}_A \rangle$ тот же самый результат, если бы предположили, что кубит A находится в одном из двух квантовых состояний. Причем с вероятностью $p_0 = |a|^2$ он находится в состоянии $|0\rangle_A$, а с вероятностью $p_1 = |b|^2$ – в состоянии $|1\rangle_A$. Если нас интересует результат любого возможного измерения, мы можем рассматривать в качестве $\mathbf{M}_A$ проектор $\mathbf{E}_A(a)$ на соответствующее собственное пространство наблюдаемой. Тогда

$$\mathrm{Prob}(a) = p_0\,{}_A\langle 0|\mathbf{E}_A(a)|0\rangle_A + p_1\,{}_A\langle 1|\mathbf{E}_A(a)|1\rangle_A, \tag{2.53}$$

что представляет собой вероятность результата a, взвешенную вероятностью каждого квантового состояния и просуммированную по всему ансамблю.

Мы уже обращали внимание на существенное различие между когерентной суперпозицией состояний $|0\rangle_A$ и $|1\rangle_A$ и вероятностным ансамблем, в котором каждое из состояний $|0\rangle_A$ и $|1\rangle_A$ может появляться с конкретной вероятностью. Например, для объекта со спином-1/2 мы видели, что если мы измеряем σ_1 в состоянии $\frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle)$, то с единичной вероятностью получим результат $|\uparrow_x\rangle$. Но ансамбль, в котором каждое из состояний $|\uparrow_z\rangle$ и $|\downarrow_z\rangle$ появляется с вероятностью 1/2, представляется оператором плотности

$$\boldsymbol{\rho} = \frac{1}{2}(|\uparrow_z\rangle\langle\uparrow_z| + |\downarrow_z\rangle\langle\downarrow_z|) = \frac{1}{2}\mathbf{1}, \tag{2.54}$$

тогда проекция на $|\uparrow_x\rangle$ имеет ожидаемое значение

$$\mathrm{tr}(|\uparrow_x\rangle\langle\uparrow_x|\boldsymbol{\rho}) = \frac{1}{2}. \tag{2.55}$$

Фактически мы видели, что любое представляемое лучом состояние одного кубита можно интерпретировать как спин, ориентированный вдоль некоторого определенного направления. Но, поскольку левая часть (2.55) не изменяется при унитарном преобразовании базиса, а состояние $|\psi(\theta,\varphi)\rangle$ можно получить унитарным преобразованием состояния $|\uparrow_z\rangle$, мы видим, что для $\boldsymbol{\rho}$, определяемого уравнением (2.54),

$$\mathrm{tr}(|\psi(\theta,\varphi)\rangle\langle\psi(\theta,\varphi)|\boldsymbol{\rho}) = \frac{1}{2}. \tag{2.56}$$

Следовательно, если приготовлено состояние $|\psi\rangle_{AB}$ (2.46) с $|a|^2 = |b|^2 = 1/2$, то при измерении спина A вдоль любой оси мы получим совершенно случайный результат; каждая из ориентаций спина, вверх или вниз, может появиться с вероятностью 1/2.

Это обсуждение коррелированного двухкубитового состояния $|\psi\rangle_{AB}$ очевидным образом распространяется на произвольное состояние любой бинарной квантовой системы (системы, состоящей из двух частей). Гильбертовым пространством бинарной системы является $\mathbf{H}_A \otimes \mathbf{H}_B$, где $\mathbf{H}_{A,B}$ — гильбертово пространство одной из составляющих систему частей. Это означает, что если $\{|i\rangle_A\}$ — ортонормированный базис в $\mathbf{H}_A$, а $\{|\mu\rangle_B\}$ — ортонормированный базис в $\mathbf{H}_B$, то $\{|i\rangle_A \otimes |\mu\rangle_B\}$ — ортонормированный базис в $\mathbf{H}_A \otimes \mathbf{H}_B$. Таким образом, произвольное чистое состояние в $\mathbf{H}_A \otimes \mathbf{H}_B$ может быть представлено в виде разложения

$$|\psi\rangle_{AB} = \sum_{i,\mu} a_{i,\mu}|i\rangle_A \otimes |\mu\rangle_B, \tag{2.57}$$

где $\sum\limits_{i,\mu} |a_{i,\mu}|^2 = 1$. Ожидаемое значение наблюдаемой $\mathbf{M}_A \otimes \mathbf{1}_B$, действующей только на подсистему A, равно

$$\begin{aligned}\langle \mathbf{M}_A \rangle &= {}_{AB}\langle\psi|\mathbf{M}_A \otimes \mathbf{1}_B|\psi\rangle_{AB} = \\ &= \sum_{j,\nu} a^*_{j,\nu}\, {}_A\langle j| \otimes {}_B\langle\nu|\mathbf{M}_A \otimes \mathbf{1}_B \sum_{i,\mu} a_{i,\mu}|i\rangle_A \otimes |\mu\rangle_B = \\ &= \sum_{i,j,\mu} a^*_{j,\mu} a_{i,\mu}\, {}_A\langle j|\mathbf{M}_A|i\rangle_A = \\ &= \mathrm{tr}(\mathbf{M}_A \boldsymbol{\rho}_A), \end{aligned} \tag{2.58}$$

где

$$\boldsymbol{\rho}_A = \mathrm{tr}_B(|\psi\rangle_{AB}\, {}_{AB}\langle\psi|) \equiv \sum_{i,j,\mu} a_{i,\mu} a^*_{j,\mu} |i\rangle_A\, {}_A\langle j|. \tag{2.59}$$

Мы говорим, что оператор плотности $\boldsymbol{\rho}_A$ подсистемы A получается взятием частичного *следа* от матрицы плотности (в рассматриваемом случае от чистого состояния) составной системы AB по переменным подсистемы B.

Из определяющего уравнения (2.59) непосредственно следует, что $\boldsymbol{\rho}_A$ обладает следующими свойствами:

1) $\boldsymbol{\rho}_A$ — самосопряженый оператор: $\boldsymbol{\rho}_A = \boldsymbol{\rho}_A^\dagger$;

2) $\boldsymbol{\rho}_A$ — положительный оператор: для любого $|\psi\rangle_A$ ${}_A\langle\psi|\boldsymbol{\rho}_A|\psi\rangle_A = \sum\limits_{\mu} |\sum\limits_{i} a_{i,\mu}\, {}_A\langle\psi|i\rangle_A|^2 \geqslant 0$;

3) $\mathrm{tr}\,\boldsymbol{\rho}_A = 1$: мы имеем $\mathrm{tr}\,\boldsymbol{\rho}_A = \sum\limits_{i,\mu} |a_{i,\mu}|^2 = 1$, поскольку состояние $|\psi\rangle_{AB}$ нормировано.

Отсюда следует, что $\boldsymbol{\rho}_A$ может быть диагонализован, что все его собственные значения вещественны и неотрицательны и что сумма его собственных значений равна единице.

Если мы наблюдаем подсистему большей системы, то, даже если состоянием большей системы является луч, состояние подсистемы таковым не будет; в общем случае оно описывается оператором плотности. В том случае, когда состоянием подсистемы является луч, мы называем его *чистым*. В противном случае — состояние является *смешанным*.

Если состоянием является чистое состояние $|\psi\rangle_A$, то матрица плотности $\rho_A = |\psi\rangle_{A\,A}\langle\psi|$ представляет собой *проектор* на одномерное пространство, натянутое на $|\psi\rangle_A$. Следовательно, матрица плотности чистого состояния обладает свойством $\rho_A^2 = \rho_A$. В общем случае матрица плотности, разложенная в базисе, в котором она диагональна, имеет вид

$$\rho_A = \sum_a p_a |\psi_a\rangle\langle\psi_a|, \tag{2.60}$$

где $0 \leqslant p_a \leqslant 1$ и $\sum_a p_a = 1$. Если состояние не является чистым, то эта сумма состоит из двух или большего числа слагаемых и $\rho_A^2 \neq \rho_A$; фактически $\operatorname{tr}\rho_A^2 = \sum p_a^2 < \sum p_a = 1$. Мы говорим, что ρ представляет *некогерентную* суперпозицию состояний $\{|\psi_a\rangle\}$; некогерентность означает, что относительные фазы $|\psi_a\rangle$ экспериментально ненаблюдаемы.

Поскольку ожидаемое значение наблюдаемой M, действующей на подсистему, может быть представлено в виде

$$\langle \mathrm{M}\rangle = \operatorname{tr}\mathrm{M}\rho = \sum_a p_a \langle\psi_a|\mathrm{M}|\psi_a\rangle, \tag{2.61}$$

мы, как и прежде, видим, что ρ можно интерпретировать как *ансамбль* чистых квантовых состояний, в котором состояние $|\psi_a\rangle$ появляется с вероятностью p_a. Таким образом, мы прошли большую часть пути к пониманию того, как в квантовой механике возникают вероятности, когда квантовая система A взаимодействует с другой системой B. Состояния A и B становятся *запутанными*, то есть коррелированными. Запутывание *разрушает когерентность* суперпозиции состояний, так что некоторые фазы становятся ненаблюдаемыми, если мы следим за одной только A. Мы можем описывать эту ситуацию, говоря, что происходит *коллапс* (*редукция*) состояния системы A — она находится в одном из множества альтернативных состояний, каждому из которых может быть приписана вероятность.

2.3.2. Сфера Блоха

Вернемся к случаю, в котором системой A является один кубит, и рассмотрим общую форму матрицы плотности. Наиболее общая самосопряженная 2×2-матрица зависит от четырех вещественных параметров и может быть разложена в базисе $\{1, \sigma_1, \sigma_2, \sigma_3\}$. Поскольку каждая из матриц σ_i является бесследовой, коэффициент перед 1 в разложении матрицы

плотности $\boldsymbol{\rho}$ должен быть равен $1/2$ (так чтобы $\operatorname{tr}\boldsymbol{\rho} = 1$), а $\boldsymbol{\rho}$ может быть разложена как

$$\begin{aligned}\boldsymbol{\rho}(\vec{P}) &= \frac{1}{2}(\mathbf{1} + \vec{P}\cdot\vec{\boldsymbol{\sigma}}) \equiv \\ &\equiv \frac{1}{2}(\mathbf{1} + P_1\boldsymbol{\sigma}_1 + P_2\boldsymbol{\sigma}_2 + P_3\boldsymbol{\sigma}_3) = \\ &= \frac{1}{2}\begin{pmatrix} 1+P_3 & P_1 - iP_2 \\ P_1 + iP_2 & 1 - P_3 \end{pmatrix}. \end{aligned} \tag{2.62}$$

Мы можем вычислить $\det\boldsymbol{\rho} = \frac{1}{4}(1 - \vec{P}^2)$. Следовательно, необходимым условием неотрицательности собственных значений $\boldsymbol{\rho}$ является неравенство $\det\boldsymbol{\rho} \geqslant 0$ или $\vec{P}^2 \leqslant 1$. Это условие также и достаточно; поскольку $\operatorname{tr}\boldsymbol{\rho} = 1$ и $\boldsymbol{\rho}$ не может иметь два отрицательных собственных значения. Таким образом, имеется взаимно однозначное соответствие между возможными матрицами плотности одиночного кубита и точками *единичного трехмерного шара* $0 \leqslant |\vec{P}| \leqslant 1$. Этот шар обычно называют *сферой Блоха* (хотя, конечно, это в действительности шар, а не сфера).

Граница ($|\vec{P}| = 1$) шара (которая как раз и является сферой) содержит матрицы плотности с нулевым детерминантом. Поскольку $\operatorname{tr}\boldsymbol{\rho} = 1$, эти матрицы плотности должны иметь собственные значения 0 и 1. Они представляют собой одномерные проекторы и, следовательно, чистые состояния. Мы уже видели, что каждое чистое состояние одного кубита имеет вид $|\psi(\theta,\varphi)\rangle$ и может рассматриваться как ориентация спина в (θ,φ)-направлении. Действительно, используя свойство

$$(\hat{n}\cdot\vec{\boldsymbol{\sigma}})^2 = 1, \tag{2.63}$$

где $\hat{n}$ — единичный вектор, мы можем легко убедиться в том, что матрица плотности чистого состояния

$$\boldsymbol{\rho}(\hat{n}) = \frac{1}{2}(\mathbf{1} + \hat{n}\cdot\vec{\boldsymbol{\sigma}}) \tag{2.64}$$

обладает свойством

$$(\hat{n}\cdot\vec{\boldsymbol{\sigma}})\boldsymbol{\rho}(\hat{n}) = \boldsymbol{\rho}(\hat{n})(\hat{n}\cdot\vec{\boldsymbol{\sigma}}) = \boldsymbol{\rho}(\hat{n}) \tag{2.65}$$

и, следовательно, является проектором

$$\boldsymbol{\rho}(\hat{n}) = |\psi(\hat{n})\rangle\langle\psi(\hat{n})|\,; \tag{2.66}$$

то есть $\hat{n}$ представляет собой направление, вдоль которого ориентируется спин. И наоборот, из выражения

$$|\psi(\theta,\varphi)\rangle = \begin{pmatrix} e^{-i\varphi/2}\cos\dfrac{\theta}{2} \\ e^{+i\varphi/2}\sin\dfrac{\theta}{2} \end{pmatrix} \tag{2.67}$$

можно непосредственно найти, что

$$\begin{aligned} \boldsymbol{\rho}(\theta,\varphi) &= |\psi(\theta,\varphi)\rangle\langle\psi(\theta,\varphi)| = \\ &= \begin{pmatrix} \cos^2\dfrac{\theta}{2} & \cos\dfrac{\theta}{2}\sin\dfrac{\theta}{2}e^{-i\varphi} \\ \cos\dfrac{\theta}{2}\sin\dfrac{\theta}{2}e^{+i\varphi} & \sin^2\dfrac{\theta}{2} \end{pmatrix} = \\ &= \frac{1}{2}\mathbf{1}+\frac{1}{2}\begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{+i\varphi}\sin\theta & -\cos\theta \end{pmatrix} = \frac{1}{2}(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}), \end{aligned} \tag{2.68}$$

где $\hat{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. Приятным свойством блоховской параметризации чистых состояний является то, что, хотя $|\psi(\theta,\varphi)\rangle$ имеет физически несущественную произвольную общую фазу, в матрице плотности $\boldsymbol{\rho}(\theta,\varphi) = |\psi(\theta,\varphi)\rangle\langle\psi(\theta,\varphi)|$ этой неоднозначности нет; все параметры в $\boldsymbol{\rho}$ имеют физический смысл.

Из свойства

$$\frac{1}{2}\operatorname{tr}\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j = \delta_{ij} \tag{2.69}$$

следует, что

$$(\hat{n}\cdot\vec{\boldsymbol{\sigma}})_{\vec{P}} = \operatorname{tr}\left(\hat{n}\cdot\vec{\boldsymbol{\sigma}}\boldsymbol{\rho}(\vec{P})\right) = \hat{n}\cdot\vec{P}. \tag{2.70}$$

Таким образом, вектор $\vec{P}$ в уравнении (2.62) параметризует *поляризацию* спина. Если в нашем распоряжении имеется множество идентично приготовленных систем, мы можем определить $\vec{P}$ [и, следовательно, полностью матрицу плотности $\boldsymbol{\rho}(\vec{P})$], измеряя $(\hat{n}\cdot\vec{\boldsymbol{\sigma}})$ вдоль каждой из трех линейно независимых осей.

2.3.3. Теорема Глизона

Отправляясь от аксиом квантовой механики и рассматривая описание части большей квантовой системы, мы пришли к понятию матрицы плотности $\boldsymbol{\rho}$ и выражению $\operatorname{tr}(\mathbf{M}\boldsymbol{\rho})$ для ожидаемого значения наблюдаемой $\mathbf{M}$. Тем более приятно узнать, что формализм матрицы плотности является очень

общим и применим в гораздо более широких пределах. В этом состоит содержание *теоремы Глизона* (1957).

Теорема Глизона исходит из предпосылки, что задачей квантовой теории является сопоставление соответствующих вероятностей всем возможным ортогональным проекциям в гильбертовом пространстве (другими словами, всем возможным измерениям наблюдаемых).

Тогда состоянием квантовой системы является отображение, которое каждой проекции ($\mathbf{E}^2 = \mathbf{E}$ и $\mathbf{E} = \mathbf{E}^\dagger$) ставит в соответствие неотрицательное вещественное число, не превосходящее единицу:

$$\mathbf{E} \to p(\mathbf{E}), \qquad 0 \leqslant p(\mathbf{E}) \leqslant 1. \tag{2.71}$$

Это отображение должно обладать свойствами:

(1) $p(0) = 0$.

(2) $p(\mathbf{1}) = 1$.

(3) Если $\mathbf{E}_1\mathbf{E}_2 = 0$, то $p(\mathbf{E}_1 + \mathbf{E}_2) = p(\mathbf{E}_1) + p(\mathbf{E}_2)$.

Решающим здесь является постулат (3). Он утверждает, что (поскольку проекции на взаимно ортогональные пространства могут рассматриваться как взаимно исключающие альтернативы) вероятности, приписываемые взаимно ортогональным проекциям, должны быть аддитивными. Это очень сильное предположение, поскольку существует много различных способов выбора $\mathbf{E}_1$ и $\mathbf{E}_2$. Грубо говоря, первые два предположения утверждают, что, какое бы измерение мы ни делали: (1) всегда есть его результат; (2) вероятность суммы всех возможных (взаимно ортогональных) исходов равна единице.

В этих предположениях Глизон показал, что если размерность гильбертова пространства больше двух, то для любого такого отображения существует эрмитовский, положительный оператор $\boldsymbol{\rho}$ с единичным следом $\mathrm{tr}\,\boldsymbol{\rho} = 1$ такой, что

$$p(\mathbf{E}) = \mathrm{tr}(\boldsymbol{\rho}\mathbf{E}). \tag{2.72}$$

Таким образом, формализм матрицы плотности действительно является *необходимым*, если мы представляем наблюдаемые самосопряженными операторами в гильбертовом пространстве и приписываем определенные вероятности каждому возможному результату измерения. Грубо говоря, требование аддитивности вероятностей взаимно исключающих результатов настолько сильно, что мы с необходимостью приходим к линейному выражению (2.72).

Случай двумерного гильбертова пространства является особым, поскольку именно в двух измерениях оказывается недостаточным взаимно исключающих проекций. Все нетривиальные проекции имеют вид

$$\mathbf{E}(\hat{n}) = \frac{1}{2}(\mathbf{1} + \hat{n} \cdot \hat{\boldsymbol{\sigma}}), \tag{2.73}$$

но

$$\mathbf{E}(\hat{n})\mathbf{E}(\hat{m}) = 0 \tag{2.74}$$

только при $\hat{m} = -\hat{n}$; следовательно, любая определенная на двумерной сфере функция $f(\hat{n})$ такая, что $f(\hat{n}) + f(-\hat{n}) = 1$, удовлетворяет условиям теоремы Глизона, а таких функций существует много. Но в трех измерениях может быть больше альтернативных способов разбиения единицы, так что предположения Глизона гораздо сильнее. Мы не приводим здесь доказательства теоремы. (Для обсуждения см. книгу Переса, стр. 190)[1].

2.3.4. Эволюция оператора плотности

До сих пор мы не обсуждали эволюцию во времени смешанных состояний. В случае бинарной системы, подчиняющейся обычным аксиомам квантовой теории, предположим, что ее гамильтониан в пространстве $\mathcal{H}_A \otimes \mathcal{H}_B$ имеет вид

$$\mathbf{H}_{AB} = \mathbf{H}_A \otimes \mathbf{1}_B + \mathbf{1}_A \otimes \mathbf{H}_B. \tag{2.75}$$

В этом предположении подсистемы A и B не связаны между собой, так что каждая из них эволюционирует независимо. Оператор эволюции комбинированной системы

$$\mathbf{U}_{AB}(t) = \mathbf{U}_A(t) \otimes \mathbf{U}_B(t) \tag{2.76}$$

расцепляется на отдельные унитарные операторы эволюции, действующие каждый на свою систему.

Тогда в шредингеровской картине динамики начальное чистое состояние $|\psi(0)\rangle_{AB}$ бинарной системы, заданное уравнением (2.57), эволюционирует как

$$|\psi(t)\rangle_{AB} = \sum_{i,\mu} a_{i\mu} |i(t)\rangle_A \otimes |\mu(t)\rangle_B, \tag{2.77}$$

где

$$\begin{aligned} |i(t)\rangle_A &= \mathbf{U}_A(t)|i(0)\rangle_A, \\ |\mu(t)\rangle_B &= \mathbf{U}_B(t)|\mu(0)\rangle_A \end{aligned} \tag{2.78}$$

[1] A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers, New York et al (2002).

определяют новый ортогональный базис в $\mathcal{H}_A$ и $\mathcal{H}_B$ [так как $\mathbf{U}_A(t)$ и $\mathbf{U}_B(t)$ унитарны]. Вычисляя, как и раньше, частичный след, находим

$$\boldsymbol{\rho}_A(t) = \sum_{i,j,\mu} a_{i\mu} a^*_{j\mu} |i(t)\rangle_A \, {}_A\langle j(t)| = \mathbf{U}_A(t) \boldsymbol{\rho}_A(0) \mathbf{U}_A^\dagger(t). \tag{2.79}$$

Таким образом, $\mathbf{U}_A(t)$ определяет эволюцию матрицы плотности.

В частности, в базисе, в котором $\boldsymbol{\rho}_A(0)$ диагональна, имеем

$$\boldsymbol{\rho}_A(t) = \sum_a p_a \mathbf{U}_A(t) |\psi_a(0)\rangle_A \, {}_A\langle\psi_a(0)| \mathbf{U}_A^\dagger(t). \tag{2.80}$$

Уравнение (2.80) говорит о том, что эволюция $\boldsymbol{\rho}_A$ полностью согласуется с интерпретацией ансамбля. Эволюция во времени каждого состояния в ансамбле управляется оператором $\mathbf{U}_A(t)$. Если состояние $|\psi_a(0)\rangle_A$ с вероятностью p_a возникает в момент времени $t = 0$, то с той же вероятностью p_a состояние $|\psi_a(t)\rangle_A$ возникает в последующий момент времени t.

С другой стороны, должно быть ясным, что уравнение (2.80) справедливо только в предположении о том, что системы A и B динамически *не связаны* между собой. Ниже мы исследуем, как эволюционирует матрица плотности при более общих условиях.

2.4. Разложение Шмидта

Чистое двухкубитовое состояние может быть представлено в стандартной форме (*разложение Шмидта*), которая часто оказывается полезной.

Чтобы прийти к этой форме, заметим, что произвольный вектор из $\mathcal{H}_A \otimes \mathcal{H}_B$ может быть разложен как

$$|\psi\rangle_{AB} = \sum_{i,\mu} a_{i\mu} |i\rangle_A |\mu\rangle_B \equiv \sum_i |i\rangle_A |\tilde{i}\rangle_B. \tag{2.81}$$

Здесь $\{|i\rangle_A\}$ и $\{|\mu\rangle_B\}$ — ортогональные базисы в $\mathcal{H}_A$ и $\mathcal{H}_B$ соответственно. Чтобы получить второе равенство в (2.81), мы положили по определению

$$|\tilde{i}\rangle_B \equiv \sum_\mu a_{i\mu} |\mu\rangle_B. \tag{2.82}$$

Заметим, что векторы $|\tilde{i}\rangle_B$ *не* обязаны быть взаимно ортогональными или нормированными.

Предположим теперь, что $\{|i\rangle_A\}$ – базис, в котором ρ_A диагональна

$$\rho_A = \sum_i p_i |i\rangle_{A\,A}\langle i|. \tag{2.83}$$

Мы можем также вычислить ρ_A, выполняя операцию взятия частичного следа

$$\begin{aligned}\rho_A &= \mathrm{tr}_B\left(|\psi\rangle_{AB\,AB}\langle\psi|\right) = \\ &= \mathrm{tr}_B\left(\sum_{ij}|i\rangle_{A\,A}\langle j|\otimes|\tilde{i}\rangle_{B\,B}\langle\tilde{j}|\right) = \sum_{ij}{}_B\langle\tilde{j}|\tilde{i}\rangle_B\left(|i\rangle_{A\,A}\langle j|\right).\end{aligned} \tag{2.84}$$

Последнее равенство в (2.84) получено с учетом того, что

$$\begin{aligned}\mathrm{tr}_B\left(|\tilde{i}\rangle_{B\,B}\langle\tilde{j}|\right) &= \sum_k {}_B\langle k|\tilde{i}\rangle_{B\,B}\langle\tilde{j}|k\rangle_B = \\ &= \sum_k {}_B\langle\tilde{j}|k\rangle_{B\,B}\langle k|\tilde{i}\rangle_B = {}_B\langle\tilde{j}|\tilde{i}\rangle_B,\end{aligned} \tag{2.85}$$

где $\{|k\rangle_B\}$ – ортонормированный базис в $\mathcal{H}_B$. Сравнивая уравнения (2.83) и (2.84), мы видим, что

$${}_B\langle\tilde{j}|\tilde{i}\rangle_B = p_i\delta_{ij}. \tag{2.86}$$

Следовательно, в конце концов оказалось, что $\{|\tilde{i}\rangle_B\}$ *взаимно ортогональны*. Соответствующим изменением масштаба мы получаем ортонормированные векторы

$$|i'\rangle_B = p_i^{-1/2}|\tilde{i}\rangle_B \tag{2.87}$$

[мы можем полагать $p_i \neq 0$, поскольку уравнение (2.87) необходимо только для фигурирующих в сумме (2.83) слагаемых] и, следовательно, разложение

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}|i\rangle_A|i'\rangle_B \tag{2.88}$$

в *специальных* ортонормированных базисах в $\mathcal{H}_A$ и $\mathcal{H}_B$.

Уравнение (2.88) представляет собой разложение Шмидта чистого двухкубитового состояния $|\psi\rangle_{AB}$[1]. Любое чистое двухкубитовое состояние может быть представлено в этой форме, но, естественно, используемые при

[1]Разложение Шмидта было получено задолго до появления квантовой механики: E. Schmidt, *Zur theorie der linearen and nichtlinearen integralgleichungen*, Math. Annalen, **63**, 433–476 (1906). – *Прим. ред.*

этом базисы зависят от разлагаемого состояния. В общем случае мы не можем одновременно разложить $|\psi\rangle_{AB}$ и $|\varphi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ в виде (2.88), используя для этого одни и те же ортонормированные базисы в $\mathcal{H}_A$ и $\mathcal{H}_B$.

Используя уравнение (2.88), мы можем также вычислить частичный след по $\mathcal{H}_A$ и получить

$$\boldsymbol{\rho}_B = \mathrm{tr}_A \left(|\psi\rangle_{AB\ AB}\langle\psi| \right) = \sum_i p_i |i'\rangle_{B\ B}\langle i'|. \tag{2.89}$$

Мы видим, что $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$ имеют *одинаковые ненулевые собственные числа*. Конечно, при этом совсем не обязательно, чтобы $\mathcal{H}_A$ и $\mathcal{H}_B$ имели одинаковые размерности, так что количество *нулевых* собственных значений $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$ может различаться.

Если $\boldsymbol{\rho}_A$ (и, следовательно, $\boldsymbol{\rho}_B$) не имеет других вырожденных собственных значений, кроме нулевых, тогда разложение Шмидта состояния $|\psi\rangle_{AB}$ существенно однозначно определяется матрицами плотности $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$. Мы можем диагонализовать $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$, чтобы найти векторы $|i\rangle_A$ и $|i'\rangle_B$. После этого мы объединим в пары собственные состояния $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$, отвечающие одинаковым собственным значениям, чтобы получить (2.88). Мы выбрали фазы наших базисных состояний так, чтобы они не возникали в коэффициентах в суммах; сохранилась лишь свобода переопределения векторов $|i\rangle_A$ и $|i'\rangle_B$ путем умножения их на противоположные фазы (что, конечно, оставляет неизменным выражение (2.88)).

Но если $\boldsymbol{\rho}_A$ имеет вырожденные ненулевые собственные значения, тогда нам необходимо больше информации, чем та, которая позволила определить разложение Шмидта состояния $|\psi\rangle_{AB}$ по $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$; нам нужно знать, какое состояние $|i'\rangle_B$ объединяется в пару с $|i\rangle_A$. Например, если $\mathcal{H}_A$ и $\mathcal{H}_B$ N-мерны, а $\mathbf{U}_{ij}$ — произвольная унитарная $N \times N$-матрица, то

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} |i\rangle_A \mathbf{U}_{ij} |j'\rangle_B \tag{2.90}$$

будет давать $\boldsymbol{\rho}_A = \boldsymbol{\rho}_B = \frac{1}{N}\mathbf{1}$ после вычисления парциальных частичных следов. Более того, мы можем выполнить одновременно унитарные преобразования в $\mathcal{H}_A$ и $\mathcal{H}_B$:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{N}} \sum_i |i\rangle_A |i'\rangle_B = \frac{1}{\sqrt{N}} \sum_{ijk} \mathbf{U}^*_{ij} |j\rangle_A \mathbf{U}_{ik} |k'\rangle_B ; \tag{2.91}$$

это сохраняет состояние $|\psi\rangle_{AB}$, но иллюстрирует имеющуюся здесь неоднозначность базиса, используемого в разложении Шмидта состояния $|\psi\rangle_{AB}$.

2.4.1. Запутанность

Каждому чистому двухкубитовому состоянию $|\psi\rangle_{AB}$ можно сопоставить положительное целое число, *число Шмидта*, представляющее собой количество ненулевых собственных значений ρ_A (или ρ_B) и, следовательно, — число слагаемых в разложении Шмидта состояния $|\psi\rangle_{AB}$. С помощью этой величины мы можем определить, что значит быть *запутанным* для чистого двухкубитового состояния: $|\psi\rangle_{AB}$ запутано (или несепарабельно) если его число Шмидта больше единицы; в противном случае оно *сепарабельно* (или не запутано). Таким образом, сепарабельное чистое состояние двух кубитов представляет собой прямое произведение чистых состояний из $\mathcal{H}_A$ и $\mathcal{H}_B$:

$$|\psi\rangle_{AB} = |\varphi\rangle_A \otimes |\chi\rangle_B ; \tag{2.92}$$

тогда и приведенные матрицы плотности $\rho_A = |\varphi\rangle_A\,{}_A\langle\varphi|$, $\rho_B = |\chi\rangle_B\,{}_B\langle\chi|$ являются чистыми. Любое состояние, которое не может быть представлено в виде такого прямого произведения, является запутанным; тогда ρ_A и ρ_B представляют собой смешанные состояния.

Одна из наших главных целей этого семестра — лучше понять смысл запутанности. Иногда не совсем строго и корректно говорят, что подсистемы A и B не коррелированы, если состояние $|\psi\rangle_{AB}$ — сепарабельно; в конце концов, два спина в сепарабельном состоянии

$$|\uparrow\rangle_A|\uparrow\rangle_B, \tag{2.93}$$

несомненно, коррелированы — оба они ориентированы в одном направлении. Однако характер корреляций между A и B в запутанном и сепарабельном состояниях различен. Вероятно, решающим различием является то, что *запутанность нелокальна*. Единственным способом запутать A и B является для двух подсистем их непосредственное взаимодействие друг с другом.

Мы можем приготовить состояние (2.93), не приводя спины A и B в контакт друг с другом. Нам нужно только послать сообщение (классическое!) двум ассистентам (Алисе и Бобу[1]), чтобы оба они приготовили спин в состоянии, ориентированном вдоль оси z. Но единственным способом превратить (2.93) в запутанное состояние, типа

$$\frac{1}{\sqrt{2}}\left(|\uparrow\rangle_A|\uparrow\rangle_B + |\downarrow\rangle_A|\downarrow\rangle_B\right), \tag{2.94}$$

является применение к нему *коллективного* унитарного преобразования. Локальные унитарные преобразования вида $\mathbf{U}_A \otimes \mathbf{U}_B$ и выполненные Алисой и Бобом локальные измерения *не могут увеличить число Шмидта*

[1] Алиса и Боб (A и B) — традиционные персонажи теории квантовой информации. — *Прим. ред.*

двухкубитового состояния, независимо от того, как долго Алиса и Боб обсуждали свои действия. Чтобы запутать два кубита, мы *должны* собрать их вместе и позволить им взаимодействовать.

Как мы обсудим позже, можно также определить различие между запутанным и сепарабельным двухкубитовыми *смешанными* состояниями. Мы также обсудим различные способы, которыми локальные операции могут модифицировать форму запутанности, а также некоторые возможности использования запутанности.

2.5. Неоднозначность интерпретации ансамблей

2.5.1. Выпуклость

Напомним, что оператор ρ, действующий в гильбертовом пространстве $\mathcal{H}$, может рассматриваться как оператор плотности, если он обладает тремя свойствами:

(1) ρ самосопряжен;

(2) ρ неотрицателен;

(3) $\mathrm{tr}(\rho) = 1$.

Отсюда непосредственно следует, что из двух данных матриц плотности ρ_1 и ρ_2 мы всегда можем построить другую матрицу плотности как выпуклую линейную комбинацию: при любом вещественном λ, удовлетворяющем $0 \leqslant \lambda \leqslant 1$,

$$\rho(\lambda) = \lambda\rho_1 + (1-\lambda)\rho_2 \tag{2.95}$$

представляет собой матрицу плотности. Легко видеть, что $\rho(\lambda)$ удовлетворяет свойствам (1) и (3), если ими обладают ρ_1 и ρ_2. Чтобы проверить (2), вычислим

$$\langle\psi|\rho(\lambda)|\psi\rangle = \lambda\langle\psi|\rho_1|\psi\rangle + (1-\lambda)\langle\psi|\rho_2|\psi\rangle \geqslant 0; \tag{2.96}$$

неотрицательность $\rho(\lambda)$ обеспечивается тем, что таковыми являются ρ_1 и ρ_2. Таким образом, мы показали, что в гильбертовом пространстве $\mathcal{H}$ размерности N операторы плотности образуют *выпуклое подмножество* вещественного векторного пространства эрмитовых $N \times N$-матриц. (Подмножество векторного пространства называется выпуклым, если оно содержит отрезки прямых линий, соединяющие любые две его точки.)

Большинство операторов плотности могут быть многими различными способами представлены в виде сумм других операторов плотности. В этом отношении чистые состояния занимают особое положение — невозможно выразить чистое состояние как выпуклую сумму двух других чистых состояний. Рассмотрим чистое состояние $\rho = |\psi\rangle\langle\psi|$; пусть $|\psi_\perp\rangle$ обозначает вектор, ортогональный $|\psi\rangle$, $\langle\psi_\perp|\psi\rangle = 0$. Предположим, что ρ можно разложить, как в уравнении (2.95); тогда

$$\langle\psi_\perp|\rho|\psi_\perp\rangle = 0 = \lambda\langle\psi_\perp|\rho_1|\psi_\perp\rangle + (1-\lambda)\langle\psi_\perp|\rho_2|\psi_\perp\rangle. \qquad (2.97)$$

Так как правая часть является суммой двух неотрицательных слагаемых, оба они должны быть одновременно равны нулю. Если λ не нуль и не единица, то отсюда следует, что ρ_1 и ρ_2 ортогональны $|\psi_\perp\rangle$. Но поскольку $|\psi_\perp\rangle$ может быть любым вектором, ортогональным $|\psi\rangle$, мы приходим к выводу, что $\rho_1 = \rho_2 = \rho$.

Векторы выпуклого множества, которые не могут быть представлены в виде линейной комбинации других векторов этого множества, называются *крайними точками* множества. Мы только что показали, что чистые состояния являются крайними точками множества матриц плотности. Более того, *только* чистые состояния являются крайними, поскольку любое смешанное состояние может быть записано как $\rho = \sum_i p_i|i\rangle\langle i|$ в базисе, в котором оно диагонально, что является выпуклой суммой чистых состояний.

Мы уже встречались с этой структурой в обсуждении частного случая сферы Блоха. Мы говорили, что операторы плотности заполняют (единичный) шар в трехмерном множестве эрмитовых 2×2-матриц с единичным следом. Шар является выпуклым, а его крайними точками являются точки на поверхности. Аналогично, $N \times N$-операторы плотности образуют выпуклое подмножество (N^2-1)-мерного множества эрмитовых $N \times N$-матриц с единичным следом, крайними точками которого являются чистые состояния.

Однако в одном отношении 2×2-случай нетипичен: при $N > 2$ точки на границе множества матриц плотности не обязательно являются чистыми состояниями. Границу множества образуют все матрицы плотности, имеющие хотя бы одно нулевое собственное значение (поскольку существуют сколь угодно близкие к ним матрицы с отрицательными собственными значениями). Такая матрица плотности при $N > 2$ не является чистой, поскольку число ее ненулевых собственных чисел может превышать единицу.

2.5.2. Приготовление ансамбля

Выпуклость множества матриц плотности имеет простую, проясняющую суть дела, физическую интерпретацию. Допустим, что ассистент со-

класси приготовить одно из двух возможных состояний; состояние ρ_1 готовится с вероятностью λ, а состояние ρ_2 — с вероятностью $1-\lambda$. (Можно воспользоваться генератором случайных чисел, чтобы осуществить этот выбор.) Чтобы вычислить ожидаемое значение любой наблюдасмой M, мы усредняем ее по *обоим* выборам приготовления, а результат квантового измерения есть

$$\begin{aligned}\langle \mathbf{M} \rangle &= \lambda \langle \mathbf{M} \rangle_1 + (1-\lambda)\langle \mathbf{M} \rangle_2 = \\ &= \lambda \operatorname{tr}(\mathbf{M}\rho_1) + (1-\lambda)\operatorname{tr}(\mathbf{M}\rho_2) = \\ &= \operatorname{tr}(\mathbf{M}\rho(\lambda)).\end{aligned} \tag{2.98}$$

Таким образом, все ожидаемые значения не отличимы от тех, что мы получили бы, если бы вместо этого было приготовлено состояние $\rho(\lambda)$ Таким образом, мы имеем операционную процедуру, данные методы приготовления состояний ρ_1 и ρ_2 для приготовления произвольной выпуклой комбинации.

Действительно, для любого смешанного состояния ρ существует бесконечное множество способов его представления в виде выпуклой комбинации других состояний и, следовательно, бесконечное разнообразие процедур, которые мы могли бы применить для приготовления ρ. И каждая из этих процедур ведет к одним и тем же следствиям для любой мыслимой наблюдаемой рассматриваемой системы. Но чистое состояние совсем другое — оно может быть приготовлено одним единственным способом. (То есть оно «чистое» относительно чистых состояний.) Каждое чистое состояние является собственным состоянием некоторой наблюдаемой, например, для состояния $\rho = |\psi\rangle\langle\psi|$ измерение проекции $\mathbf{E} = |\psi\rangle\langle\psi|$ гарантирует результат 1. (В качестве примера вспомним, что каждое чистое состояние одного кубита представляет собой состояние типа «спин вверх» вдоль некоторой оси). Поскольку ρ является состоянием, для которого результат измерения $\mathbf{E}$ со 100% вероятностью равен единице, нет никакой возможности воспроизвести это наблюдаемое свойство, выбирая один из нескольких возможных способов его приготовления. Таким образом, приготовление чистого состояния совершенно однозначно (мы можем установить этот единственный способ его приготовления, если имеем множество копий состояния, чтобы поэкспериментировать с ними), тогда как в приготовлении смешанного состояния всегда возможны варианты.

Как велика эта неоднозначность? Так как любое ρ можно представить в виде суммы чистых состояний, задержим наше внимание на следующем вопросе: насколько большим числом способов может быть представ-

лен оператор плотности как выпуклая сумма чистых состояний? Математически этот вопрос звучит так: насколько большим числом способов можно записать ρ в виде суммы *крайних состояний*?

В качестве первого примера рассмотрим «максимально смешанное» состояние одного кубита:

$$\rho = \frac{1}{2}\mathbf{1}. \tag{2.99}$$

Такое состояние действительно может быть приготовлено бесконечным числом способов как ансамбль чистых состояний. Например,

$$\rho = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\downarrow_z\rangle\langle\downarrow_z|, \tag{2.100}$$

такое ρ мы получим, если приготовим состояния или $|\uparrow_z\rangle$, или $|\downarrow_z\rangle$, появляющиеся с вероятностью $\frac{1}{2}$ каждое. Но также мы имеем

$$\rho = \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| + \frac{1}{2}|\downarrow_x\rangle\langle\downarrow_x|, \tag{2.101}$$

такое ρ мы получим, если приготовим состояния или $|\uparrow_x\rangle$, или $|\downarrow_x\rangle$, появляющиеся с вероятностью $\frac{1}{2}$ каждое. Эта процедура приготовления, бесспорно, *другая*. Тем не менее, наблюдая за одним спином, разницу между ними обнаружить невозможно.

И вообще, центральная точка шара Блоха является суммой любых двух диаметрально противоположных точек на сфере, поэтому, приготовив или $|\uparrow_A\rangle$, или $|\downarrow_A\rangle$, появляющиеся с вероятностью $\frac{1}{2}$ каждое, мы тем самым приготовим состояние $\rho = \frac{1}{2}\mathbf{1}$.

Различные способы приготовления ρ из ансамбля *взаимно ортогональных* чистых состояний существуют только тогда, когда ρ имеет два (или более) вырожденных собственных значения. Однако у нас нет серьезных оснований ограничивать наше внимание этими ансамблях. Мы можем рассмотреть точку внутри шара Блоха

$$\rho(\vec{P}) = \frac{1}{2}(\mathbf{1} + \vec{P}\cdot\vec{\sigma}) \tag{2.102}$$

с $0 < |\vec{P}| < 1$; это состояние также можно выразить как

$$\rho(\vec{P}) = \lambda\rho(\hat{n}_1) + (1-\lambda)\rho(\hat{n}_2), \tag{2.103}$$

если $\vec{P} = \lambda\hat{n}_1 + (1-\lambda)\hat{n}_2$ (или, другими словами, $\vec{P}$ лежит где-нибудь на отрезке прямой, соединяющей точки сферы $\hat{n}_1$ и $\hat{n}_2$). Очевидно, что для любого $\vec{P}$ существует такое решение, связанное с любой хордой сферы, проходящей через точку $\vec{P}$; все такие хорды образуют двухпараметрическое семейство.

Эта высокая степень неоднозначности приготовления смешанного квантового состояния является одной из характерных особенностей квантовой информации, которая резко контрастирует с классическими вероятностными распределениями. Рассмотрим в качестве примера распределение вероятностей одного классического бита. Существует два крайних распределения таких, в которых 0 или 1 возникают со 100% вероятностью. *Любое* распределение вероятностей для бита является выпуклой суммой этих двух крайних точек. Аналогично, если имеется N возможных состояний, то существует N крайних распределений, а любое распределение вероятностей имеет *единственное* разложение по этим крайним распределениям (выпуклое множество распределений вероятностей образует *симплекс*). Если 0 появляется с вероятностью 21%, 1 — с вероятностью 33%, а 2 — с вероятностью 46%, то существует единственная процедура приготовления, которая дает это распределение вероятностей!

2.5.3. Быстрее света?

Вернемся теперь к нашей ранней точке зрения — смешанное состояние системы A возникает вследствие *запутывания* A с системой B — чтобы продолжить рассмотрение последствий неоднозначности приготовления смешанных состояний. Если кубит имеет матрицу плотности

$$\rho_A = \frac{1}{2}|\uparrow_z\rangle_A\,{}_A\langle\uparrow_z| + \frac{1}{2}|\downarrow_z\rangle_A\,{}_A\langle\downarrow_z|, \tag{2.104}$$

эта матрица плотности могла бы возникнуть в результате запутывания двухкубитового чистого состояния $|\psi\rangle_{AB}$, представимого в виде разложения Шмидта:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle_A|\uparrow_z\rangle_B + |\downarrow_z\rangle_A|\downarrow_z\rangle_B\right). \tag{2.105}$$

Следовательно, интерпретация ансамбля ρ_A, в котором приготовлено или состояние $|\uparrow_z\rangle_A$, или состояние $|\downarrow_z\rangle_A$ (каждое с вероятностью $p = 1/2$), может быть реализована выполнением измерения кубита B. Мы измеряем кубит B в базисе $\{|\uparrow_z\rangle_B, |\downarrow_z\rangle_B\}$; если получается результат $|\uparrow_z\rangle_B$, то приготовлено состояние $|\uparrow_z\rangle_A$, если же получается результат $|\downarrow_z\rangle_B$, то приготовлено $|\downarrow_z\rangle_A$.

Но, как уже отмечалось, в этом случае базис Шмидта не единственен, поскольку ρ_A имеет вырожденные собственные значения. Мы можем одновременно применить унитарные преобразования к кубитам A и B (если мы применяем $\mathbf{U}$ к A, то к B мы должны применить $\mathbf{U}^*$), не меняя при этом двухкубитовое чистое состояние $|\psi\rangle_{AB}$. Следовательно, для *любого* единичного трехмерного вектора $\hat{n}$ состояние $|\psi\rangle_{AB}$ имеет разложение Шмидта вида

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\uparrow_{\hat{n}}\rangle_A|\uparrow_{\hat{n}'}\rangle_B + |\downarrow_{\hat{n}}\rangle_A|\downarrow_{\hat{n}'}\rangle_B\right). \qquad (2.106)$$

Отсюда видно, что, измеряя кубит B в подходящем базисе, мы можем реализовать *любую* интерпретацию ρ_A как ансамбля двух чистых состояний.

Вдумчивые студенты после знакомства с этим свойством иногда загораются идеей предложить механизм сверхсветовой системы связи. Готовится много копий состояния $|\psi\rangle_{AB}$. Алиса забирает кубиты A с собой на туманность Андромеды, а Боб оставляет все кубиты B на Земле. Когда Боб хочет послать Алисе однобитовое сообщение, он выбирает, измерить ли ему σ_1 или σ_3 на всех своих спинах, приготовив таким образом спины Алисы в одном из двух ансамблей: $\{|\uparrow_x\rangle_A, |\downarrow_x\rangle_A\}$ или $\{|\uparrow_z\rangle_A, |\downarrow_z\rangle_A\}$.[1] Чтобы прочитать сообщение, Алиса сразу вслед за этим измеряет свои спины, чтобы увидеть, какой ансамбль был приготовлен.

Но *еще более* вдумчивые студенты (или студенты, слышавшие предыдущую лекцию) могут разглядеть изъян в этой схеме. Несмотря на то, что оба метода приготовления несомненно различны, оба ансамбля описываются в точности одной и той же матрицей плотности ρ_A. Таким образом, Алиса не может сделать никакого мыслимого измерения, чтобы различить эти два ансамбля, и нет возможности сообщить ей, какое действие совершил Боб. Сообщение «нечитабельно».

Почему тогда мы так уверенно утверждаем, что «оба метода приготовления несомненно различны»? Чтобы развеять любые сомнения относительно этого, представим, что Боб: (1) измеряет все свои спины вдоль оси $\hat{x}$ или (2) измеряет все свои спины вдоль оси $\hat{z}$, а затем вызывает Алису по межгалактическому телефону. Он *не говорит* Алисе, какое измерение он выполнил, (1) или (2), но сообщает ей все их результаты: «первый спин направлен вверх, второй — вниз» и т. д.. Теперь Алиса выполняет измерения (1) или (2) со *своими* спинами. Если они оба измеряли вдоль одной и той же оси, то Алиса обнаружит, что каждый из результатов ее измерений согласуется с тем, что нашел Боб. Но если Алиса и Боб выполняли измерения вдоль разных (ортогональных) осей, то Алиса *не обнаружит никаких*

[1] В этом случае $\mathbf{U}$ вещественно, поэтому $\mathbf{U} = \mathbf{U}^*$, а $\hat{n} = \hat{n}'$.

корреляций между их результатами. Примерно половина результатов ее измерений будет согласоваться с результатами Боба, примерно половина — противоречить им. Если Боб обещает выполнить (1) или (2) и предполагается отсутствие ошибок в приготовлении или измерении, тогда Алиса будет знать, что их действия были различными (даже если Боб не сообщал ей этой информации), сразу, как только результат одного из ее измерений вступит в противоречие с тем, что нашел Боб. Если же результаты всех их измерений согласуются, тогда, если было проведено достаточно много измерений, с очень высоким уровнем значимости Алиса будет считать, что она и Боб выполняли измерения вдоль одной и той же оси. (Даже с учетом возможных ошибок измерения этот статистический тест будет оставаться надежным, если частота появления ошибок достаточно низка). Таким образом, Алиса имеет возможность различить два использованных Бобом метода приготовления, однако в этом случае нет сверхсветовой связи, поскольку прежде чем Алиса смогла выполнить свою проверку, ею получен телефонный вызов от Боба.

2.5.4. Квантовое удаление (информации)

Мы говорили, что матрица плотности $\rho_A = \frac{1}{2}\mathbf{1}_A$ описывает спин в *некогерентной* суперпозиции чистых состояний $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$. Оно отличимо от *когерентной* суперпозиции этих состояний, такой как

$$|\uparrow_x,\downarrow_x\rangle = \frac{1}{2}\left(|\uparrow_z\rangle \pm |\downarrow_z\rangle\right); \tag{2.107}$$

в этом случае относительная фаза двух состояний имеет наблюдаемые следствия (отличает $|\uparrow_x\rangle$ от $|\downarrow_x\rangle$). В случае некогерентной суперпозиции относительная фаза полностью ненаблюдаема. Суперпозиция становится некогерентной, если спин A запутывается с другим спином B, недоступным для наблюдения.

С эвристической точки зрения состояния $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$ могут *интерферировать* (может быть наблюдаемой относительная фаза этих состояний) только тогда, когда мы не имеем информации о том, находится ли спин в состоянии $|\uparrow_z\rangle_A$ или в состоянии $|\downarrow_z\rangle_A$. Даже более того, интерференция может наблюдаться только тогда, когда *в принципе нет никакой возможности* определить, находится ли спин в состоянии вверх или вниз вдоль оси $\hat{z}$. Запутывание спина A со спином B разрушает интерференцию (по причине *декогерентизации* A), поскольку у нас появляется принципиальная возможность определить, находится ли спин A в состоянии вверх или вниз вдоль оси $\hat{z}$, выполняя соответствующее измерение спина B.

Но сейчас мы увидим, что утверждение о том, что запутывание является причиной декогерентизации, требует оговорки. Допустим, что Боб измеряет спин B вдоль оси $\hat{x}$, получая в качестве результата или $|\uparrow_x\rangle_B$, или $|\downarrow_x\rangle_B$, и посылает результат своего измерения Алисе. *Теперь* спин Алисы находится в чистом состоянии (или в $|\uparrow_x\rangle_A$, или в $|\downarrow_x\rangle_A$), а фактически в когерентной суперпозиции $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$. Мы сумели восстановить чистоту состояния спина Алисы, прежде чем ее скрыло облако декогерентизации!

Предположим, что Боб позволил своему спину пройти через прибор Штерна–Герлаха, ориентированный вдоль оси $\hat{z}$. Ну, конечно, спин Алисы не может вести себя, как в состоянии когерентной суперпозиции $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$; всего лишь проследив за тем, по какому пути прошел его спин, Боб будет знать, ориентирован ли спин Алисы вверх или вниз вдоль оси $\hat{z}$. Но допустим, что Боб не производит наблюдений. Вместо этого он вновь тщательно фокусирует два пучка, не делая никакой записи о том, вверх или вниз переместился его спин, после чего позволяет пройти спину через второй прибор Штерна–Герлаха, ориентированный вдоль оси $\hat{x}$. На этот раз он производит наблюдение и сообщает Алисе результат своего измерения σ_1. Теперь когерентность состояния спина Алисы восстановлена!

Эта ситуация была названа *квантовым ластиком*. Запутывание двух спинов создает «ситуацию измерения», в котором теряется когерентность $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$, вследствие чего, наблюдая за спином B, мы можем выяснить, ориентирован спин A вверх или вниз вдоль оси $\hat{z}$. Но когда мы (вслед за этим) измеряем спин B вдоль оси $\hat{x}$, эта информация «стирается». Ни результат $|\uparrow_x\rangle_B$, ни $|\downarrow_x\rangle_B$ — ничего не сообщают нам о том, ориентирован ли спин A вверх или вниз вдоль оси $\hat{z}$, поскольку Боб не позаботился сохранить информацию «какой путь», что можно было сделать, наблюдая за движением спина в первом приборе Штерна – Герлаха[1]. Следовательно, для спина A вновь возможно поведение типа когерентной суперпозиции $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$ (и это *после того*, как Алиса узнает о результате Боба).

Мы можем лучше понять квантовый ластик с точки зрения ансамбля. Алиса имеет множество спинов, выбранных из ансамбля $\boldsymbol{\rho}_A = \frac{1}{2}\mathbf{1}_A$, и у нее нет возможности наблюдать интерференцию между состояниями $|\uparrow_z\rangle_A$ и $|\downarrow_z\rangle_A$. Когда Боб выполняет свое измерение вдоль оси $\hat{x}$, реализуется конкретно приготовленный ансамбль. Однако это не дает эффекта, который может почувствовать Алиса, — состояние ее спинов, *как и прежде*, описы-

[1]Часто говорят, что была стерта «welcher weg»-информация, поскольку по-немецки это звучит более изысканно [«welcher weg» (*нем.*) – «какой путь» (перев.)].

вается ансамблем $\rho_A = \frac{1}{2}\mathbf{1}_A$. Но когда Алиса получает телефонное сообщение от Боба, она может отобрать *субансамбль* из тех своих спинов, которые находятся в чистом состоянии $|\uparrow_x\rangle_A$. Сообщенная Бобом информация позволяет Алисе отделить чистые состояния от максимально смешанных.

Другой намек на квантовый ластик иногда называют *отложенным выбором*. Это значит, что описанная нами ситуация в действительности полностью симметрична относительно Алисы и Боба, то есть невозможно определить, кто первым произвел измерение. (Действительно, если измерения Алисы и Боба являются событиями, разделенными пространственно-подобным интервалом, то их следование друг за другом во времени неинвариантно, оно зависит от системы отсчета, используемой наблюдателем.) Алиса могла бы измерить все свои спины сегодня (скажем, вдоль оси $\hat{x}$), прежде чем Боб решит, как он будет измерять свои спины. На следующей неделе Боб решает «приготовить» спины Алисы в состояниях $|\uparrow_{\hat{n}}\rangle_A$ и $|\downarrow_{\hat{n}}\rangle_A$ (это и есть «отложенный выбор»). После этого он сообщает Алисе о том, какие спины были в состоянии $|\uparrow_{\hat{n}}\rangle_A$, а она может проконтролировать запись его измерения, чтобы убедиться, что

$$(\sigma_1)_{\hat{n}} = \hat{n} \cdot \hat{x}. \tag{2.108}$$

Результат будет один и тот же независимо от того, «приготовил» Боб спины до или после измерения Алисы.

Мы утверждали, что матрица плотности представляет полное физическое описание подсистемы A, поскольку она характеризует все возможные измерения, которые на ней могут быть выполнены. Иногда раздаются возражения[1], что явление квантового ластика демонстрирует обратное. Так как полученная от Боба информация даст Алисе возможность извлечь чистое состояние из смеси, то как можно утверждать, что в ρ_A закодировано все, что Алиса может узнать об A?

Я не считаю это правильным выводом. Скорее я хочу сказать, что квантовый ластик предоставляет еще одну возможность повторить наше заклинание: «Информация материальна». Состояние ρ_A системы A и состояние ρ_A, дополненное информацией, полученной Алисой от Боба, — не одно и то же. Эта информация (которая снабжает метками субансамбли) изменяет физическое описание. Чтобы выразить это математически, мы должны включить в наше описание «знание Алисы о состоянии». Ансамбль спинов, о котором Алиса не имеет информации, вверх или вниз направлен каждый

[1]Например, в книге: Roger Penrose, *Shadows of the Mind. A Search for the Missing Science of Consciousness*, Oxford University Press, New York et al, 1994; перевод Роджер Пенроуз, *Тени разума. В поисках науки о сознании*. Москва–Ижевск: ИКИ (2003)

спин, представляет собой другое состояние, нежели ансамбль, в котором Алисе известно состояние каждого спина[1].

2.5.5. Теорема ЖХЙВ

До сих пор мы рассматривали квантовый ластик только в связи с одиночным кубитом, описываемым ансамблем равновероятных, взаимно ортогональных состояний (то есть $\rho_A = \frac{1}{2}\mathbf{1}_A$). Это обсуждение можно существенно обобщить.

Мы уже видели, что смешанное состояние любой квантовой системы можно бесконечным числом различных способов реализовать как ансамбль чистых состояний. Рассмотрим одну такую реализацию для матрицы плотности ρ_A

$$\rho_A = \sum_i p_i |\varphi_i\rangle_A\, {}_A\langle\varphi_i|, \qquad \sum_i p_i = 1. \tag{2.109}$$

Здесь все состояния $\{|\varphi_i\rangle_A\}$ являются нормированными, но не обязательно взаимно ортогональными векторами. Тем не менее ρ_A можно понимать как ансамбль, в котором каждое чистое состояние $|\varphi_i\rangle_A\, {}_A\langle\varphi_i|$ возникает с вероятностью p_i.

Конечно, для любого такого ρ_A мы можем построить его «очищение», двухкубитовое чистое состояние $|\Phi_1\rangle_{AB}$, которое дает ρ_A при вычислении частичного следа в пространстве $\mathcal{H}_B$. Такое очищение имеет вид

$$|\Phi_1\rangle_{AB} = \sum_i \sqrt{p_i}\,|\varphi_i\rangle_A |\alpha_i\rangle_B, \tag{2.110}$$

где векторы $|\alpha_i\rangle_B \in \mathcal{H}_B$ взаимно ортогональны и нормированы:

$${}_B\langle\alpha_i|\alpha_j\rangle_B = \delta_{ij}. \tag{2.111}$$

Очевидно, что

$$\mathrm{tr}_B\left(|\Phi_1\rangle_{AB}\, {}_{AB}\langle\Phi_1|\right) = \rho_A. \tag{2.112}$$

Более того, мы можем представить выполнение ортогонального измерения в системе B, проецирующее на базис $|\alpha_i\rangle_B$.[2] С вероятностью p_i получится результат $|\alpha_i\rangle_B$ и приготовит чистое состояние $|\varphi_i\rangle_A\, {}_A\langle\varphi_i|$ системы A.

[1] Это «знание состояния» не должно быть действительно состоянием человеческого разума; достаточной будет любая (неодушевленная) запись, помечающая субансамбль.

[2] Пространство $\mathcal{H}_B$ может быть и не натянутым на векторы $|\alpha_i\rangle_B$, однако в состоянии $|\Phi_1\rangle_{AB}$ никогда не появляются результаты измерения, ортогональные всем $|\alpha_i\rangle_B$.

Таким образом, для данного очищения $|\Phi_1\rangle_{AB}$ состояния ρ_A существует такое измерение в системе B, при выполнении которого реализуется состояние $|\varphi_i\rangle_A$ ансамбля ρ_A. По известному результату измерения в B мы успешно выделили одно из чистых состояний $|\varphi_i\rangle_A$ смеси ρ_A.

Только что описанное представляет собой обобщение процедуры приготовления состояния $|\uparrow_z\rangle_A$ путем измерения спина B вдоль оси $\hat{z}$ (в нашем обсуждении двух запутанных кубитов). Но чтобы обобщить понятие квантового ластика, мы хотим видеть, что, выполняя разные измерения B в состоянии $|\Phi_1\rangle_{AB}$, можно реализовать *разные* интерпретации ансамбля ρ_A. Пусть

$$\rho_A = \sum_\mu q_\mu |\psi_\mu\rangle_A\,{}_A\langle\psi_\mu| \tag{2.113}$$

— другая реализация той же самой матрицы плотности ρ_A как ансамбля чистых состояний. Для этого ансамбля тоже существует соответствующее очищение

$$|\Phi_2\rangle_{AB} = \sum_\mu \sqrt{q_\mu}|\psi_\mu\rangle_A|\beta_\mu\rangle_B, \tag{2.114}$$

где вновь $\{|\beta_\mu\rangle\}$ — ортонормированные векторы из $\mathcal{H}_B$. Итак, в состоянии $|\Phi_2\rangle_{AB}$ мы можем реализовать ансамбль, выполняя в $\mathcal{H}_B$ измерение, проецирующее на базис $\{|\beta_\mu\rangle_B\}$.

Как связаны состояния $|\Phi_1\rangle_{AB}$ и $|\Phi_2\rangle_{AB}$? Фактически мы можем легко показать, что

$$|\Phi_1\rangle_{AB} = (\mathbf{1}_A \otimes \mathbf{U}_B)|\Phi_2\rangle_{AB}; \tag{2.115}$$

два состояния отличаются унитарным изменением базиса, действующим только в $\mathcal{H}_B$, или

$$|\Phi_1\rangle_{AB} = \sum_\mu \sqrt{q_\mu}|\psi_\mu\rangle_A|\gamma_\mu\rangle_B, \tag{2.116}$$

где

$$|\gamma_\mu\rangle_B = \mathbf{U}_B|\beta_\mu\rangle_B \tag{2.117}$$

— другой ортонормированный базис в $\mathcal{H}_B$. Итак, мы видим, что существует *единственное* очищение $|\Phi_1\rangle_{AB}$ смешанного состояния ρ_A такое что, выбирая измерение соответствующей наблюдаемой в системе B, мы можем реализовать либо $\{|\varphi_i\rangle\}$-ансамбль, либо $\{|\psi_\mu\rangle\}$-ансамбль!

Аналогично мы можем рассмотреть множество ансамблей, реализующих смешанное состояние ρ_A, где максимальное число чистых состояний, возникающих в любом из этих ансамблей, равно n. Тогда мы мо-

жем выбрать гильбертово пространство $\mathcal{H}_B$ размерности n и чистое состояние $|\Phi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ такое, что любой из ансамблей может быть реализован путем измерения соответствующей наблюдаемой в $\mathcal{H}_B$. В этом и состоит *теорема ЖХЙВ*[1]. Она выражает явление квантового ластика в его наиболее общей форме.

Фактически теорема ЖХЙВ является почти тривиальным следствием разложения Шмидта. Оба состояния, $|\Phi_1\rangle_{AB}$ и $|\Phi_2\rangle_{AB}$, имеют разложение Шмидта, а поскольку при вычислении частичного следа по B оба они дают одно и то же смешанное состояние ρ_A, эти разложения должны иметь вид

$$|\Phi_1\rangle_{AB} = \sum_k \sqrt{\lambda_k}|k\rangle_A|k_1'\rangle_B,$$
$$|\Phi_2\rangle_{AB} = \sum_k \sqrt{\lambda_k}|k\rangle_A|k_2'\rangle_B, \qquad (2.118)$$

где λ_k — собственные значения, а $|k\rangle_A$ — соответствующие им собственные векторы ρ_A. Но поскольку $\{|k_1'\rangle_B\}$ и $\{|k_2'\rangle_B\}$ — ортонормированные базисы в $\mathcal{H}_B$, существует такое унитарное преобразование $\mathbf{U}_B$, что

$$|k_1'\rangle_B = \mathbf{U}_B|k_2'\rangle_B, \qquad (2.119)$$

откуда непосредственно следует уравнение (2.115).

В ансамбле чистых состояний, описываемом уравнением (2.109), мы хотели бы сказать, что чистые состояния $|\varphi_i\rangle_A$ образуют в нем *некогерентную* суперпозицию — наблюдатель в системе A не может детектировать относительные фазы этих состояний. С эвристической точки зрения причина того, что эти состояния не могут интерферировать, состоит в том, что, выполняя в системе B измерение, проецирующее на ортонормированный базис $\{|\alpha_i\rangle_B\}$, в принципе мы имеем возможность обнаружить, какой представитель ансамбля реализован на самом деле. Однако проецируя вместо этого на базис $\{|\gamma_\mu\rangle_B\}$ и передавая в систему A информацию о результате измерения, мы можем выделить из ансамбля одно из чистых состояний $|\psi_\mu\rangle_A$, даже если оно может быть когерентной суперпозицией состояний $|\varphi_i\rangle_A$. В сущности, измерение B в базисе $\{|\gamma_\mu\rangle_B\}$ «стирает» «welcher weg»-информацию (состоянием A является либо $|\varphi_i\rangle_A$, либо $|\varphi_j\rangle_A$). В этом смысле теорема ЖХЙВ характеризует обобщенный квантовый «ластик». Еще раз повторим мораль, что *информация материальна* — информация, полученная в системе B, после ее передачи в A изменяет физическое описание состояния A.

[1]ЖХДжВ: Жизан, Хастон, Джозса, Вутерс.

2.6. Резюме

Аксиомы. Ареной квантовой механики служит гильбертово пространство $\mathcal{H}$. Основными предположениями являются:

(1) *Состояние* представляет собой *луч* в $\mathcal{H}$.

(2) *Наблюдаемая* представляется *самосопряженным оператором* в $\mathcal{H}$.

(3) *Измерением* является ортогональная *проекция*.

(4) *Эволюция во времени унитарна*.

Оператор плотности. Если мы ограничиваем наше внимание только на части большей квантовой системы, предположения (1)–(4) не выполняются. В частности, квантовое состояние описывается не лучом, а оператором плотности $\boldsymbol{\rho}$, неотрицательным оператором с единичным следом. Оператор плотности описывает *чистое состояние* (состояние может быть описано лучом), если $\boldsymbol{\rho}^2 = \boldsymbol{\rho}$; в противном случае состояние является *смешанным*. В этом состоянии наблюдаемая $\mathbf{M}$ имеет ожидаемое значение $\mathrm{tr}(\mathbf{M}\boldsymbol{\rho})$.

Кубит. Квантовая система, определенная в двумерном гильбертовом пространстве, называется *кубитом*. Наиболее общая матрица плотности кубита имеет вид

$$\boldsymbol{\rho}(\vec{P}) = \frac{1}{2}\left(\mathbf{1} + \vec{P}\cdot\vec{\boldsymbol{\sigma}}\right), \tag{2.120}$$

где $\vec{P}$ — трехкомпонентный вектор длины $|\vec{P}| \leqslant 1$. В чистом состоянии $|\vec{P}| = 1$.

Разложение Шмидта. Для любой квантовой системы, состоящей из двух частей A и B (*бинарная* система), гильбертово пространство является тензорным произведением $\mathcal{H}_A \otimes \mathcal{H}_B$. Для любого чистого состояния $|\psi\rangle_{AB}$ бинарной системы существуют ортонормированные базисы $\{|i\rangle_A\}$ в $\mathcal{H}_A$ и $\{|i'\rangle_B\}$ в $\mathcal{H}_B$ такие, что

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}|i\rangle_A|i'\rangle_B. \tag{2.121}$$

Уравнение (2.121) называется *разложением Шмидта* состояния $|\psi\rangle_{AB}$. В двухкубитовом чистом состоянии подсистемы A и B по отдельности описываются операторами плотности $\boldsymbol{\rho}_A$ и $\boldsymbol{\rho}_B$; из уравнения (2.121)

следует, что ρ_A и ρ_B имеют одинаковые ненулевые собственные значения (p_i). Количество ненулевых собственных значений называется *числом Шмидта* состояния $|\psi\rangle_{AB}$. Двухкубитовое чистое состояние называется *запутанным*, если его число Шмидта больше единицы.

Ансамбли. Операторы плотности в гильбертовом пространстве образуют выпуклое множество, а чистые состояния представляют собой *крайние точки* этого множества. Смешанное состояние системы A может быть приготовлено как *ансамбль* чистых состояний многими различными способами, неразличимыми экспериментально, если мы наблюдаем только систему A. Для любого смешанного состояния ρ_A системы A любое приготовление ρ_A как ансамбля чистых состояний, в принципе можно реализовать, выполняя измерение в другой системе B, с которой запутана система A. Фактически для множества таких приготовлений смешанного состояния ρ_A существует единственное запутанное состояние A и B такое, что любое из этих приготовлений может быть реализовано путем измерения соответствующей наблюдаемой в B (*теорема ЖХЙВ*). Выполняя измерение в системе B и сообщая его результат в систему A, мы можем выделить из смеси чистое состояние, выбранное из одного из ансамблей.

2.7. Упражнения

2.1. Точность воспроизведения вероятностной гипотезы. Один кубит (объект со спином-$\frac{1}{2}$) находится в неизвестном *чистом* состоянии $|\psi\rangle$, случайно выбранном из равномерно распределенного по сфере Блоха ансамбля. Мы наугад считаем, что этим состоянием является $|\phi\rangle$. Чему в среднем равна определяемая соотношением

$$F = |\langle\phi|\psi\rangle|^2 \tag{2.122}$$

точность воспроизведения нашей гипотезы?

2.2. Точность воспроизведения после измерения. После случайного выбора однокубитового чистого состояния, как в предыдущей задаче, мы выполняем измерение спина вдоль оси $\hat{z}$. Это измерение приготавливает состояние, описываемое матрицей плотности

$$\boldsymbol{\rho} = \mathbf{P}_\uparrow\langle\psi|\mathbf{P}_\uparrow|\psi\rangle + \mathbf{P}_\downarrow\langle\psi|\mathbf{P}_\downarrow|\psi\rangle \tag{2.123}$$

(где $\mathbf{P}_{\uparrow,\downarrow}$ обозначает проектор на состояние спин-вверх или спин-вниз вдоль оси $\hat{z}$). С какой в среднем точностью

$$F \equiv \langle\psi|\boldsymbol{\rho}|\psi\rangle \tag{2.124}$$

эта матрица плотности воспроизводит начальное состояние $|\psi\rangle$? (Улучшение F по сравнению с ответом к предыдущей задаче является грубой мерой того, как много мы узнаем, выполнив измерение).

2.3. Разложение Шмидта. Для двухкубитового состояния

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{2}}|\uparrow\rangle_A \left(\frac{1}{2}|\uparrow\rangle_B + \frac{\sqrt{3}}{2}|\downarrow\rangle_B\right) + \\ + \frac{1}{\sqrt{2}}|\downarrow\rangle_A \left(\frac{\sqrt{3}}{2}|\uparrow\rangle_B + \frac{1}{2}|\downarrow\rangle_B\right). \tag{2.125}$$

1) Вычислите $\boldsymbol{\rho}_A = \mathrm{tr}_B(|\Phi\rangle_{AB\,AB}\langle\Phi|)$ и $\boldsymbol{\rho}_B = \mathrm{tr}_A(|\Phi\rangle_{AB\,AB}\langle\Phi|)$.
2) Найдите разложение Шмидта состояния $|\Phi\rangle_{AB}$.

2.4. Трехкубитовое чистое состояние. Существует ли разложение Шмидта произвольного трехкубитового чистого состояния? То есть если $|\psi\rangle_{ABC}$ — произвольный вектор в $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, то можем ли мы найти ортогональные базисы $\{|i\rangle_A\}$, $\{|i\rangle_B\}$ и $\{|i\rangle_C\}$, такие что

$$|\psi\rangle_{ABC} = \sum_i \sqrt{p_i}|i\rangle_A \otimes |i\rangle_B \otimes |i\rangle_C\,? \tag{2.126}$$

Истолкуйте ваш ответ.

2.5. Квантовые корреляции в смешанном состоянии. Рассмотрим матрицу плотности двух кубитов:

$$\boldsymbol{\rho} = \frac{1}{8}\mathbf{1} + \frac{1}{2}|\psi^-\rangle\langle\psi^-|, \tag{2.127}$$

где $\mathbf{1}$ обозначает единичную 4×4-матрицу, а

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}\big(|\uparrow\rangle|\downarrow\rangle - |\downarrow\rangle|\uparrow\rangle\big) \tag{2.128}$$

Пусть мы измеряем первый спин вдоль оси $\hat{n}$, а второй — вдоль оси $\hat{m}$, где $\hat{n} \cdot \hat{m} = \cos\theta$. Какова вероятность того, что оба спина находятся в состоянии «спин-вверх» вдоль соответствующих осей?

Глава 3

Основы II: Измерение и эволюция

3.1. За пределами ортогональных измерений

3.1.1. Ортогональные измерения

Мы хотели бы исследовать свойства *обобщенных* измерений, которые можно реализовать в системе A, выполняя ортогональные измерения в большей системе, содержащей A. Но сначала, следуя классической трактовке фон Неймана, кратко обсудим, как в принципе могут быть выполнены ортогональные измерения произвольной наблюдаемой.

Чтобы измерить наблюдаемую M, мы модифицируем гамильтониан Вселенной, включая в него связь между этой наблюдаемой и «переменной-указателем» («pointer» variable), которая будет играть роль прибора. Связь запутывает собственные состояния наблюдаемой с различимыми состояниями прибора, так что, «наблюдая» за прибором, мы можем приготовить собственное состояние наблюдаемой.

Конечно, это не вполне удовлетворительная модель измерения, поскольку мы не объяснили, как можно измерить «переменную-указатель». Как можно видеть, позиция фон Неймана состояла в том, что в принципе возможно скоррелировать состояния микроскопической квантовой системы со значениями макроскопической классической переменной, которые, само собой разумеется, мы способны различать. Конечно, возможно и желательно более обстоятельное объяснение; мы еще вернемся к этой проблеме ниже.

Мы можем представлять себе прибор как пробную частицу, распространяющуюся свободно, за исключением ее регулируемой связи с измеряемой квантовой системой. Если мы намерены измерять положение этой пробной частицы, то в начальном состоянии должен быть приготовлен достаточно узкий волновой пакет; но не слишком, поскольку очень узкий волновой пакет будет слишком быстро расплываться. Если начальная ширина волнового пакета равна Δx, неопределенность его скорости будет иметь

порядок $\Delta v = \Delta p/m \sim \hbar/m\Delta x$, так что, спустя время t, пакет расползется до ширины

$$\Delta x(t) = \Delta x + \frac{\hbar t}{m\Delta x}, \tag{3.1}$$

минимальное значение которой имеет прядок $[\Delta x(t)]^2 \sim [\Delta x]^2 \sim \hbar t/m$. Следовательно, если эксперимент продолжается в течение времени t, то разрешение, которого мы можем добиться в определении конечного положения пробной частицы, ограничено «стандартным квантовым пределом»:

$$\Delta x \gtrsim (\Delta x)_{SQL} \sim \sqrt{\frac{\hbar t}{m}}. \tag{3.2}$$

Будем считать нашу пробную частицу достаточно тяжелой, чтобы это ограничение было несущественным.

Гамильтониан, описывающий взаимодействие квантовой системы с пробной частицей, имеет вид:

$$\mathbf{H} = \mathbf{H}_0 + \frac{1}{2m}\mathbf{P}^2 + \lambda \mathbf{M}\mathbf{P}, \tag{3.3}$$

где $\mathbf{P}^2/2m$ — гамильтониан свободной пробной частицы (который в дальнейшем будет игнорироваться, поскольку пробная частица настолько тяжела, что расплыванием ее волнового пакета можно пренебречь), $\mathbf{H}_0$ — невозмущенный гамильтониан измеряемой системы, λ — константа связи, которую мы можем менять по своему усмотрению. Измеряемая наблюдаемая $\mathbf{M}$ связана с импульсом $\mathbf{P}$ пробной частицы.

Если $\mathbf{M}$ не коммутирует с $\mathbf{H}_0$, то нас будет беспокоить, как эта наблюдаемая изменяется в процессе измерения. Чтобы упростить анализ, предположим, что или $[\mathbf{M}, \mathbf{H}_0] = 0$, или же измерение выполняется настолько быстро, что в ходе его можно пренебречь свободной эволюцией системы. Тогда гамильтониан (3.3) можно аппроксимировать одним слагаемым $\mathbf{H} \simeq \lambda \mathbf{M}\mathbf{P}$ (где, конечно же, $[\mathbf{M}, \mathbf{P}] = 0$, так как $\mathbf{M}$ — наблюдаемая системы, а $\mathbf{P}$ — наблюдаемая пробной частицы), а оператор эволюции во времени —

$$\mathbf{U}(t) \simeq \exp\big[-i\lambda t \mathbf{M}\mathbf{P}\big]. \tag{3.4}$$

Разлагая в базисе, в котором наблюдаемая $\mathbf{M}$ диагональна

$$\mathbf{M} = \sum_a |a\rangle M_a \langle a|, \tag{3.5}$$

представим $\mathbf{U}(t)$ в виде

$$\mathbf{U}(t) = \sum_a |a\rangle \exp\big[-i\lambda t M_a \mathbf{P}\big]\langle a|. \tag{3.6}$$

Вспомним теперь, что $\mathbf{P}$ генерирует параллельные переносы *положения* пробной частицы: в координатном представлении $\mathbf{P} = -i\frac{d}{dx}$, так что $\exp(-ix_0\mathbf{P}) = \exp\left(-x_0\frac{d}{dx}\right)$ и разложение в ряд Тейлора дает

$$\exp(-ix_0\mathbf{P})\psi(x) = \psi(x - x_0). \tag{3.7}$$

Другими словами, оператор $\exp(-ix_0\mathbf{P})$, действуя на волновой пакет, перемещает его на x_0. Отсюда видно, что если наша квантовая система начинает эволюцию из суперпозиции собственных состояний оператора $\mathbf{M}$, первоначально не запутанной с пространственным волновым пакетом пробной частицы $|\psi(x)\rangle$, то по истечении времени t ее квантовое состояние эволюционирует в

$$\mathbf{U}(t)\left(\sum_a \alpha_a|a\rangle \otimes |\psi(x)\rangle\right) = \sum_a \alpha_a|a\rangle \otimes |\psi(x - \lambda t M_a)\rangle. \tag{3.8}$$

Теперь положение пробной частицы коррелирует со значениями наблюдаемой $\mathbf{M}$. Если волновой пакет достаточно узок, чтобы мы могли разрешить все значения M_a (что имеет место при $|\Delta x| \lesssim |\lambda t M_a|$), тогда всякий раз, измеряя (неважно как!) положение пробной частицы, мы будем готовить собственное состояние наблюдаемой M_a. С вероятностью $|\alpha_a|^2$ мы обнаружим, что положение пробной частицы сместилось на величину $\lambda t M_a$, и, следовательно, приготовим собственное состояние $|a\rangle$ оператора $\mathbf{M}$. Таким образом, мы приходим к выводу, что начальное состояние $|\varphi\rangle$ квантовой системы проецируется на $|a\rangle$ с вероятностью $|\langle a|\varphi\rangle|^2$. Это и есть модель ортогонального измерения фон Неймана.

Классическим примером служит прибор Штерна–Герлаха. Чтобы измерить $\boldsymbol{\sigma}_3$ объекта со спином-$1/2$, мы пропускаем его через область неоднородного магнитного поля

$$B_3 = \lambda z. \tag{3.9}$$

Магнитный момент объекта равен $\mu\vec{\boldsymbol{\sigma}}$, а индуцируемая магнитным полем связь –

$$\mathbf{H} = -\lambda\mu z\boldsymbol{\sigma}_3. \tag{3.10}$$

В этом случае $\boldsymbol{\sigma}_3$ является измеряемой наблюдаемой, связанной с положением z, а не с импульсом пробной частицы. В этом нет ничего страшного, поскольку z генерирует трансляции импульса $\mathbf{P}_z$ и, следовательно, эта

связь сообщает *импульс* пробной частице. Мы можем различить, сдвинулся объект вверх или вниз и таким образом определить спиновое состояние $|\uparrow_z\rangle$ или $|\downarrow_z\rangle$. Конечно, поворачивая магнит, можно измерить наблюдаемую $\hat{n}\cdot\vec{\sigma}$.

Как показало обсуждение квантового ластика, одного только факта возникновения запутанного состояния (3.8) еще *не достаточно* для объяснения, почему процедура измерения готовит собственное состояние оператора $\mathbf{M}$. В принципе измерение пробной частицы могло бы спроецировать ее состояние на некоторую специфическую суперпозицию собственных состояний оператора положения, и таким образом приготовить квантовую систему в суперпозиции собственных состояний оператора $\mathbf{M}$. Чтобы достичь более глубокого понимания процесса измерения, нужно объяснить, почему базис собственных состояний положения пробной частицы имеет особый статус среди других возможных базисов.

Если мы действительно можем, как это только что было описано, связать любую наблюдаемую с измеримой «переменной-указателем», тогда мы в состоянии выполнить любую мыслимую ортогональную проекцию в гильбертовом пространстве. Для данного набора операторов $\{\mathbf{E}_a\}$ таких, что

$$\mathbf{E}_a = \mathbf{E}_a^{\dagger}, \qquad \mathbf{E}_a\mathbf{E}_b = \delta_{ab}\mathbf{E}_a, \qquad \sum_a \mathbf{E}_a = \mathbf{1}, \tag{3.11}$$

мы можем выполнить процедуру измерения, которое с вероятностью

$$\mathrm{Prob}(a) = \langle\psi|\mathbf{E}_a|\psi\rangle \tag{3.12}$$

преобразует чистое состояние $|\psi\rangle\langle\psi|$ в

$$\frac{\mathbf{E}_a|\psi\rangle\langle\psi|\mathbf{E}_a}{\langle\psi|\mathbf{E}_a|\psi\rangle}. \tag{3.13}$$

Результаты такого измерения можно описать матрицей плотности, которая получается суммированием всех возможных результатов (3.13) (а не выбором одного частного результата), взвешенных вероятностями их появления (3.12). В таком случае измерение преобразует начальное чистое состояние в соответствии с

$$|\psi\rangle\langle\psi| \rightarrow \sum_a \mathbf{E}_a|\psi\rangle\langle\psi|\mathbf{E}_a. \tag{3.14}$$

Это *ансамбль* чистых состояний, описывающих результаты измерения. Он описывает состояние, в котором известно выполненное в системе измерение, но неизвестен его результат. Следовательно, начальное чистое состояние превращается в смешанное, за исключением тех случаев, когда оно

оказывается собственным состоянием измеряемой наблюдаемой. Если же начальным состоянием перед измерением было смешанное состояние, описываемое матрицей плотности ρ, тогда, представляя ρ как ансамбль чистых состояний, мы обнаружим, что в результате измерения

$$\rho \to \sum_a \mathbf{E}_a \rho \mathbf{E}_a. \tag{3.15}$$

3.1.2. Обобщенные измерения

Теперь мы хотели бы обобщить понятие измерения, выйдя за пределы рассмотренных фон Нейманом ортогональных измерений. Путь, ведущий к идее обобщенного измерения, состоит в предположении, что наша система A расширена до тензорного произведения $\mathcal{H}_A \otimes \mathcal{H}_B$ и мы выполняем в нем ортогональные измерения, которые в самой системе A уже не обязательно ортогональны. Сначала мы пойдем несколько другим путем, который, хотя и не имеет физической мотивации, выглядит более естественно и просто с математической точки зрения.

Предположим, что гильбертово пространство $\mathcal{H}_A$ является частью более широкого пространства, имеющего структуру *прямой суммы*:

$$\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_A^{\perp}. \tag{3.16}$$

Нашим наблюдателям, «живущим» в $\mathcal{H}_A$, доступны только наблюдаемые с носителем в $\mathcal{H}_A$, то есть такие наблюдаемые $\mathbf{M}_A$, что для любого $|\psi^{\perp}\rangle \in \mathcal{H}_A^{\perp}$

$$\mathbf{M}_A|\psi^{\perp}\rangle = 0 = \langle\psi^{\perp}|\mathbf{M}_A. \tag{3.17}$$

Например, в двухкубитовом мире мы можем представить, что наши наблюдаемые имеют носители только в той части пространства, в которой второй кубит находится в состоянии $|0\rangle_2$. Тогда $\mathcal{H}_A = \mathcal{H}_1 \otimes |0\rangle_2$, а $\mathcal{H}_A^{\perp} = \mathcal{H}_1 \otimes |1\rangle_2$, где $\mathcal{H}_1$ — гильбертово пространство первого кубита. (Эта ситуация может показаться несколько искусственной, что я и подразумевал, говоря о немотивированности такого разложения пространства на прямую сумму.) Всякий раз, когда мы выполняем ортогональное измерение в $\mathcal{H}$, приготавливая в нем одно из множества взаимно ортогональных состояний, наш наблюдатель будет знать только о компоненте состояния, принадлежащей его пространству $\mathcal{H}_A$. Поскольку в $\mathcal{H}_A$ эти компоненты не обязательно ортогональны, он придет к выводу, что измерение готовит одно из множества неортогональных состояний.

Пусть $\{|i\rangle\}$ обозначает базис в $\mathcal{H}_A$, а $\{|\mu\rangle\}$ — базис в $\mathcal{H}_A^{\perp}$. Допустим, что начальная матрица плотности ρ_A имеет носитель в $\mathcal{H}_A$ и что мы выполняем ортогональное измерение в $\mathcal{H}$. Рассмотрим случай, когда каждый

оператор $\mathbf{E}_a$ является одномерным проектором, что, вообще говоря, достаточно для наших целей. Таким образом, $\mathbf{E}_a = |u_a\rangle\langle u_a|$, где $|u_a\rangle$ — нормированный вектор в $\mathcal{H}$. Этот вектор имеет единственное ортогональное разложение

$$|u_a\rangle = |\tilde{\psi}_a\rangle + |\tilde{\psi}_a^{\perp}\rangle, \tag{3.18}$$

где $|\tilde{\psi}_a\rangle$ и $|\tilde{\psi}_a^{\perp}\rangle$ — (ненормированные) векторы из $\mathcal{H}_A$ и $\mathcal{H}_A^{\perp}$ соответственно. После измерения новой матрицей плотности будет $|u_a\rangle\langle u_a|$ с вероятностью $\langle u_a|\boldsymbol{\rho}_A|u_a\rangle = \langle\tilde{\psi}_a|\boldsymbol{\rho}_A|\tilde{\psi}_a\rangle$ (поскольку $\boldsymbol{\rho}_A$ не имеет носителя в $\mathcal{H}_A^{\perp}$).

Но для нашего наблюдателя, не подозревающего о существовании $\mathcal{H}_A^{\perp}$, нет физической разницы между $|u_a\rangle$ и $|\tilde{\psi}_a\rangle$ (за исключением нормировки). Если записать $|\tilde{\psi}_a\rangle = \sqrt{\lambda_a}|\psi_a\rangle$, где $|\psi_a\rangle$ — нормированное состояние, тогда вдобавок к этому можно сказать, что для ограниченного пространством $\mathcal{H}_A$ наблюдателя с вероятностью $\langle\psi_a|\boldsymbol{\rho}_A|\tilde{\psi}_a\rangle$ результатом измерения является $|\psi_a\rangle\langle\psi_a|$.

Определим оператор

$$\mathbf{F}_a = \mathbf{E}_A\mathbf{E}_a\mathbf{E}_A = |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a| = \lambda_a|\psi_a\rangle\langle\psi_a| \tag{3.19}$$

(где $\mathbf{E}_A$ — ортогональный проектор, проецирующий $\mathcal{H}$ на $\mathcal{H}_A$). Тогда мы можем сказать, что результат a имеет вероятность $\operatorname{tr}\mathbf{F}_a\boldsymbol{\rho}_A$. Очевидно, что каждый оператор $\mathbf{F}_a$ эрмитов и неотрицателен, но $\mathbf{F}_a$ не является проектором, за исключением случая, когда $\lambda_a = 1$. Более того

$$\sum_a \mathbf{F}_a = \mathbf{E}_A\left(\sum_a \mathbf{E}_a\right)\mathbf{E}_A = \mathbf{E}_A = \mathbf{1}_A\,; \tag{3.20}$$

сумма всех $\mathbf{F}_a$ равна единичному оператору в $\mathcal{H}_A$.

Разложение единицы на сумму неотрицательных операторов называется *положительной операторно-значной мерой* (ПОЗМ)[1]. (Термин мера несколько неуклюж в нашем конечномерном случае; он более уместен, когда индекс a может меняться непрерывным образом). В этом обсуждении мы пришли к частному случаю ПОЗМ, построенной из одномерных операторов (операторов с одним отличным от нуля собственным значением).

[1]Строгое определение ПОЗМ или, как ее чаще называют в русской литературе, разложения единицы в гильбертовом пространстве напоминает определение вероятностной меры (имеющей, однако, не числовые, а операторные значения). См. в книге: А. С. Холево. *Вероятностные и статистические аспекты квантовой теории*, Москва–Ижевск, ИКИ (2003). *Прим. ред.*]

В обобщенной теории измерений каждый результат имеет вероятность, которую можно представить в виде

$$\text{Prob}(a) = \text{tr}\,\boldsymbol{\rho}_A \mathbf{F}_a. \tag{3.21}$$

Положительность $\mathbf{F}_a$ и равенство $\sum_a \mathbf{F}_a = \mathbf{1}_A$ необходимы, чтобы обеспечить положительность вероятностей и равенство единице их суммы.

Как ПОЗМ общего вида влияет на квантовое состояние? Простого общего ответа на этот чрезвычайно важный вопрос нет, но в случае (только что обсуждавшемся) ПОЗМ, построенной из одномерных операторов, когда результат $|\psi_a\rangle\langle\psi_a|$ возникает с вероятностью $\text{tr}\,\boldsymbol{\rho}_A\mathbf{F}_a$, суммирование по всем исходам дает

$$\begin{aligned}\boldsymbol{\rho}_A \to \boldsymbol{\rho}'_A &= \sum_a |\psi_a\rangle\langle\psi_a|\big(\lambda_a\langle\psi_a|\boldsymbol{\rho}_A|\psi_a\rangle\big)\\ &= \sum_a \big(\sqrt{\lambda_a}|\psi_a\rangle\langle\psi_a|\big)\boldsymbol{\rho}_A\big(\sqrt{\lambda_a}|\psi_a\rangle\langle\psi_a|\big)\\ &= \sum_a \sqrt{\mathbf{F}_a}\boldsymbol{\rho}_A\sqrt{\mathbf{F}_a},\end{aligned} \tag{3.22}$$

[что обобщает неймановское $\sum_a \mathbf{E}_a\boldsymbol{\rho}\mathbf{E}_a$ (3.15) на случай, когда $\mathbf{F}_a$ не являются проекторами]. Заметим, что $\text{tr}\,\boldsymbol{\rho}_A = \text{tr}\,\boldsymbol{\rho}'_A = 1$, поскольку $\sum_a \mathbf{F}_a = \mathbf{1}_A$.

3.1.3. Однокубитовая ПОЗМ

В качестве примера рассмотрим один кубит и предположим, что $\{\hat{n}_a\}$ – N единичных трехмерных векторов, удовлетворяющих условию

$$\sum_a \lambda_a \hat{n}_a = 0, \tag{3.23}$$

где λ_a – положительные вещественные числа $0 < \lambda_a < 1$ такие, что $\sum_a \lambda_a = 1$. Пусть

$$\mathbf{F}_a = \lambda_a\big(\mathbf{1} + \hat{n}_a\cdot\vec{\boldsymbol{\sigma}}\big) = 2\lambda_a\mathbf{E}(\hat{n}_a) \tag{3.24}$$

[где $\mathbf{E}(\hat{n}_a)$ – проектор $|\uparrow_{\hat{n}_a}\rangle\langle\uparrow_{\hat{n}_a}|$]. Тогда

$$\sum_a \mathbf{F}_a = \left(\sum_a \lambda_a\right)\mathbf{1} + \left(\sum_a \lambda_a\hat{n}_a\right)\cdot\vec{\boldsymbol{\sigma}} = \mathbf{1}, \tag{3.25}$$

следовательно, $\mathbf{F}_a$ определяют ПОЗМ.

В случае $N = 2$ имеем $\hat{n}_1 + \hat{n}_2 = 0$,[1] следовательно, ПОЗМ является ортогональным измерением вдоль оси $\hat{n}_1$. При $N = 3$ в симметричном случае $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$ имеем $\hat{n}_1 + \hat{n}_2 + \hat{n}_3 = 0$ и

$$\mathbf{F}_a = \frac{1}{3}(\mathbf{1} + \hat{n}_a \cdot \vec{\sigma}) = \frac{2}{3}\mathbf{E}(\hat{n}_a). \tag{3.26}$$

3.1.4. Теорема Наймарка

Мы пришли к понятию ПОЗМ, рассматривая ортогональные измерения в более широком, чем $\mathcal{H}_A$, пространстве. Теперь обратим наши рассуждения и покажем, что таким образом может быть реализована любая ПОЗМ.

Рассмотрим произвольную ПОЗМ с n одномерными положительными операторами $\mathbf{F}_a$, удовлетворяющими условию $\sum_{a=1}^{n} \mathbf{F}_a = \mathbf{1}$. Покажем, что эту ПОЗМ всегда можно реализовать путем расширения гильбертова пространства и выполнения ортогонального измерения в этом более широком пространстве. Это утверждение называется *теоремой Наймарка* [2].

Чтобы доказать это, рассмотрим гильбертово пространство $\mathcal{H}$ с $\dim \mathcal{H} = N$ и ПОЗМ $\{\mathbf{F}_a\}$, $a = 1, 2, \ldots, n$ с $n \geqslant N$. Каждый одномерный положительный оператор может быть записан в виде

$$\mathbf{F}_a = |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a|, \tag{3.27}$$

где вектор $|\tilde{\psi}_a\rangle$ не нормирован. Выписывая в явном виде матричные элементы равенства $\sum_{a=1}^{n} \mathbf{F}_a = \mathbf{1}$, получим:

$$\sum_{a=1}^{n} (\mathbf{F}_a)_{ij} = \sum_{a=1}^{n} \tilde{\psi}_{ai}^{*} \tilde{\psi}_{aj} = \delta_{ij}. \tag{3.28}$$

Теперь изменим точку зрения на уравнение (3.28). Будем интерпретировать $(\tilde{\psi}_a)_i$ не как $n \geqslant N$ векторов в N-мерном пространстве, а как $N \leqslant n$

[1] Конечно, это равенство справедливо лишь в симметричном случае $\lambda_1 = \lambda_2 = \frac{1}{2}$, который здесь по-видимому подразумевается. — *Прим. ред.*

[2] Обсуждение ПОЗМ и теоремы Наймарка можно найти в книге: A. Peres. *Quantum theory: Concepts and Methods*, Kluwer Academic Publishers, New York et al (2002) [На русском языке см.: Н.И. Ахиезер, И.М. Глазман, *Теория операторов в гильбертовом пространстве*, Наука, М.: (1966). — *Прим. ред.*]

векторов $\left(\tilde{\psi}_i^T\right)_a$ в n-мерном пространстве. Тогда уравнение (3.28) утверждает, что эти N векторов образуют ортогональный набор. Естественно, что он может быть расширен до ортогонального базиса в n-мерном пространстве. Другими словами, существует $n \times n$-матрица u_{ai} с $u_{ai} = \tilde{\psi}_{ai}$ при $a = 1, 2, \ldots, N$ такая, что

$$\sum_a u_{ai}^* u_{aj} = \delta_{ij}, \tag{3.29}$$

или, в матричной форме, $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$. Отсюда следует, что $\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$, поскольку для любого вектора $|\psi\rangle$

$$\mathbf{U}(\mathbf{U}^\dagger\mathbf{U})|\psi\rangle = (\mathbf{U}\mathbf{U}^\dagger)\mathbf{U}|\psi\rangle = \mathbf{U}|\psi\rangle, \tag{3.30}$$

а областью действия $\mathbf{U}$ (по крайней мере для конечномерных матриц) является все n-мерное пространство. Возвращаясь к компонентной записи, имеем

$$\sum_j u_{aj} u_{bj}^* = \delta_{ab}. \tag{3.31}$$

Следовательно, $\left(u_a\right)_i$ образуют полный ортонормированный набор векторов[1].

Пусть теперь в пространстве размерности $n \geqslant N$ выполняется ортогональное измерение:

$$\mathbf{E}_a = |u_a\rangle\langle u_a|. \tag{3.32}$$

Мы построили векторы $|u_a\rangle$ так, чтобы каждый из них имел ортогональное разложение

$$|u_a\rangle = |\tilde{\psi}_a\rangle + |\tilde{\psi}_a^\perp\rangle, \tag{3.33}$$

где $|\tilde{\psi}_a\rangle \in \mathcal{H}$, а $|\tilde{\psi}_a^\perp\rangle \in \mathcal{H}^\perp$. Тогда ортогональным проецированием этого базиса на $\mathcal{H}$ мы воспроизводим ПОЗМ $\{\mathbf{F}_a\}$. Этим завершается доказательство теоремы Наймарка.

Чтобы проиллюстрировать теорему Наймарка в действии, рассмотрим еще раз однокубитовую ПОЗМ

$$\mathbf{F}_a = \frac{2}{3}|\uparrow_{\hat{n}_a}\rangle\langle\uparrow_{\hat{n}_a}|, \tag{3.34}$$

$a = 1, 2, 3$, где $\hat{n}_1 + \hat{n}_2 + \hat{n}_3 = 0$. Согласно теореме ПОЗМ может быть реализована как ортогональное измерение «кутрита» — квантовой системы в трехмерном гильбертовом пространстве.

[1] Другими словами, мы показали, что если строки $n \times n$-матрицы ортонормированы, то таковыми же будут их столбцы.

Пусть $\hat{n}_1 = (0,0,1)$, $\hat{n}_2 = (\sqrt{3}/2, 0, -1/2)$, $\hat{n}_3 = (-\sqrt{3}/2, 0, -1/2)$, тогда, вспоминая что

$$|\theta, \varphi = 0\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}, \tag{3.35}$$

мы можем записать три вектора $|\tilde{\psi}_a\rangle = \sqrt{2/3}|\theta, \varphi = 0\rangle$ (где $\theta_1, \theta_2, \theta_3 = 0, 2\pi/3, 4\pi/3$) как

$$|\tilde{\psi}_1\rangle, |\tilde{\psi}_2\rangle, |\tilde{\psi}_3\rangle = \begin{pmatrix} \sqrt{2/3} \\ 0 \end{pmatrix}, \begin{pmatrix} \sqrt{1/6} \\ \sqrt{1/2} \end{pmatrix}, \begin{pmatrix} -\sqrt{1/6} \\ \sqrt{1/2} \end{pmatrix}. \tag{3.36}$$

Теперь мы можем интерпретировать эти три двумерных вектора, как 2×3-матрицу, а теорема Наймарка гарантирует, что две ее строки ортонормированы. Следовательно, мы можем добавить еще одну ортонормированную строку:

$$|u_1\rangle, |u_2\rangle, |u_3\rangle = \begin{pmatrix} \sqrt{2/3} \\ 0 \\ \sqrt{1/3} \end{pmatrix}, \begin{pmatrix} \sqrt{1/6} \\ \sqrt{1/2} \\ -\sqrt{1/3} \end{pmatrix}, \begin{pmatrix} -\sqrt{1/6} \\ \sqrt{1/2} \\ \sqrt{1/3} \end{pmatrix}. \tag{3.37}$$

Мы видим, что (как и утверждает теорема) столбцы $|u_a\rangle$ также ортонормированы. Если мы выполним ортогональное измерение в базисе векторов $|u_a\rangle$, то живущий в двумерном подпространстве наблюдатель придет к выводу, что мы реализовали ПОЗМ $\{\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3\}$. Мы показали, что если наш кубит фактически является двумя компонентами *кутрита*, то ПОЗМ может быть реализована как ортогональное измерение этого кутрита.

3.1.5. Ортогональное измерение на тензорном произведении

Тем не менее типичный кубит не скрывает никакого секрета. Чтобы выполнить обобщенное измерение, нам необходимо запастись дополнительными кубитами и выполнить совместные ортогональные измерения на нескольких кубитах сразу.

Рассмотрим случай двух (изолированных) систем A и B, описываемых тензорным произведением $\mathcal{H}_A \otimes \mathcal{H}_B$. Предположим, что на этом тензорном произведении мы выполняем ортогональное измерение, характеризуемое полным набором взаимно ортогональных проекторов $\mathbf{E}_a$:

$$\sum_a \mathbf{E}_a = \mathbf{1}. \tag{3.38}$$

Представим, что начальным состоянием квантовой системы является «некоррелированное» тензорное произведение состояний

$$\rho_{AB} = \rho_A \otimes \rho_B. \tag{3.39}$$

Тогда с вероятностью

$$\mathrm{Prob}(a) = \mathrm{tr}_{AB}\left[\mathbf{E}_a(\rho_A \otimes \rho_B)\right] \tag{3.40}$$

результатом измерения будет a. В этом случае новая матрица плотности будет равна[1]

$$\rho'_{AB} = \frac{\mathbf{E}_a(\rho_A \otimes \rho_B)\mathbf{E}_a}{\mathrm{tr}_{AB}\left[\mathbf{E}_a(\rho_A \otimes \rho_B)\right]}. \tag{3.41}$$

Для наблюдателя, имеющего доступ только к системе A, ее новую матрицу плотности дает частичный след только что записанной матрицы плотности, или

$$\rho'_A = \frac{\mathrm{tr}_B\left[\mathbf{E}_a(\rho_A \otimes \rho_B)\mathbf{E}_a\right]}{\mathrm{tr}_{AB}\left[\mathbf{E}_a(\rho_A \otimes \rho_B)\right]}. \tag{3.42}$$

Выражение (3.40) для вероятности результата a также может быть записано в виде

$$\mathrm{Prob}(a) = \mathrm{tr}_A\left[\mathrm{tr}_B\left(\mathbf{E}_a(\rho_A \otimes \rho_B)\right)\right] = \mathrm{tr}_A\left(\mathbf{F}_a\rho_A\right); \tag{3.43}$$

если ввести ортогональные базисы $\{|i\rangle_A\}$ в $\mathcal{H}_A$, а $\{|\mu\rangle_B\}$ в $\mathcal{H}_B$, то

$$\sum_{ij\mu\nu}\left(\mathbf{E}_a\right)_{j\nu,i\mu}(\rho_A)_{ij}(\rho_B)_{\mu\nu} = \sum_{ij}\left(\mathbf{F}_a\right)_{ji}(\rho_A)_{ij} \tag{3.44}$$

или

$$\left(\mathbf{F}_a\right)_{ji} = \sum_{\mu\nu}\left(\mathbf{E}_a\right)_{j\nu,i\mu}(\rho_B)_{\mu\nu}. \tag{3.45}$$

Из уравнения (3.45) следует, что каждый оператор обладает $\mathbf{F}_a$ свойствами:

(1) Эрмитовость:

$$\left(\mathbf{F}_a\right)^*_{ij} = \sum_{\mu\nu}\left(\mathbf{E}_a\right)^*_{i\nu,j\mu}(\rho_B)^*_{\mu\nu} = \sum_{\mu\nu}\left(\mathbf{E}_a\right)_{j\mu,i\nu}(\rho_B)_{\nu\mu} = \left(\mathbf{F}_a\right)_{ji},$$

поскольку эрмитовы $\mathbf{E}_a$ и ρ_B.

[1] Такой вид матрица плотности будет иметь после измерения, результат которого *зафиксирован* и имеет значение a. Если же результат не был «записан», то при таком измерении матрица плотности эволюционирует в соответствии с уравнением (3.15). — *Прим. ред.*

(2) Положительность: В базисе, диагонализующем $\rho_B = \sum_\mu p_\mu |\mu\rangle_{B\,B}\langle\mu|$,

$$ {}_A\langle\psi|\mathbf{F}_a|\psi\rangle_A = \sum_\mu p_\mu ({}_A\langle\psi| \otimes {}_B\langle\mu|)\mathbf{E}_a(|\psi\rangle_A \otimes |\mu\rangle_B) \geqslant 0, $$

поскольку положителен $\mathbf{E}_a$.

(3) Полнота:

$$ \sum_a \mathbf{F}_a = \sum_\mu p_\mu \left\langle \mu \middle| \sum_a \mathbf{E}_a \middle| \mu \right\rangle_B = \mathbf{1}_A, $$

поскольку $\sum_a \mathbf{E}_a = \mathbf{1}_{AB}$, а $\operatorname{tr}\rho_B = 1$.

Однако операторы $\mathbf{F}_a$ не обязательно взаимно ортогональны. Фактически количество $\mathbf{F}_a$ ограничено только размерностью $\mathcal{H}_A \otimes \mathcal{H}_B$, большей (и, возможно, гораздо большей) чем размерность $\mathcal{H}_A$.

В общем случае нет простого способа выразить конечную матрицу плотности ρ'_A через ρ_A и $\mathbf{F}_a$. Не будем, однако, обращать внимание на то, как ПОЗМ изменяет матрицу плотности, а вместо этого зададимся вопросом. Допустим, что $\mathcal{H}_A$ имеет размерность N, и рассмотрим ПОЗМ, состоящую из n одномерных неотрицательных операторов $\mathbf{F}_a$, удовлетворяющих условию $\sum_{a=1}^{n} \mathbf{F}_a = \mathbf{1}_A$. Можем ли мы выбрать пространство $\mathcal{H}_B$, матрицу плотности ρ_B в $\mathcal{H}_B$ и проекционные операторы $\mathbf{E}_a$ в $\mathcal{H}_A \otimes \mathcal{H}_B$ (где количество $\mathbf{E}_a$ может превосходить количество $\mathbf{F}_a$) такие, чтобы вероятность результата a ортогонального измерения удовлетворяла условию[1]

$$ \operatorname{tr}\mathbf{E}_a(\rho_A \otimes \rho_B) = \operatorname{tr}(\mathbf{F}_a\rho_A)\,? \tag{3.46} $$

(Неважно, как ортогональная проекция модифицирует ρ_A!). Мы будем считать это «реализацией» ПОЗМ с помощью ортогонального измерения, поскольку нас не интересует, какова ρ'_A для каждого результата измерения; мы только требуем, чтобы *вероятности* результатов согласовались с определенной таким образом ПОЗМ.

Такая реализация ПОЗМ действительно возможна; чтобы показать это, еще раз обратимся к теореме Наймарка. Каждый одномерный оператор $\mathbf{F}_a$,

[1] Если количество $\mathbf{E}_a$ больше, чем $\mathbf{F}_a$, то почти все n результатов имеют вероятность, *равную нулю*.

$a = 1, 2, \ldots, n$, может быть представлен в виде $\mathbf{F}_a = |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a|$. Согласно Наймарку существуют n ортонормированных n-компонентных векторов $|u_a\rangle$ таких, что

$$|u_a\rangle = |\tilde{\psi}_a\rangle + |\tilde{\psi}_a^{\perp}\rangle. \tag{3.47}$$

Начнем с того, что рассмотрим частный случай $n = rN$, где r — положительное целое число. Тогда удобно разложить $|\tilde{\psi}_a^{\perp}\rangle$ на прямую сумму N-компонентных векторов

$$|\tilde{\psi}_a^{\perp}\rangle = |\tilde{\psi}_{1,a}^{\perp}\rangle \oplus |\tilde{\psi}_{2,a}^{\perp}\rangle \oplus \cdots \oplus |\tilde{\psi}_{r-1,a}^{\perp}\rangle. \tag{3.48}$$

Здесь $|\tilde{\psi}_{1,a}^{\perp}\rangle$ обозначает первые N компонент вектора $|\tilde{\psi}_a^{\perp}\rangle$, $|\tilde{\psi}_{2,a}^{\perp}\rangle$ обозначает следующие N компонент и так далее. Тогда ортонормированность векторов $|u_a\rangle$ означает, что

$$\delta_{ab} = \langle u_a|u_b\rangle = \langle\tilde{\psi}_a|\tilde{\psi}_b\rangle + \sum_{\mu=1}^{r-1}\langle\tilde{\psi}_{\mu,a}^{\perp}|\tilde{\psi}_{\mu,b}^{\perp}\rangle. \tag{3.49}$$

Выберем теперь $\mathcal{H}_B$ имеющим размерность r и обозначим ортонормированный базис в $\mathcal{H}_B$:

$$\{|\mu\rangle_B\}, \qquad \mu = 0, 1, 2, \ldots, r-1. \tag{3.50}$$

Тогда из уравнения (3.49) следует, что

$$|\Phi_a\rangle_{AB} = |\tilde{\psi}_a\rangle_A|0\rangle_B + \sum_{\mu=1}^{r-1}|\tilde{\psi}_{\mu,a}^{\perp}\rangle_A|\mu\rangle_B, \qquad a = 1, 2, \ldots, n, \tag{3.51}$$

– ортонормированный базис в $\mathcal{H}_A \otimes \mathcal{H}_B$.

Предположим теперь, что состоянием в $\mathcal{H}_A \otimes \mathcal{H}_B$ является

$$\boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes |0\rangle_{B\,B}\langle 0|, \tag{3.52}$$

и мы выполняем ортогональное проецирование на базис $\{|\Phi_a\rangle_{AB}\}$ в $\mathcal{H}_A \otimes \mathcal{H}_B$. Тогда, поскольку при $\mu \neq 0$ ${}_B\langle 0|\mu\rangle_B = 0$, результат $|\Phi_a\rangle_{AB}$ появляется с вероятностью

$${}_{AB}\langle\Phi_a|\boldsymbol{\rho}_{AB}|\Phi_a\rangle_{AB} = {}_A\langle\tilde{\psi}_a|\boldsymbol{\rho}_A|\tilde{\psi}_a\rangle_A \tag{3.53}$$

и, следовательно,

$${}_{AB}\langle\Phi_a|\boldsymbol{\rho}_{AB}|\Phi_a\rangle_{AB} = \mathrm{tr}\,(\mathbf{F}_a\boldsymbol{\rho}_A). \tag{3.54}$$

Мы действительно успешно «реализовали» ПОЗМ, выполняя ортогональное измерение в $\mathcal{H}_A \otimes \mathcal{H}_B$. Эта конструкция так же эффективна, как и описанная выше конструкция «прямой суммы»; мы выполнили ортогональное измерение в пространстве размерности $n = rN$.

Если появился результат a, тогда с помощью измерения приготовлено состояние

$$\rho'_{AB} = |\Phi_a\rangle_{AB\,AB}\langle\Phi_a|. \tag{3.55}$$

Матрица плотности, видимая наблюдателю, которому доступна только система A, получается взятием частичного следа по $\mathcal{H}_B$:

$$\begin{aligned}\rho'_A &= \mathrm{tr}_B\left(|\Phi_a\rangle_{AB\,AB}\langle\Phi_a|\right) = \\ &= |\tilde{\psi}_a\rangle_{A\,A}\langle\tilde{\psi}_a| + \sum_{\mu=1}^{r-1} |\tilde{\psi}^{\perp}_{\mu,a}\rangle_{A\,A}\langle\tilde{\psi}^{\perp}_{\mu,a}|,\end{aligned} \tag{3.56}$$

что не совсем то же самое, что было получено в нашей конструкции «прямой суммы». Во всяком случае существует множество способов реализовать ПОЗМ с помощью ортогональных измерений, и уравнение (3.56) применимо только к выбранной здесь частной конструкции.

Тем не менее в действительности эта конструкция идеально подходит для реализации ПОЗМ, в которой состояние $|\psi_a\rangle_{A\,A}\langle\psi_a|$ приготавливается в результате появления исхода a. Трудным моментом осуществления ПОЗМ является обеспечение того, что результат a появляется с требуемой вероятностью. После этого уже легко прийти к соглашению о том, что *следствием* появления результата a является состояние $|\psi_a\rangle_{A\,A}\langle\psi_a|$; если угодно, сразу как только измерение выполнено и результат a получен, мы можем просто отбросить ρ_A и приступить к приготовлению требуемого состояния! Фактически, в случае проекции на базис $|\Phi_a\rangle_{AB}$ мы можем полностью построить ПОЗМ, проецируя систему B на базис $\{|\mu\rangle_B\}$ и сообщая результат в систему A. Если результатом является $|0\rangle_B$, тогда не нужно предпринимать никаких действий. Если же результатом является $|\mu\rangle_B$, $\mu > 0$, тогда было приготовлено состояние $|\tilde{\psi}^{\perp}_{\mu,a}\rangle_A$, которое затем может быть преобразовано в $|\psi_a\rangle_A$.

До сих пор мы обсуждали частный случай $n = r \cdot N$. Если в действительности $n = r \cdot N - c$, $0 < c < N$, то нам нужно лишь выбрать равными нулю последние c компонент вектора $|\tilde{\psi}^{\perp}_{r-1,a}\rangle_A$ и состояния $|\Phi\rangle_{AB}$ по-прежнему будут взаимно ортогональными. Чтобы получить полный базис, мы можем добавить c состояний:

$$|e_i\rangle_A|r-1\rangle_B, \qquad i = rN - c + 1, rN - c + 2, \ldots, rN; \tag{3.57}$$

здесь $|e_i\rangle_A$ — вектор, у которого отлична от нуля только одна i-ая компонента, так что $|e_i\rangle_A$ гарантированно ортогонален вектору $|\tilde{\psi}^{\perp}_{r-1,a}\rangle_A$. В этом случае ПОЗМ реализуется как ортогональное измерение в пространстве размерности $rN = n + c$.

В качестве примера конструкции тензорного произведения мы вновь можем рассмотреть однокубитовую ПОЗМ с

$$\mathbf{F}_a = \frac{2}{3}|\uparrow_{\hat{n}_a}\rangle_A\,{}_A\langle\uparrow_{\hat{n}_a}|, \qquad a = 1,2,3. \tag{3.58}$$

Мы можем реализовать эту ПОЗМ, вводя второй кубит B. В двухкубитовом гильбертовом пространстве мы можем проецировать на ортонормированный базис [1]

$$\begin{aligned} |\Phi_a\rangle &= \sqrt{\frac{2}{3}}|\uparrow_{\hat{n}_a}\rangle_A|0\rangle_B + \sqrt{\frac{1}{3}}|0\rangle_A|1\rangle_B, \quad a = 1,2,3, \\ |\Phi_0\rangle &= |1\rangle_A|1\rangle_B. \end{aligned} \tag{3.59}$$

Если начальным состоянием является $\boldsymbol{\rho}_{AB} = \boldsymbol{\rho}_A \otimes |0\rangle_B\,{}_B\langle 0|$, то мы имеем

$$\langle\Phi_a|\boldsymbol{\rho}_{AB}|\Phi_a\rangle = \frac{2}{3}\,{}_A\langle\uparrow_{\hat{n}_a}|\boldsymbol{\rho}_A|\uparrow_{\hat{n}_a}\rangle_A, \tag{3.60}$$

следовательно, эта проекция осуществляет ПОЗМ в $\mathcal{H}_A$. (Здесь мы выполнили ортогональные измерения в четырехмерном пространстве; в предыдущей конструкции «прямой суммы» мы нуждались только в трех измерениях).

3.1.6. ЖХЙВ с ПОЗМ

Обсуждая теорему ЖХЙВ, мы говорили, что, приготовив состояние

$$|\Phi_a\rangle_{AB} = \sum_{\mu} \sqrt{q_\mu}|\psi_\mu\rangle_A|\beta_\mu\rangle_B, \tag{3.61}$$

[1] Здесь фаза $|\tilde{\psi}_2\rangle = \sqrt{2/3}|\uparrow_{\hat{n}_a}\rangle_A$ отличается на -1 от соответствующей фазы в уравнении (3.36); она выбрана таким образом, чтобы ${}_A\langle\uparrow_{\hat{n}_a}|\uparrow_{\hat{n}_b}\rangle_A = -1/2$ при $a \neq b$. Мы сделали этот выбор затем, чтобы коэффициент перед $|0\rangle_A|1\rangle_B$ был положителен во всех трех $|\Phi_1\rangle$, $|\Phi_2\rangle$, $|\Phi_3\rangle$.

мы можем реализовать ансамбль

$$\boldsymbol{\rho}_A = \sum_\mu q_\mu |\psi_\mu\rangle_{A\,A}\langle\psi_\mu|, \tag{3.62}$$

выполняя ортогональные измерения в $\mathcal{H}_B$. Более того, если $\dim \mathcal{H}_B = n$, то, измеряя подходящие наблюдаемые в $\mathcal{H}_B$, мы можем с этим одним чистым состоянием $|\Phi_a\rangle_{AB}$ реализовать любое приготовление $\boldsymbol{\rho}_A$ как ансамбля, содержащего вплоть до n чистых состояний.

Теперь можно видеть, что если мы готовы допустить в $\mathcal{H}_B$ ПОЗМы, а не только ортогональные измерения, то даже при $\dim \mathcal{H}_B = N$ можно реализовать любое приготовление $\boldsymbol{\rho}_A$ с помощью подходящего выбора ПОЗМ в $\mathcal{H}_B$. Суть в том, что $\boldsymbol{\rho}_B$ имеет носитель в пространстве размерности самое большее N. Следовательно, можно переписать

$$|\Phi_a\rangle_{AB} = \sum_\mu \sqrt{q_\mu}|\psi_\mu\rangle_A|\tilde{\beta}_\mu\rangle_B, \tag{3.63}$$

где $|\tilde{\beta}_\mu\rangle_B$ — ортогональная проекция вектора $|\beta_\mu\rangle_B$ на носитель матрицы плотности $\boldsymbol{\rho}_B$. Мы можем выполнить ПОЗМ на носителе $\boldsymbol{\rho}_B$ с $\mathbf{F}_a$ — $|\tilde{\beta}_\mu\rangle_{B\,B}\langle\tilde{\beta}_\mu|$ и таким образом приготовить состояние $|\psi_\mu\rangle_A$ с вероятностью q_μ.

3.2. Супероператоры

3.2.1. Представление операторной суммы

Перейдем к следующему этапу нашей программы понимания поведения части бинарной системы. Мы видели, что чистое состояние бинарной системы может вести себя подобно смешанному состоянию, когда мы наблюдаем только одну ее подсистему A, а ортогональное измерение бинарной системы внутри ее подсистемы A может быть (неортогональной) ПОЗМ. Зададимся вопросом: если состояние бинарной системы совершает унитарную эволюцию, то как тогда описать эволюцию одной только ее подсистемы A?

Пусть начальная матрица плотности бинарной системы представляет собой тензорное произведение состояний вида

$$\boldsymbol{\rho}_A \otimes |0\rangle_{B\,B}\langle 0|; \tag{3.64}$$

система A имеет матрицу плотности $\boldsymbol{\rho}_A$, а система B предполагается находящейся в чистом состоянии, которое мы обозначили $|0\rangle_B$. Бинарная система эволюционирует в течение конечного промежутка времени, управляемая унитарным оператором эволюции

$$\mathbf{U}_{AB}(\boldsymbol{\rho}_A \otimes |0\rangle_{B\,B}\langle 0|)\mathbf{U}_{AB}. \tag{3.65}$$

Выполним вычисление частичного следа в $\mathcal{H}_B$, чтобы найти конечную матрицу плотности системы A:

$$\begin{aligned}\boldsymbol{\rho}'_A &= \mathrm{tr}_B\left(\mathbf{U}_{AB}(\boldsymbol{\rho}_A \otimes |0\rangle_{B\,B}\langle 0|)\mathbf{U}^\dagger_{AB}\right)\\ &= \sum_\mu {}_B\langle\mu|\mathbf{U}_{AB}|0\rangle_B\,\boldsymbol{\rho}_A\,{}_B\langle 0|\mathbf{U}^\dagger_{AB}|\mu\rangle_B,\end{aligned} \tag{3.66}$$

где $\{|\mu\rangle_B\}$ — ортонормированный базис в $\mathcal{H}_B$, а ${}_B\langle\mu|\mathbf{U}_{AB}|0\rangle_B$ — оператор, действующий в $\mathcal{H}_A$. [Если $\{|i\rangle_A \otimes |\mu\rangle_B\}$ — ортонормированный базис в $\mathcal{H}_A \otimes \mathcal{H}_B$, то ${}_B\langle\mu|\mathbf{U}_{AB}|\nu\rangle_B$ обозначает оператор, матричные элементы которого равны

$${}_A\langle i|{}_B\langle\mu|\mathbf{U}_{AB}|\nu\rangle_B|j\rangle_A = ({}_A\langle i| \otimes {}_B\langle\mu|)\mathbf{U}_{AB}(|\nu\rangle_B \otimes |j\rangle_A).] \tag{3.67}$$

Если обозначить

$$\mathbf{M}_\mu = {}_B\langle\mu|\mathbf{U}_{AB}|0\rangle_B, \tag{3.68}$$

то $\boldsymbol{\rho}'_A$ можно представить в виде

$$\$(\boldsymbol{\rho}_A) \equiv \boldsymbol{\rho}'_A = \sum_\mu \mathbf{M}_\mu \boldsymbol{\rho}_A \mathbf{M}^\dagger_\mu. \tag{3.69}$$

Из унитарности $\mathbf{U}_{AB}$ следует, что $\mathbf{M}_\mu$ обладает свойством

$$\begin{aligned}\sum_\mu \mathbf{M}^\dagger_\mu \mathbf{M}_\mu &= \sum_\mu {}_B\langle 0|\mathbf{U}^\dagger_{AB}|\mu\rangle_{B\,B}\langle\mu|\mathbf{U}_{AB}|0\rangle_B =\\ &= {}_B\langle 0|\mathbf{U}^\dagger_{AB}\mathbf{U}_{AB}|0\rangle_B = \mathbf{1}_A.\end{aligned} \tag{3.70}$$

Уравнение (3.69) определяет линейное отображение \$, преобразующее один линейный оператор в другой. Если выполняется свойство (3.70), то такое отображение называется *супероператором*, а уравнение (3.69) назы-

вается представлением супероператора операторной суммой (или представлением *Крауса*). Супероператор можно рассматривать как линейное отображение, преобразующее операторы плотности в операторы плотности, поскольку из (3.69) и (3.70) следует, что $\boldsymbol{\rho}'_A$ — оператор плотности, если им является $\boldsymbol{\rho}_A$:

(1) $\boldsymbol{\rho}'_A$ эрмитов: $\boldsymbol{\rho}'^{\dagger}_A = \sum_\mu \mathbf{M}_\mu \boldsymbol{\rho}^\dagger_A \mathbf{M}^\dagger_\mu = \sum_\mu \mathbf{M}_\mu \boldsymbol{\rho}_A \mathbf{M}^\dagger_\mu = \boldsymbol{\rho}'_A$.

(2) $\boldsymbol{\rho}'_A$ имеет единичный след: $\operatorname{tr} \boldsymbol{\rho}'_A = \sum_\mu \operatorname{tr}\left(\boldsymbol{\rho}_A \mathbf{M}^\dagger_\mu \mathbf{M}_\mu\right) = \operatorname{tr} \boldsymbol{\rho}_A = 1$.

(3) $\boldsymbol{\rho}'_A$ положительно определен:

$$ {}_A\langle\psi|\boldsymbol{\rho}'_A|\psi\rangle_A = \sum_\mu \left({}_A\langle\psi|\mathbf{M}_\mu\right)\boldsymbol{\rho}_A\left(\mathbf{M}^\dagger_\mu|\psi\rangle_A\right) \geqslant 0. $$

Мы показали, что представление операторной суммы (3.69) следует из «унитарного представления» (3.66). Более того, по данному представлению супероператора в виде операторной суммы всегда можно построить соответствующее унитарное представление. Выберем в качестве $\mathcal{H}_B$ гильбертово пространство, размерность которого, по крайней мере, больше числа слагаемых в операторной сумме. Если $|\varphi\rangle_A$ — произвольный вектор в $\mathcal{H}_A$, $\{|\mu\rangle_B\}$ — ортонормированный базис в $\mathcal{H}_B$, а $|0\rangle_B$ — некоторое нормированное состояние в $\mathcal{H}_B$, то определим действие $\mathbf{U}_{AB}$ соотношением

$$ \mathbf{U}_{AB}\left(|\varphi\rangle_A \otimes |0\rangle_B\right) = \sum_\mu \mathbf{M}_\mu |\varphi\rangle_A \otimes |\mu\rangle_B. \tag{3.71} $$

Это действие сохраняет внутреннее произведение

$$ \left(\sum_\nu {}_A\langle\varphi_2|\mathbf{M}^\dagger_\nu \otimes {}_B\langle\nu|\right)\left(\sum_\mu \mathbf{M}_\mu|\varphi_1\rangle_A \otimes |\mu\rangle_B\right) = $$

$$ = {}_A\left\langle \varphi_2 \left| \sum_\mu \mathbf{M}^\dagger_\mu \mathbf{M}_\mu \right| \varphi_1 \right\rangle_A = \langle\varphi_2|\varphi_1\rangle_A, \tag{3.72} $$

следовательно, $\mathbf{U}_{AB}$ может быть расширен до унитарного оператора, действующего на всем $\mathcal{H}_A \otimes \mathcal{H}_B$. Взяв частичный след, мы найдем

$$ \operatorname{tr}_B\left(\mathbf{U}_{AB}\left(|\varphi\rangle_A \otimes |0\rangle_B\right)\left({}_A\langle\varphi| \otimes {}_B\langle 0|\right)\mathbf{U}^\dagger_{AB}\right) = \sum_\mu \mathbf{M}_\mu \left(|\varphi\rangle_A\, {}_A\langle\varphi|\right)\mathbf{M}^\dagger_\mu. \tag{3.73} $$

Поскольку любая матрица плотности ρ_A может быть представлена как ансамбль чистых состояний, мы воспроизводим представление операторной суммы, действующей на произвольную ρ_A.

Очевидно, что представление операторной суммы данного супероператора \$ не единственно. Мы можем вычислить частичный след в любом базисе, в каком пожелаем. Если мы используем базис $\left\{ {}_B\langle\nu| = \sum_\mu \mathbf{U}_{\nu\mu}\, {}_B\langle\mu| \right\}$, то получим представление

$$\$(\rho_A) = \sum_\nu \mathbf{N}_\nu \rho_A \mathbf{N}_\nu^\dagger, \tag{3.74}$$

где $\mathbf{N}_\nu = \mathbf{U}_{\nu\mu}\mathbf{M}_\mu$. Вскоре мы увидим, что так связаны *любые* два представления операторных сумм одного супероператора.

Супероператоры важны, поскольку они обеспечивают нас формализмом для обсуждения общей теории *декогерентизации*, эволюции чистых состояний в смешанные. *Унитарная* эволюция ρ_A является частным случаем, когда в операторной сумме имеется только одно слагаемое. Если в ней присутствуют два или более слагаемых, тогда в ходе эволюции, управляемой оператором $\mathbf{U}_{AB}$, чистые начальные состояния из $\mathcal{H}_A$ *запутываются* с $\mathcal{H}_B$. То есть если возникающие в операторной сумме операторы $\mathbf{M}_1$ и $\mathbf{M}_2$ линейно независимы, то существует такой вектор $|\varphi\rangle_A$, что векторы $|\tilde{\varphi}_1\rangle_A = \mathbf{M}_1|\varphi\rangle_A$ и $|\tilde{\varphi}_2\rangle_A = \mathbf{M}_2|\varphi\rangle_A$ линейно независимы, следовательно, состояние $|\tilde{\varphi}_1\rangle_A|1\rangle_B + |\tilde{\varphi}_2\rangle_A|2\rangle_B + \cdots$ имеет число Шмидта больше единицы. Следовательно, чистое состояние $|\varphi\rangle_A\,{}_A\langle\varphi|$ эволюционирует к смешанному *конечному* состоянию ρ'_A.

Из двух супероператоров $\$_1$ и $\$_2$ можно построить композицию, представляющую собой другой супероператор $\$_1 \circ \$_2$; если $\$_1$ описывает эволюцию от вчерашнего дня до сегодняшнего, а $\$_2$ — от сегодняшнего дня до завтрашнего, то $\$_1 \circ \$_2$ описывает эволюцию от вчерашнего дня до завтрашнего. Но является ли обратный супероператор также супероператором? То есть существует ли супероператор, описывающий эволюцию из сегодняшнего дня во вчерашний? Вы покажете в домашнем упражнении, что на самом деле супероператор обратим только тогда, когда он унитарен.

Операторы унитарной эволюции образуют группу, а супероператоры определяют динамическую *полугруппу*. Когда возникает декогерентизация, существует стрела времени; даже на микроскопическом уровне можно говорить о различии между движением вперед и назад во времени. Декогерентизация вызывает неизбежную потерю квантовой информации – однажды вынув (мертвого) кота из ящика, мы не можем вернуть его в исходное состояние.

3.2.2. Линейность

Теперь посмотрим на эту проблему немного шире и обсудим основные свойства, которым должен удовлетворять любой «разумный» закон эволюции матрицы плотности. Мы увидим, что любой такой закон допускает представление операторной суммы, то есть, в известном смысле, выделенное нами динамическое поведение рассматриваемой части бинарной системы действительно является наиболее общим.

Отображение $\$: \rho \to \rho'$, преобразующее исходную матрицу плотности ρ в конечную ρ', представляет собой отображение операторов в операторы, обладающее следующими свойствами:

(1) $\$$ *сохраняет эрмитовость*: ρ' эрмитова, если таковой является ρ.

(2) $\$$ *сохраняет след*: $\operatorname{tr}\rho' = 1$, если $\operatorname{tr}\rho = 1$.

(3) $\$$ *положителен*: ρ' неотрицательна, если таковой является ρ.

Обычно также предполагают, что

(0) $\$$ *линейный оператор*.

В то время как условия (1), (2) и (3) действительно необходимы для того, чтобы ρ' оставалась матрицей плотности, (0) остается открытым вопросом. Почему линейность?

Возможный ответ состоит в том, что нелинейную эволюцию матрицы плотности было бы сложно согласовать с любой интерпретацией ансамбля. Если

$$\$\big[\rho(\lambda)\big] = \$\big[\lambda\rho_1 + (1-\lambda)\rho_2\big] = \lambda\$[\rho_1] + (1-\lambda)\$[\rho_2], \tag{3.75}$$

тогда временная эволюция согласуется с вероятностной интерпретацией $\rho(\lambda)$: или (с вероятностью λ) было приготовлено начальное состояние ρ_1, которое эволюционировало в состояние $\$[\rho_1]$, или (с вероятностью $1-\lambda$) было приготовлено начальное состояние ρ_2, которое эволюционировало в состояние $\$[\rho_2]$. Нелинейный супероператор $\$$, по-видимому, ведет к парадоксальным следствиям.

В качестве примера рассмотрим один кубит, эволюционирующий согласно

$$\$(\rho) = \exp\big[i\pi\sigma_1 \operatorname{tr}(\sigma_1\rho)\big]\rho\exp\big[-i\pi\sigma_1\operatorname{tr}(\sigma_1\rho)\big]. \tag{3.76}$$

Нетрудно проверить, что $ — положительный и сохраняющий след оператор. Предположим, что начальной матрицей плотности является $\rho = \frac{1}{2}\mathbf{1}$, реализованная как ансамбль

$$\rho = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\downarrow_z\rangle\langle\downarrow_z|. \tag{3.77}$$

Поскольку $\mathrm{tr}(\sigma_1\rho) = 0$, эволюция ρ тривиальна и оба представителя ансамбля остаются неизменными. Если спин был приготовлен в состоянии $|\uparrow_z\rangle$, то в нем он и останется.

Теперь представим, что непосредственно после приготовления ансамбля мы ничего не делаем, если было приготовлено состояние $|\uparrow_z\rangle$, но если оказалось, что приготовлено $|\downarrow_z\rangle$, мы поворачиваем его в состояние $|\uparrow_x\rangle$. Теперь матрица плотности равна

$$\rho' = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x|, \tag{3.78}$$

так что $\mathrm{tr}(\sigma_1\rho') = \frac{1}{2}$. В результате эволюции, управляемой оператором $, она преобразуется в $\$(\rho') = \sigma_1\rho\sigma_1$. Тогда если спин был приготовлен в состоянии $|\uparrow_z\rangle$, то он эволюционирует в ортогональное состояние $|\downarrow_z\rangle$.

Следуя этим двум сценариям, первоначально приготовленное состояние $|\uparrow_z\rangle$ эволюционирует различным образом. Но в чем разница между этими двумя случаями? Разница в том, что *если* приготовлено начальное состояние спина $|\downarrow_z\rangle$, то мы совершаем различные действия: ничего не делаем в случае (1), но поворачиваем спин в случае (2). Тем не менее, мы обнаружили, что в этих двух случаях спин ведет себя по-разному, даже если первоначально было приготовлено состояние $|\uparrow_z\rangle$!

Мы привыкли говорить, что ρ описывает две (или более) различные альтернативы приготовления чистого состояния, только одна из которых действительно реализуется всякий раз, когда мы готовим кубит. Но мы обнаружили, что если мы готовим $|\uparrow_z\rangle$, то происходящее действительно *зависит от того, что мы были бы должны сделать*, если бы вместо этого было приготовлено $|\downarrow_z\rangle$. По-видимому, становится неразумно рассматривать два возможных приготовления как взаимно исключающие альтернативы. Эволюция альтернатив действительно зависит от других альтернатив, которые предположительно не были реализованы. Джо Полчински назвал это явление «телефоном Эверетта», поскольку различные «ветви волновой функции» выглядят способными «общаться» между собой.

Тогда нелинейная эволюция матрицы плотности имела бы странные, возможно, даже абсурдные следствия. И все-таки это не факт, что нели-

нейная эволюция должна быть исключена. Действительно, Джим Хартл доказывал, что существуют варианты «обобщенных квантовых механик», в которых допустима нелинейная эволюция, но тем не менее можно сохранить последовательную вероятностную интерпретацию. Несмотря на это, здесь мы будем следовать традиции и требовать, чтобы $ был линейным оператором.

3.2.3. Полная положительность

Было бы приятно прийти к выводу, что любой $, удовлетворяющий условиям (0)–(3), имеет представление операторной суммы и, следовательно, может быть реализован унитарной эволюцией подходящей бинарной системы. К сожалению, это не всегда возможно. И все же, к счастью, оказывается, что, добавив одно достаточно безобидно звучащее предположение, можно показать, что $ имеет представление операторной суммы.

Необходимым нам дополнительным предположением [в действительности более сильной версией (3)] является:

(3′) $ *вполне положителен.*

Полная положительность определяется следующим образом. Рассмотрим любое возможное расширение $\mathcal{H}_A$ до тензорного произведения $\mathcal{H}_A \otimes \mathcal{H}_B$; тогда $ вполне положителен в $\mathcal{H}_A$, если $\$_A \otimes \mathbf{1}_B$ является положительным для любого такого расширения.

Полная положительность, несомненно, является разумным свойством, чтобы нуждаться в физических основаниях. Если мы изучаем эволюцию системы A, то никогда нельзя быть уверенным в том, что нет взаимодействующей с ней системы B, о существовании которой мы не подозреваем. Полная положительность (в комбинации с другими предположениями) утверждает лишь то, что если система A эволюционирует, а система B — нет, то любая начальная матрица плотности составной системы эволюционирует в другую матрицу плотности.

Докажем, что предположений (0), (1), (2) и (3′) достаточно для того, чтобы $ был супероператором (имел представление операторной суммы). [Действительно, свойства (0)–(3′) могут рассматриваться как альтернативное определение супероператора.] Однако, прежде чем приступать к доказательству, попробуем пояснить понятие полной положительности на примере положительного, но не вполне положительного оператора. Таким примером служит оператор транспонирования

$$T : \rho \to \rho^T. \tag{3.79}$$

T сохраняет собственные значения оператора ρ и, следовательно, очевидно положителен. Но является ли T *вполне* положительным (положителен ли любой оператор $\mathbf{T}_A \otimes \mathbf{1}_B$)? Выберем $\dim \mathcal{H}_A = \dim \mathcal{H}_B = N$ и рассмотрим максимально запутанное состояние

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |i\rangle_A \otimes |i'\rangle_B, \tag{3.80}$$

где $\{|i\rangle_A\}$ и $\{|i'\rangle_B\}$ – ортонормированные базисы в $\mathcal{H}_A$ и $\mathcal{H}_B$ соответственно. Тогда

$$\mathbf{T}_A \otimes \mathbf{1}_B : \rho|\Phi\rangle_{AB\,AB}\langle\Phi| = \frac{1}{N} \sum_{i,j} \left(|i\rangle_{A\,A}\langle j|\right) \otimes \left(|i'\rangle_{B\,B}\langle j'|\right) \to$$

$$\to \rho' = \frac{1}{N} \sum_{i,j} \left(|j\rangle_{A\,A}\langle i|\right) \otimes \left(|i'\rangle_{B\,B}\langle j'|\right). \tag{3.81}$$

Мы видим, что оператор $N\rho'$ действует как

$$N\rho' : \left(\sum_i a_i|i\rangle_A\right) \otimes \left(\sum_j b_j|j'\rangle_B\right) \to \left(\sum_i a_i|i'\rangle_B\right) \otimes \left(\sum_j b_j|j\rangle_A\right), \tag{3.82}$$

или

$$N\rho'\left(|\varphi\rangle_A \otimes |\psi\rangle_B\right) = |\psi\rangle_A \otimes |\varphi\rangle_B. \tag{3.83}$$

Следовательно, $N\rho'$ – оператор перестановки (квадрат которого является тождественным оператором). Собственными состояниями $N\rho'$ являются *симметричные* относительно $A \leftrightarrow B$ состояния, которым отвечает собственное значение $+1$, и *антисимметричные* состояния, которым отвечает собственное значение -1. Поскольку ρ' имеет отрицательные собственные значения, он не является положительным и (поскольку ρ несомненно положителен), следовательно, $\mathbf{T}_A \otimes \mathbf{1}_B$ не сохраняет положительность. Таким образом, $\mathbf{T}_A$ – положительный оператор, но он не является вполне положительным.

3.2.4. ПОЗМ как супероператор

Унитарное преобразование, запутывающее A с B, после ортогонального измерения B может быть описано как ПОЗМ в A. Фактически положительные операторы, включая ПОЗМ, можно построить из операторов Крауса. Если $|\varphi\rangle_A$ эволюционирует как

$$|\varphi\rangle_A|0\rangle_B \to \sum_\mu \mathbf{M}_\mu|\varphi\rangle_A|\mu\rangle_B, \tag{3.84}$$

тогда измерение в B, проецирующее на базис $\{|\mu\rangle_B\}$, с вероятностью

$$\mathrm{Prob}(\mu) = {}_A\langle\varphi|\mathbf{M}_\mu^\dagger\mathbf{M}_\mu|\varphi\rangle_A \tag{3.85}$$

дает результат μ. Выражая $\boldsymbol{\rho}_A$ как ансамбль чистых состояний, мы находим вероятность

$$\mathrm{Prob}(\mu) = \mathrm{tr}\left(\mathbf{F}_\mu\boldsymbol{\rho}_A\right), \qquad \mathbf{F}_\mu = \mathbf{M}_\mu^\dagger\mathbf{M}_\mu \tag{3.86}$$

результата μ; очевидно, что $\mathbf{F}_\mu$ положителен, а равенство $\sum\limits_\mu \mathbf{F}_\mu = \mathbf{1}$ следует из нормировки операторов Крауса. Следовательно, это действительно реализация ПОЗМ.

В частности, ПОЗМ, модифицирующая матрицу плотности согласно

$$\boldsymbol{\rho} \to \sum_\mu \sqrt{\mathbf{F}_\mu}\boldsymbol{\rho}\sqrt{\mathbf{F}_\mu}, \tag{3.87}$$

является частным случаем супероператора. Так как каждый $\sqrt{\mathbf{F}_\mu}$ эрмитов, требование

$$\sum_\mu \mathbf{F}_\mu = \mathbf{1} \tag{3.88}$$

в точности совпадает с условием нормировки операторной суммы. Следовательно, ПОЗМ имеет «унитарное представление»: существует унитарный оператор $\mathbf{U}_{AB}$, действующий как

$$\mathbf{U}_{AB} \,:\, |\varphi\rangle_A \otimes |0\rangle_B \to \sum_\mu \sqrt{\mathbf{F}_\mu}|\varphi\rangle_A \otimes |\mu\rangle_B, \tag{3.89}$$

где $|\varphi\rangle_A$ — чистое состояние в A. Очевидно, что, выполняя ортогональное измерение в системе B, проецирующее на базис $\{|\mu\rangle_B\}$, мы можем реализовать ПОЗМ, которая готовит состояние

$$\boldsymbol{\rho}'_A = \frac{\sqrt{\mathbf{F}_\mu}\boldsymbol{\rho}_A\sqrt{\mathbf{F}_\mu}}{\mathrm{tr}\left(\mathbf{F}_\mu\boldsymbol{\rho}_A\right)} \tag{3.90}$$

с вероятностью

$$\mathrm{Prob}(\mu) = \mathrm{tr}\left(\mathbf{F}_\mu\boldsymbol{\rho}_A\right). \tag{3.91}$$

Эта реализация ПОЗМ, возможно, не самая эффективная (мы требуем, чтобы гильбертово пространство $\mathcal{H}_A \otimes \mathcal{H}_B$ имело размерность $N \cdot n$, если ПОЗМ имеет n возможных результатов), но в некоторых отношениях она наиболее удобна. ПОЗМ представляет собой наиболее общее измерение, которое мы можем выполнить в системе A, сначала запутывая ее с системой B, а затем выполняя ортогональное измерение в системе B.

3.3. Теорема о представлении Крауса

Теперь мы практически готовы доказать, что любой $, удовлетворяющий условиям (0), (1), (2) и (3′), имеет представление операторной суммы (теорема о представлении Крауса)[1]. Но сначала мы обсудим полезный трюк, который будет использован в доказательстве. Поскольку этот прием широко применяется, имеет смысл описать его отдельно.

Этот трюк (который мы будем называть «методом соответственного состояния») позволяет полностью охарактеризовать оператор $\mathbf{M}_A$, действующий в $\mathcal{H}_A$, описывая действие оператора $\mathbf{M}_A \otimes \mathbf{1}_B$ на единственное чистое максимально запутанное состояние[2] в $\mathcal{H}_A \otimes \mathcal{H}_B$ (где $\dim \mathcal{H}_B \geqslant \dim \mathcal{H}_A \equiv N$). Рассмотрим состояние

$$|\tilde{\psi}\rangle_{AB} = \sum_{i=1}^{N} |i\rangle_A \otimes |i'\rangle_B, \tag{3.92}$$

где $\{|i\rangle_A\}$ и $\{|i'\rangle_B\}$ — ортонормированные базисы в $\mathcal{H}_A$ и $\mathcal{H}_B$. (Мы выбрали $|\tilde{\psi}\rangle_{AB}$ нормированным таким образом, чтобы ${}_{AB}\langle\tilde{\psi}|\tilde{\psi}\rangle_{AB} = N$; это избавляет нас от необходимости писать множители $\sqrt{N}$ в формулах ниже.) Заметим, что любой вектор

$$|\varphi\rangle_A = \sum_i a_i |i\rangle_A \tag{3.93}$$

в $\mathcal{H}_A$ может быть представлен в виде «частичного» внутреннего произведения

$$|\varphi\rangle_A = {}_B\langle\varphi^*|\tilde{\psi}\rangle_{AB}, \tag{3.94}$$

где

$$|\varphi^*\rangle_B = \sum_i a_i^* |i'\rangle_B. \tag{3.95}$$

Мы говорим, что $|\varphi\rangle_A$ является «соответственным состоянием» «состояния-указателя» $|\varphi^*\rangle_B$. Отображение

$$|\varphi\rangle_A \to |\varphi^*\rangle_B, \tag{3.96}$$

[1] Приводимое здесь доказательство следует работе B. W. Schumacher, *Sending Entanglement Through Noisy Quantum Channels*, Phys. Rev., A54, 2614–2628 (1996); quant-ph/9604023 (см. Appendix A в этой работе).

[2] Мы говорим, что состояние $|\psi\rangle_{AB}$ максимально запутано, если $\mathrm{tr}_B(|\psi\rangle_{AB\,AB}\langle\psi|) \propto \mathbf{1}_A$.

очевидно, является *антилинейным* и фактически *антиунитарным* отображением из $\mathcal{H}_A$ в подпространство $\mathcal{H}_B$. Оператор $\mathbf{M}_A \otimes \mathbf{1}_B$, действуя на $|\tilde{\psi}\rangle_{AB}$, дает

$$(\mathbf{M}_A \otimes \mathbf{1}_B)|\tilde{\psi}\rangle_{AB} = \sum_i \mathbf{M}_A|i\rangle_A \otimes |i'\rangle_B. \tag{3.97}$$

Из этого состояния мы можем выделить $\mathbf{M}_A|\varphi\rangle_A$ в качестве соответственного состояния

$$_B\langle\varphi^*|\mathbf{M}_A \otimes \mathbf{1}_B|\tilde{\psi}\rangle_{AB} = \mathbf{M}_A|\varphi\rangle_A. \tag{3.98}$$

Мы можем интерпретировать формализм соответственного состояния, говоря что можно реализовать ансамбль чистых состояний в $\mathcal{H}_A$, выполняя измерения в $\mathcal{H}_B$ на запутанном состоянии — если измерение в $\mathcal{H}_B$ дает результат $|\varphi^*\rangle_B$, то приготовленным состоянием является $|\varphi\rangle_A$. Если мы намерены применить некоторый линейный оператор в $\mathcal{H}_A$, то обнаружим, что результат не зависит от того, было ли сначала приготовлено состояние, а затем на него подействовали оператором, или сначала был применен оператор, а затем приготовлено состояние. Конечно, этот вывод имеет физический смысл. Можно даже представить, что приготовление и действие оператора являются событиями, разделенными пространственно-подобным интервалом, так что временное упорядочение становится нековариантным (зависящим от наблюдателя).

Мы покажем, что $\$_A$ имеет представление операторной суммы, применяя метод соответственного состояния не к операторам, а к супероператорам. *Поскольку* мы предполагаем, что $\$_A$ вполне положителен, мы знаем, что $\$_A \otimes \mathbf{1}_B$ положителен. Следовательно, если мы применяем $\$_A \otimes \mathbf{1}_B$ к $\tilde{\rho}_{AB} = |\tilde{\psi}\rangle_{AB\,AB}\langle\tilde{\psi}|$, то результатом будет положительный оператор, (ненормированная) матрица плотности $\tilde{\rho}'_{AB}$ в $\mathcal{H}_A \otimes \mathcal{H}_B$. Подобно любой матрице плотности, $\tilde{\rho}'_{AB}$ может быть представлена как ансамбль чистых состояний. Следовательно,

$$(\$_A \otimes \mathbf{1}_B)(|\tilde{\psi}\rangle_{AB\,AB}\langle\tilde{\psi}|) = \sum_\mu q_\mu |\tilde{\Phi}_\mu\rangle_{AB\,AB}\langle\tilde{\Phi}_\mu| \tag{3.99}$$

где $q_\mu > 0$, $\sum_\mu q_\mu = 1$, а каждый вектор $|\tilde{\Phi}_\mu\rangle_{AB}$, подобно $|\tilde{\psi}\rangle_{AB}$, нормирован таким образом, что $_{AB}\langle\tilde{\Phi}_\mu|\tilde{\Phi}_\mu\rangle_{AB} = N$. Применяя метод соответственного состояния, имеем

$$\begin{aligned}\$_A(|\varphi\rangle_{A\,A}\langle\varphi|) &= {}_B\langle\varphi^*|(\$_A \otimes \mathbf{1}_B)(|\tilde{\psi}\rangle_{AB\,AB}\langle\tilde{\psi}|)|\varphi^*\rangle_B = \\ &= \sum_\mu q_\mu {}_B\langle\varphi^*|\tilde{\Phi}_\mu\rangle_{AB\,AB}\langle\tilde{\Phi}_\mu|\varphi^*\rangle_B.\end{aligned} \tag{3.100}$$

Мы почти у цели; определим оператор $\mathbf{M}_\mu$ в $\mathcal{H}_A$ соотношением

$$\mathbf{M}_\mu : |\varphi\rangle_A \to \sqrt{q_\mu}\,{}_B\langle\varphi^*|\tilde{\Phi}_\mu\rangle_{AB}. \tag{3.101}$$

Можно проверить, что

1) Оператор $\mathbf{M}_\mu$ *линеен*, поскольку отображение $|\varphi\rangle_A \to |\varphi^*\rangle_B$ *антилинейно*.

2) $\$_A(|\varphi\rangle_{A\,A}\langle\varphi|) = \sum_\mu \mathbf{M}_\mu(|\varphi\rangle_{A\,A}\langle\varphi|)\mathbf{M}_\mu^\dagger$ для любого чистого состояния $|\varphi\rangle_A \in \mathcal{H}_A$.

3) $\$_A(\boldsymbol{\rho}_A) = \sum_\mu \mathbf{M}_\mu \boldsymbol{\rho}_A \mathbf{M}_\mu^\dagger$ для любой матрицы плотности $\boldsymbol{\rho}_A$, поскольку $\boldsymbol{\rho}_A$ может быть представлена как ансамбль чистых состояний, а $\$_A$ линеен.

4) $\sum_\mu \mathbf{M}_\mu \mathbf{M}_\mu^\dagger = \mathbf{1}_A$, поскольку $\$_A$ сохраняет след для любого $\boldsymbol{\rho}_A$.

Таким образом, мы построили представление операторной суммы для $\$_A$.

Вкратце, доказательство состоит в следующем. Поскольку $\$_A$ вполне положителен, то $\$_A \otimes \mathbf{1}_B$ преобразует максимально запутанную матрицу плотности в $\mathcal{H}_A \otimes \mathcal{H}_B$ в другую матрицу плотности. Эта матрица плотности может быть выражена, как ансамбль чистых состояний. Каждому из этих чистых состояний в $\mathcal{H}_A \otimes \mathcal{H}_B$ можно сопоставить (с помощью метода соответственных состояний) слагаемое операторной суммы.

Рассматривая таким образом представление операторной суммы, можно легко установить два важных следствия:

Как много операторов Крауса? Каждый оператор $\mathbf{M}_\mu$ связан с состоянием $|\Phi_\mu\rangle$ в представлении ансамбля $\boldsymbol{\rho}'_{AB}$. Так как максимальный ранг $\boldsymbol{\rho}'_{AB}$ равен N^2 (где $N = \dim \mathcal{H}_A$), $\$_A$ всегда имеет представление операторной суммы с максимальным числом операторов Крауса, равным N^2.

Какова неоднозначность? Выше мы отмечали, что операторы Крауса

$$\mathbf{N}_a = \mathbf{M}_\mu \mathbf{U}_{\mu a} \tag{3.102}$$

($\mathbf{U}_{\mu a}$ — унитарное преобразование) представляют тот же, что и $\mathbf{M}_\mu$, супероператор $\$_A$. Теперь можно сказать, что любые два представления Крауса должны быть связаны таким образом. (Если число операторов $\mathbf{N}_a$ оказывается больше, чем $\mathbf{M}_\mu$, тогда набору $\mathbf{M}_\mu$, очевидно, нужно добавить соответствующее количество нулевых операторов, так чтобы эти два набора

операторов имели одинаковую мощность.) Это свойство можно рассматривать как следствие теоремы ЖХЙВ.

Описанная выше конструкция соответственного состояния устанавливает взаимно однозначное соответствие между представлениями ансамблей (ненормированной) матрицы плотности $(\$_A \otimes \mathbf{1}_B)(|\tilde{\psi}\rangle_{AB\,AB}\langle\tilde{\psi}|)$ и представлениями операторных сумм $\$_A$. (Мы явно описали, как перейти от представления ансамбля к представлению операторной суммы, но, очевидно, можно пойти и другим путем. Если

$$\$_A(|i\rangle_{A\,A}\langle j|) = \sum_\mu \mathbf{M}_\mu |i\rangle_{A\,A}\langle j|\mathbf{M}_\mu^\dagger, \tag{3.103}$$

то

$$\begin{aligned}(\$_A \otimes \mathbf{1}_B)(|\tilde{\psi}\rangle_{AB\,AB}\langle\tilde{\psi}|) &= \sum_{i,j,\mu} (\mathbf{M}_\mu|i\rangle_A|i'\rangle_B)({}_A\langle j|{}_B\langle j'|\mathbf{M}_\mu^\dagger{}_B\langle j'|) \\ &= \sum_\mu q_\mu |\tilde{\Phi}_\mu\rangle_{AB\,AB}\langle\tilde{\Phi}_\mu|,\end{aligned} \tag{3.104}$$

где

$$\sqrt{q_\mu}|\tilde{\Phi}_\mu\rangle_{AB} = \sum_i \mathbf{M}_\mu|i\rangle_A|i'\rangle_B. \tag{3.105}$$

Рассмотрим теперь два таких ансамбля (или соответственно два представления $\$_A$ операторными суммами) $\{\sqrt{q_\mu}|\tilde{\Phi}_\mu\rangle_{AB}\}$ и $\{\sqrt{p_a}|\tilde{\Upsilon}_a\rangle_{AB}\}$. Для каждого ансамбля в $\mathcal{H}_{AB} \otimes \mathcal{H}_C$ существует соответствующее «очищение»:

$$\begin{aligned}&\sum_\mu \sqrt{q_\mu}|\tilde{\Phi}_\mu\rangle_{AB}|\alpha_\mu\rangle_C, \\ &\sum_a \sqrt{p_a}|\tilde{\Upsilon}_a\rangle_{AB}|\beta_a\rangle_C,\end{aligned} \tag{3.106}$$

где $\{|\alpha_\mu\rangle_C\}$ и $\{|\beta_a\rangle_C\}$ — два разных ортонормированных набора из $\mathcal{H}_C$. Теорема ЖХЙВ утверждает, что эти два «очищения» связаны между собой действующим в $\mathcal{H}_C$ унитарным преобразованием $\mathbf{1}_{AB} \otimes \mathbf{U}'_C$. Следовательно,

$$\begin{aligned}\sum_a \sqrt{p_a}|\tilde{\Upsilon}_a\rangle_{AB}|\beta_a\rangle_C \quad &\sum_\mu \sqrt{q_\mu}|\tilde{\Phi}_\mu\rangle_{AB}\mathbf{U}'_C|\alpha_\mu\rangle_C \\ &= \sum_{\mu,a} \sqrt{q_\mu}|\tilde{\Phi}_\mu\rangle_{AB}U_{\mu a}|\beta_a\rangle_C.\end{aligned} \tag{3.107}$$

Здесь второе равенство мы получили, заметив что ортонормированные базисы $\{|\alpha_\mu\rangle_C\}$ и $\{|\beta_a\rangle_C\}$ связаны между собой унитарным преобразованием, а произведение преобразований, в свою очередь, унитарно. Мы приходим к выводу, что

$$\sqrt{p_a}|\tilde{\Upsilon}_a\rangle_{AB} = \sum_\mu \sqrt{q_\mu}|\tilde{\Phi}_\mu\rangle_{AB} U_{\mu a} \tag{3.108}$$

(где $U_{\mu a}$ — унитарное преобразование), откуда следует, что

$$\mathbf{N}_a = \sum_\mu \mathbf{M}_\mu \mathbf{U}_{\mu a}. \tag{3.109}$$

Замечание. Поскольку мы уже установили, что можем перейти от представления операторной суммы для $\$_A$ к унитарному представлению, мы нашли, что любой «разумный» закон эволюции оператора плотности в $\mathcal{H}_A$ может быть реализован унитарным преобразованием $\mathbf{U}_{AB}$, действующим в $\mathcal{H}_A \otimes \mathcal{H}_B$ как

$$\mathbf{U}_{AB} : |\psi\rangle_A \otimes |0\rangle_B \to \sum_\mu |\varphi\rangle_A \otimes |\mu\rangle_B. \tag{3.110}$$

Является ли этот результат неожиданным? Возможно, да. Мы можем интерпретировать супероператор как *описывающий эволюцию системы* (A), взаимодействующей с окружением (B). В общем случае состояния системы запутаны с ее окружением. Но в (3.110) предполагается, что начальное состояние не запутано. Несмотря на то что реальная система всегда связана запутыванием с ее окружением, при описании эволюции ее матрицы плотности без потери общности можно *представлять*, что в момент, когда мы начинаем ее наблюдать, предварительное запутывание отсутствует!

Замечание. Представление операторной суммы даст очень удобный способ выражения любого вполне положительного $\$$. Но положительный $\$$ не допускает такого представления, если не является вполне положительным. Насколько мне известно, не существует удобного, сопоставимого с представлением Крауса, способа выразить наиболее общий *положительный* $\$$.

3.4. Три квантовых канала

Лучше всего познакомиться с понятием супероператора, изучив несколько примеров. Мы рассмотрим три примера (все они интересны и по-

лезны) супероператоров для одного кубита. Из уважения к традиционной терминологии (классической) теории связи я буду ссылаться на эти супероператоры как на *квантовые каналы*. Мы можем представлять, что $ описывает судьбу квантовой информации, которая с некоторой потерей точности воспроизведения посылается от передатчика к приемнику. Или, если угодно, можно считать (в духе предыдущего обсуждения), что передача идет во времени, а не в пространстве, то есть $ описывает эволюцию квантовой системы, взаимодействующей с ее окружением.

3.4.1. Деполяризующий канал

Деполяризующий канал представляет собой модель декогерентизации кубита, имеющую особенно тонкие свойства симметрии. Мы можем описать его, говоря что с вероятностью $1-p$ кубит остается неповрежденным, тогда как с вероятностью p возникает ошибка. Она может быть любой из трех типов, причем все три типа ошибок равновероятны. Если $\{|0\rangle, |1\rangle\}$ — ортонормированный базис кубита, их можно описать следующим образом:

1. Ошибка инвертирования бита $\begin{matrix}|0\rangle \to |1\rangle \\ |1\rangle \to |0\rangle\end{matrix}$ или $|\psi\rangle \to \sigma_1|\psi\rangle$, $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

2. Ошибка обращения фазы $\begin{matrix}|0\rangle \to |0\rangle \\ |1\rangle \to -|1\rangle\end{matrix}$ или $|\psi\rangle \to \sigma_3|\psi\rangle$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

3. Обе ошибки $\begin{matrix}|0\rangle \to +i|1\rangle \\ |1\rangle \to -i|0\rangle\end{matrix}$ или $|\psi\rangle \to \sigma_2|\psi\rangle$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$.

При появлении ошибки $|\psi\rangle$ превращается в ансамбль трех равновероятных состояний: $\sigma_1|\psi\rangle$, $\sigma_2|\psi\rangle$ и $\sigma_3|\psi\rangle$.

Унитарное представление

Деполяризующий канал может быть представлен унитарным оператором, действующим в $\mathcal{H}_A \otimes \mathcal{H}_E$, где размерность пространства $\mathcal{H}_E$ равна четырем. (Я обозначаю здесь это пространство $\mathcal{H}_E$, чтобы подтолкнуть вас

к мысли о вспомогательной системе как окружении.) Унитарный оператор $\mathbf{U}_{AE}$ действует как

$$\mathbf{U}_{AE}\colon |\psi\rangle_A \otimes |0\rangle_E \to \sqrt{1-p}|\psi\rangle_A \otimes |0\rangle_E + \\ + \sqrt{\frac{p}{3}}\Big[\boldsymbol{\sigma}_1|\psi\rangle_A \otimes |1\rangle_E + \boldsymbol{\sigma}_2|\psi\rangle_A \otimes |2\rangle_E + \boldsymbol{\sigma}_3|\psi\rangle_A \otimes |3\rangle_E\Big]. \quad (3.111)$$

(Поскольку $\mathbf{U}_{AE}$ сохраняет внутреннее произведение, он имеет унитарное расширение на все пространство $\mathcal{H}_A \otimes \mathcal{H}_E$.) Окружение эволюционирует к одному из четырех взаимно ортогональных состояний, «хранящих запись» о том, что произошло; если бы мы могли измерить окружение в базисе $\{|\mu\rangle_E,\ \mu = 0,1,2,3\}$, мы узнали бы, какого сорта ошибка возникла (тогда мы были бы в состоянии вмешаться и устранить ошибку).

Представление Крауса

Чтобы получить представление канала в виде операторной суммы, вычислим частичный след по окружению в базисе $\{|\mu\rangle_E\}$. Тогда

$$\mathbf{M}_\mu = {}_E\langle\mu|\mathbf{U}_{AE}|0\rangle_E, \quad (3.112)$$

где

$$\mathbf{M}_0 = \sqrt{1-p}\,\mathbf{1}, \quad \mathbf{M}_1 = \sqrt{\frac{p}{3}}\,\boldsymbol{\sigma}_1, \quad \mathbf{M}_2 = \sqrt{\frac{p}{3}}\,\boldsymbol{\sigma}_2, \quad \mathbf{M}_3 = \sqrt{\frac{p}{3}}\,\boldsymbol{\sigma}_3. \quad (3.113)$$

Используя $\boldsymbol{\sigma}_i^2 = \mathbf{1}$, можно непосредственно проверить условие нормировки:

$$\sum_\mu \mathbf{M}_\mu^\dagger \mathbf{M}_\mu = \left[(1-p) + 3\cdot\frac{p}{3}\right]\mathbf{1} = \mathbf{1}. \quad (3.114)$$

Произвольная начальная матрица плотности кубита $\boldsymbol{\rho}_A$ преобразуется как

$$\boldsymbol{\rho}_A \to \boldsymbol{\rho}'_A = (1-p)\boldsymbol{\rho}_A + \frac{p}{3}\big(\boldsymbol{\sigma}_1\boldsymbol{\rho}_A\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2\boldsymbol{\rho}_A\boldsymbol{\sigma}_2 + \boldsymbol{\sigma}_3\boldsymbol{\rho}_A\boldsymbol{\sigma}_3\big), \quad (3.115)$$

где мы суммируем по четырем (в принципе различимым) путям, по которым могло бы эволюционировать окружение.

Представление соответственного состояния

Канал можно также охарактеризовать, описывая как в нем преобразуется максимально запутанное состояние двух кубитов, если канал действует

только на первый кубит. Существует четыре взаимно ортогональных максимально запутанных состояния, которые можно записать в виде

$$
\begin{aligned}
|\phi^+\rangle_{AB} &= \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} + |11\rangle_{AB}\right),\\
|\phi^-\rangle_{AB} &= \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} - |11\rangle_{AB}\right),\\
|\psi^+\rangle_{AB} &= \frac{1}{\sqrt{2}}\left(|01\rangle_{AB} + |10\rangle_{AB}\right),\\
|\psi^-\rangle_{AB} &= \frac{1}{\sqrt{2}}\left(|01\rangle_{AB} - |10\rangle_{AB}\right).
\end{aligned}
\tag{3.116}
$$

Если начальным состоянием является $|\phi^-\rangle_{AB}$, то, когда деполяризующий канал действует на первый кубит, запутанное состояние эволюционирует как

$$
\begin{aligned}
|\phi^+\rangle_{AB\,AB}\langle\phi^+| \to (1-p)|\phi^+\rangle_{AB\,AB}\langle\phi^+| +\\
+\frac{p}{3}\left(|\psi^+\rangle_{AB\,AB}\langle\psi^+| + |\psi^-\rangle_{AB\,AB}\langle\psi^-| + |\phi^-\rangle_{AB\,AB}\langle\phi^-|\right).
\end{aligned}
\tag{3.117}
$$

В «наихудшем» квантовом канале $p = 3/4$, в этом случае начальное запутанное состояние эволюционирует в

$$
\begin{aligned}
|\phi^+\rangle_{AB\,AB}\langle\phi^+| \to \frac{1}{4}\left(|\phi^+\rangle_{AB\,AB}\langle\phi^+| + |\phi^-\rangle_{AB\,AB}\langle\phi^-| +\right.\\
\left. + |\psi^+\rangle_{AB\,AB}\langle\psi^+| + |\psi^-\rangle_{AB\,AB}\langle\psi^-|\right) = \frac{1}{4}\mathbf{1}_{AB}.
\end{aligned}
\tag{3.118}
$$

Оно становится полностью случайной матрицей плотности в $\mathcal{H}_A \otimes \mathcal{H}_B$. Тогда, применяя метод соответственного состояния, можно увидеть, что чистое состояние $|\varphi\rangle_A$ одного кубита A эволюционирует как

$$
|\varphi\rangle_{A\,A}\langle\varphi| \to {}_B\left\langle\varphi^*\left|2\left(\frac{1}{4}\mathbf{1}_{AB}\right)\right|\varphi^*\right\rangle_B = \frac{1}{2}\mathbf{1}_A; \tag{3.119}
$$

оно становится случайной матрицей в $\mathcal{H}_A$, независимо от значения начального состояния $|\varphi\rangle_A$. Как если бы канал выбросил начальное состояние и заменил его совершенно случайным мусором.

Альтернативным представлением эволюции максимально запутанного состояния является

$$|\phi^+\rangle_{AB\,AB}\langle\phi^+| \to \left(1-\frac{4}{3}p\right)|\phi^+\rangle_{AB\,AB}\langle\phi^+| + \frac{4}{3}p\left(\frac{1}{4}\mathbf{1}_{AB}\right). \quad (3.120)$$

Таким образом, вместо того, чтобы говорить о трех типах равновероятных ошибок, появляющихся с вероятностью p каждая, мы могли бы говорить, что с вероятностью $4p/3$ возникает ошибка, полностью «рандомизирующая» состояние (мы можем так говорить по крайней мере при $p \leqslant 3/4$). Наличие двух естественных способов определения «вероятности ошибки» в этом канале иногда может приводить к путанице и недоразумениям.

Полезной мерой того, насколько хорошо канал сохраняет исходную квантовую информацию, является так называемая «точность воспроизведения запутанности» F_e. Она количественно определяет, насколько конечная матрица плотности «близка» к исходному максимально запутанному состоянию $|\phi^+\rangle$:

$$F_e = \langle\phi^+|\boldsymbol{\rho}'|\phi^+\rangle. \quad (3.121)$$

Для деполяризующего канала мы имеем $F_e = 1-p$ и, следовательно, можем интерпретировать F_e как вероятность отсутствия ошибки.

Представление сферы Блоха

Также поучительно рассмотреть, как деполяризующий канал действует на сфере Блоха. Произвольная матрица плотности одного кубита может быть записана в виде

$$\boldsymbol{\rho} = \frac{1}{2}(\mathbf{1} + \vec{P}\cdot\vec{\boldsymbol{\sigma}}), \quad (3.122)$$

где $\vec{P}$ — «спиновая поляризация» кубита. Повернем оси таким образом, чтобы $\vec{P} = P_3\hat{e}_3$, а $\boldsymbol{\rho} = \frac{1}{2}(\mathbf{1} + P_3\boldsymbol{\sigma}_3)$. Тогда, поскольку $\boldsymbol{\sigma}_3\boldsymbol{\sigma}_3\boldsymbol{\sigma}_3 = \boldsymbol{\sigma}_3$, а $\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3\boldsymbol{\sigma}_1 = -\boldsymbol{\sigma}_3 = \boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_2$, найдем, что

$$\boldsymbol{\rho}' = \left(1-p+\frac{p}{3}\right)\frac{1}{2}(1+P_3\boldsymbol{\sigma}_3) + \frac{2p}{3}\frac{1}{2}(\mathbf{1}-P_3\boldsymbol{\sigma}_3) \quad (3.123)$$

или $P_3' = (1-4p/3)P_3$. С учетом симметрии относительно поворотов видно, что независимо от ориентации $\vec{P}$

$$\vec{P}' = \left(1-\frac{4}{3}p\right)\vec{P}. \quad (3.124)$$

Следовательно, под действием деполяризующего канала происходит однородное сжатие сферы Блоха; спиновая поляризация уменьшается на множитель $(1 - 4p/3)$ (вот почему мы называем этот канал деполяризующим). Этот результат следовало ожидать в связи со сделанным ранее заключением о том, что с вероятностью $4p/3$ в канале происходит полная «рандомизация» спина.

Обратимость?

Почему мы говорим, что супероператор необратим? Очевидно, мы можем обратить однородное сжатие сферы однородным же раздуванием. Но беда в том, что раздувание сферы Блоха не положительно и потому не является супероператором. Раздувание преобразует $\vec{P}$ длины $|\vec{P}| \leqslant 1$ в вектор длины $|\vec{P}| \geqslant 1$, преобразуя таким образом оператор плотности в оператор с отрицательным собственным значением. Декогерентизация может сжать шар, но нет физического процесса, способного снова надуть его! Супероператор, бегущий назад во времени, не является супероператором.

3.4.2. Канал затухания фазы

Нашим следующим примером является канал *затухания фазы*. Этот случай интересен с практической точки зрения, поскольку представляет голую, свободную от несущественных математических деталей, карикатуру декогерентизации в реальной физической ситуации.

Унитарное представление

Унитарным представлением канала является

$$\begin{aligned} |0\rangle_A|0\rangle_E &\to \sqrt{1-p}|0\rangle_A|0\rangle_E + \sqrt{p}|0\rangle_A|1\rangle_E, \\ |1\rangle_A|0\rangle_E &\to \sqrt{1-p}|1\rangle_A|0\rangle_E + \sqrt{p}|1\rangle_A|2\rangle_E. \end{aligned} \tag{3.125}$$

В этом случае, в отличие от деполяризующего канала, кубит A не совершает никаких переходов. Вместо этого он время от времени (с вероятностью p) «рассеивает» окружение, толкая его в состояние $|1\rangle_E$, если A находится в состоянии $|0\rangle_A$, и — в состояние $|2\rangle_E$, если A находится в состоянии $|1\rangle_A$. Более того, также в отличие от деполяризующего канала, этот канал выделяет предпочтительный базис для кубита A; только в базисе $\{|0\rangle_A\ |1\rangle_A\}$ не происходит опрокидывание спина кубита A.

Представление Крауса

Вычисляя частичный след по $\mathcal{H}_E$ в базисе $\{|0\rangle_E, |1\rangle_E, |2\rangle_E\}$, получим операторы Крауса

$$\mathbf{M}_0 = \sqrt{1-p}\,\mathbf{1}, \quad \mathbf{M}_1 = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{M}_2 = \sqrt{p}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.126)$$

Нетрудно проверить, что $\mathbf{M}_0^2 + \mathbf{M}_1^2 + \mathbf{M}_2^2 = \mathbf{1}$. В этом случае не обязательно иметь три оператора Крауса; как вы покажете в домашнем упражнении, возможно представление двумя операторами Крауса.

Начальная матрица плотности $\boldsymbol{\rho}$ эволюционирует к

$$\$(\boldsymbol{\rho}) = \mathbf{M}_0\boldsymbol{\rho}\mathbf{M}_0 + \mathbf{M}_1\boldsymbol{\rho}\mathbf{M}_1 + \mathbf{M}_2\boldsymbol{\rho}\mathbf{M}_2 = \\ = (1-p)\boldsymbol{\rho} + p\begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix} = \begin{pmatrix} \rho_{00} & (1-p)\rho_{01} \\ (1-p)\rho_{10} & \rho_{11} \end{pmatrix}; \quad (3.127)$$

таким образом, диагональные элементы $\boldsymbol{\rho}$ остаются неизменными, тогда как недиагональные — затухают.

Предположим, что отнесенная к единице времени вероятность акта рассеяния Γ такова, что вероятность рассеяния за время Δt гораздо меньше единицы ($p = \Gamma\Delta t \ll 1$). Эволюция в течение времени $t = n\Delta t$ управляется супероператором $\$^n$, так что недиагональные элементы матрицы плотности подавляются по закону $(1-p)^n = (1-\Gamma\Delta t)^{t/\Delta t} \to \exp(-\Gamma t)$ (при $\Delta t \to 0$). Таким образом, если мы приготовили начальное чистое состояние $a|0\rangle + b|1\rangle$, то спустя время $t \gg \Gamma^{-1}$ оно распадается в некогерентную суперпозицию $\boldsymbol{\rho}' = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|$. Декогерентизация возникает в выделенном базисе $\{|0\rangle, |1\rangle\}$.

Представление сферы Блоха

Эту задачу вы исследуете в домашнем упражнении.

Интерпретация

Канал затухания фазы можно интерпретировать как описывающий тяжелую «классическую» частицу (например, частицу межзвездной пыли), взаимодействующую с фоновым газом легких частиц (например, с фотонами реликтового микроволнового излучения). Можно представить, что первоначально пылинка была приготовлена в суперпозиции собственных со-

стояний оператора координаты $|\psi\rangle = \frac{1}{\sqrt{2}}(|x\rangle + |-x\rangle)$ (или в более общей суперпозиции слабо перекрывающихся, пространственно-локализованных волновых пакетов). Можно контролировать поведение частички пыли, но безнадежно пытаться следить за квантовым состоянием всех фотонов, рассеиваемых этой частицей; для наших целей ее квантовое состояние описывается матрицей плотности ρ, полученной после вычисления следа по фотонным степеням свободы.

Наш анализ канала затухания фазы показывает, что если фотоны рассеиваются частицей с частотой Γ, то недиагональные элементы матрицы плотности ρ затухают как $\exp(-\Gamma t)$ и становятся полностью пренебрежимыми при $t \gg \Gamma^{-1}$. Начиная с этого момента, когерентная суперпозиция собственных состояний оператора положения полностью разрушена — нет никакой возможности восстановить волновые пакеты и заставить их интерферировать. (Если мы пытаемся получить с помощью частиц пыли картину интерференции на двух щелях, то мы не увидим ее, если пылинкам необходимо время $t \gg \Gamma^{-1}$, чтобы пройти путь от источника до экрана.)

Частицы пыли тяжелы. Вследствие большой инерции, их состояние движения мало подвержено влиянию со стороны рассеиваемых фотонов. Таким образом, имеется два несоизмеримых временных масштаба, имеющих отношение к динамике частиц пыли. С одной стороны, это время затухания, то есть время, за которое значительная часть импульса частиц передается фотонам: это большое время, если частицы достаточно тяжелы. С другой стороны, существует временной масштаб декогерентизации. В этой модели он имеет порядок Γ^{-1} — времени, в течение которого на частице пыли происходит рассеяние *одного* фотона и которое гораздо короче временного масштаба затухания. В макроскопическом объекте декогерентизация протекает *быстро*.

Как мы уже отмечали, канал затухания фазы выделяет предпочтительный базис для декогерентизации, в нашей «интерпретации» мы предположили, что им является базис собственных состояний оператора положения. С физической точки зрения декогерентизация выделяет пространственно локализованные состояния частиц пыли, поскольку их *взаимодействие* с фотонами локализовано в пространстве. Частицы, находящиеся в различимых пространственных положениях, стремятся рассеивать фотоны во взаимно ортогональные состояния.

Даже если «частицы» разделены настолько мало, что они не разрешаются рассеиваемыми фотонами, процесс декогерентизации все еще работает подобным образом. Возможно, фотоны, рассеянные частицами, находящимися в точках $+x$ и $-x$, не являются взаимно ортогональными, а вместо

этого имеют ненулевое перекрытие

$$\langle \gamma + | \gamma - \rangle = 1 - \varepsilon, \qquad \varepsilon \ll 1. \tag{3.128}$$

Тем не менее канал затухания фазы описывает эту ситуацию, но теперь с p, замененным на εp (если p – по-прежнему вероятность акта рассеяния). Таким образом, темп декогерентизации становится равным $\Gamma_{\rm dec} = \varepsilon \Gamma_{\rm scat}$, где $\Gamma_{\rm scat}$ – частота рассеяния (см. домашнее задание).

Интуитивное понимание, извлекаемое из этой простой модели, применимо к огромному множеству физических ситуаций. Распад когерентной суперпозиции макроскопически различимых состояний «тяжелых» объектов происходит гораздо быстрее их затухания. Пространственная локализация взаимодействия системы с ее окружением делает предпочтительным для декогерентизации «локальный» базис. По-видимому, подобные принципы можно применить к декогерентизации «состояния кота» $\frac{1}{\sqrt{2}}(|\text{dead}\rangle + |\text{alive}\rangle)$, поскольку состояния «мертвый» и «живой» можно различить локальными испытаниями.

3.4.3. Канал затухания амплитуды

Канал затухания амплитуды представляет собой схематическую модель распада возбужденного состояния (двухуровневого) атома вследствие спонтанного излучения фотона. Регистрируя излучаемый фотон («наблюдая за окружением»), мы можем выполнить ПОЗМ, которая дает информацию о начальном состоянии атома.

Унитарное представление

Обозначим как $|0\rangle_A$ основное состояние атома, а интересующее нас возбужденное состояние – $|1\rangle_A$. Роль «окружения» играет электромагнитное поле, начальным состоянием которого предполагается основное $|0\rangle_E$. Существует вероятность p того, что некоторое время спустя возбужденное состояние распадается в основное $|0\rangle_A$, что сопровождается излучением фотона и, следовательно, переходом окружения из состояния $|0\rangle_E$ («нет фотонов») в состояние $|1\rangle_E$ («один фотон»). Эта эволюция описывается унитарным преобразованием, действующим на атом и окружение как

$$\begin{aligned} |0\rangle_A|0\rangle_E &\to |0\rangle_A|0\rangle_E, \\ |1\rangle_A|0\rangle_E &\to \sqrt{1-p}|1\rangle_A|0\rangle_E + \sqrt{p}|0\rangle_A|1\rangle_E. \end{aligned} \tag{3.129}$$

(Естественно, если начальным состоянием атома является основное, а окружение находится при нулевой температуре, то никакие переходы не происходят).

Операторы Крауса

Вычисляя частичный след по окружению в базисе $\{|0\rangle_E,\ |1\rangle_E\}$, найдем операторы Крауса

$$\mathbf{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \qquad \mathbf{M}_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}. \tag{3.130}$$

Нетрудно проверить, что

$$\mathbf{M}_0^\dagger \mathbf{M}_0 + \mathbf{M}_1^\dagger \mathbf{M}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix} = \mathbf{1}. \tag{3.131}$$

Оператор $\mathbf{M}_1$ индуцирует «квантовый скачок» — распад состояния $|1\rangle_A$ в $|0\rangle_A$, а $\mathbf{M}_0$ описывает эволюцию состояния в отсутствии скачков. Матрица плотности изменяется как

$$\begin{aligned} \boldsymbol{\rho} \ \to\ \$(\boldsymbol{\rho}) &= \mathbf{M}_0 \boldsymbol{\rho} \mathbf{M}_0^\dagger + \mathbf{M}_1 \boldsymbol{\rho} \mathbf{M}_1^\dagger = \\ &= \begin{pmatrix} \rho_{00} & \sqrt{1-p}\rho_{01} \\ \sqrt{1-p}\rho_{10} & (1-p)\rho_{11} \end{pmatrix} + \begin{pmatrix} p\rho_{11} & 0 \\ 0 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\rho_{01} \\ \sqrt{1-p}\rho_{10} & (1-p)\rho_{11} \end{pmatrix}. \end{aligned} \tag{3.132}$$

Если мы применим канал n раз подряд, то матричный элемент ρ_{11} уменьшится согласно

$$\rho_{11} \ \to\ (1-p)^n \rho_{11}. \tag{3.133}$$

Следовательно, если вероятность перехода в течение времени Δt равна $\Gamma \Delta t$, то вероятность того, что возбужденное состояние проживет в течение времени t равна $(1-\Gamma\Delta t)^{t/\Delta t} \to e^{-\Gamma t}$, ожидаемый экспоненциальный закон затухания.

При $t \to \infty$ вероятность затухания стремится к единице, следовательно:

$$\$(\boldsymbol{\rho}) = \begin{pmatrix} \rho_{00} + \rho_{11} & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.134}$$

Атом всегда сваливается в свое основное состояние. Этот пример показывает, что иногда оказывается возможным, что супероператор преобразует начальное смешанное состояние, например:

$$\rho = \begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix} \tag{3.135}$$

в чистое конечное состояние.

Контроль окружения

В случае распада возбужденного атомного состояния, сопровождающегося излучением фотона, полезно следить за состоянием окружения с помощью детектора фотонов. Измерение окружения готовит чистое состояние атома и, в сущности, предотвращает процесс декогерентизации.

Возвращаясь к унитарному представлению канала затухания амплитуды, мы видим, что когерентная суперпозиция основного и возбужденного атомных состояний эволюционирует как

$$\left(a|0\rangle_A + b|1\rangle_B\right)|0\rangle_E \to \left(a|0\rangle_A + b\sqrt{1-p}|1\rangle_B\right)|0\rangle_E + b\sqrt{p}|0\rangle_A|1\rangle_E. \tag{3.136}$$

Регистрируя фотон и, следовательно, проецируя окружение на состояние $|1\rangle_E$, мы готовим атомное состояние $|0\rangle_A$. Фактически мы приготовили состояние, относительно которого нам точно известно, что оно было порождено начальным возбужденным атомным состоянием $|1\rangle_A$, — основное состояние не распадается.

С другой стороны, если мы не зарегистрировали фотон, а наш детектор обладает идеальной чувствительностью, то мы спроецировали окружение на состояние $|0\rangle_E$ и, следовательно, приготовили атомное состояние

$$a|0\rangle_A + b\sqrt{1-p}|1\rangle_B. \tag{3.137}$$

Ввиду *неудачи* в регистрации фотона становится более вероятным, что начальным атомным состоянием было основное!

Как уже отмечалось, унитарное преобразование, которое запутывает A с E вслед за ортогональным измерением E, может быть описано как ПОЗМ в A. Если $|\varphi\rangle_A$ изменяется как

$$|\varphi\rangle_A|0\rangle_E \to \sum_\mu \mathbf{M}_\mu|\varphi\rangle_A|\mu\rangle_E, \tag{3.138}$$

то ортогональное измерение в E, которое проецирует на базис $\{|\mu\rangle_E\}$, для каждого результата μ реализует ПОЗМ с

$$\text{Prob}(\mu) = \text{tr}\left(\mathbf{F}_\mu \rho_A\right), \qquad \mathbf{F}_\mu = \mathbf{M}_\mu^\dagger \mathbf{M}_\mu. \tag{3.139}$$

В случае канала затухания амплитуды находим:

$$\mathbf{F}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix}, \qquad \mathbf{F}_1 = \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix}, \tag{3.140}$$

где $\mathbf{F}_0$ определяет вероятность успешного детектирования фотона, а $\mathbf{F}_1$ — дополнительную к ней вероятность того, что фотон не зарегистрирован.

Если мы ожидаем в течение времени $t \gg \Gamma^{-1}$, так что p стремится к единице, наша ПОЗМ приближается к ортогональному измерению, измерению начального атомного состояния в базисе $\{|0\rangle_A, |1\rangle_A\}$. Необычной чертой этого измерения является то, что мы можем проецировать на состояние $|0\rangle_A$, *не регистрируя* фотон. Это пример того, что Дикке называл «измерением без взаимодействия» — наблюдая *отсутствие изменения* в состоянии окружения, мы делаем вывод, каким должно было быть атомное состояние. Термин «измерение без взаимодействия» является общеупотребительным, хотя он в некоторой степени вводит в заблуждение; очевидно, что если бы гамильтониан Вселенной не включал связь атома с электромагнитным полем, то измерение было бы невозможно.

3.5. Основное уравнение

3.5.1. Марковская эволюция

Формализм супероператоров предоставляет общее описание эволюции матрицы плотности, в том числе эволюции чистого состояния в смешанное (декогерентизация). В том же смысле, в каком унитарное преобразование дает общее описание когерентной квантовой эволюции. В последнем случае динамику квантовой системы удобно характеризовать *гамильтонианом*, описывающим эволюцию в бесконечно малом интервале времени. Тогда динамика описывается дифференциальным уравнением, *уравнением Шредингера*. Интегрируя это уравнение или, иначе говоря, складывая эволюции на множестве инфинитезимальных интервалов, мы можем рассчитать эволюцию в течение конечного интервала времени.

Часто, по крайней мере в хорошем приближении, оказывается возможным описание эволюции (не обязательно когерентной) матрицы плотности

дифференциальным уравнением. Это так называемое *основное уравнение* (master equation) будет нашей следующей темой.

В самом деле, непонятно, почему для описания декогерентизации необходимо дифференциальное уравнение. Такое описание возможно, если только эволюция квантовой системы является «марковской» или, другими словами, *локальной* во времени. Если эволюция во времени t оператора плотности $\rho(t)$ управляется дифференциальным уравнением (первого порядка), то это значит, что оператор $\rho(t+dt)$ полностью определяется оператором $\rho(t)$.

Мы видели, что всегда можем описать эволюцию оператора плотности ρ_A в гильбертовом пространстве $\mathcal{H}_A$, если предположить, что в расширенном гильбертовом пространстве $\mathcal{H}_A \otimes \mathcal{H}_B$ она в действительности является унитарной. Но, даже если эволюция в $\mathcal{H}_A \otimes \mathcal{H}_B$ управляется уравнением Шредингера, этого не достаточно, чтобы обеспечить *локальность* во времени эволюции $\rho_A(t)$. Действительно, если мы знаем только $\rho_A(t)$, мы не имеем полной системы начальных условий для уравнения Шредингера; кроме этого нам необходимо знать состояние «окружения». Так как из общей теории супероператоров известно, что мы вправе потребовать, что в момент времени $t = 0$ квантовым состоянием в пространстве $\mathcal{H}_A \otimes \mathcal{H}_B$ является

$$\rho_A \otimes |0\rangle_{E\,E}\langle 0|, \tag{3.141}$$

то наиболее ярким выражением этой трудности является то, что оператор плотности $\rho_A(t+dt)$ зависит не только от $\rho_A(t)$, но также и от ρ_A в более ранние моменты времени, поскольку резервуар E[1] некоторое время сохраняет память об этой информации и может вернуть ее обратно в систему A.

Это затруднение возникает вследствие того, что информация течет по улице с двухсторонним движением. Открытая система (классическая или квантовая) является *диссипативной*, поскольку информация может перетекать из системы в резервуар. Но это значит, что информация может также течь обратно из резервуара в систему, приводя к немарковским *флуктуациям* в системе[2].

Таким образом, за исключением случая когерентной (унитарной) эволюции, флуктуации неизбежны, а строго марковское описание квантовой динамики невозможно. Тем не менее во многих отношениях марковское описание является хорошим приближением. Ключевая идея здесь в том, что возможно разделение между типичным корреляционным временем флуктуа-

[1] Обсуждая основное уравнение, окружение обычно называют *резервуаром* в знак уважения к глубоко укоренившейся терминологии статистической физики.

[2] Эта неизбежная связь лежит в основе *флуктуационно-диссипационной теоремы*, мощного инструмента статистической физики.

аций и временным масштабом наблюдаемой нами эволюции. Грубо говоря, мы можем обозначить через $(\Delta t)_{\rm res}$ время, которое требуется резервуару, чтобы «забыть» полученную от системы информацию, — спустя время $(\Delta t)_{\rm res}$ мы можем считать, что информация навсегда потеряна, и пренебрегать возможностью того, что она вновь вернется, чтобы повлиять на дальнейшую эволюцию системы.

Наше описание эволюции системы будет включать в себя «сглаживание» («coarse graining») во времени: мы воспринимаем динамику сквозь фильтр, скрывающий высокочастотную часть движения с $\omega \gg (\Delta t)^{-1}_{\rm coarse}$. Тогда марковское описание должно быть приближенно справедливым, если $(\Delta t)_{\rm res} \ll (\Delta t)_{\rm coarse}$; мы можем пренебречь памятью резервуара, поскольку не в состоянии обнаружить ее влияние. Это «марковское приближение» *полезно*, если временной масштаб наблюдаемой нами динамики велик по сравнению с $(\Delta t)_{\rm coarse}$, например, если временной масштаб затухания $(\Delta t)_{\rm damp}$ удовлетворяет неравенству

$$(\Delta t)_{\rm damp} \gg (\Delta t)_{\rm coarse} \gg (\Delta t)_{\rm res}. \tag{3.142}$$

Это условие часто выполняется на практике, например, в атомной физике, где $(\Delta t)_{\rm res} \sim \hbar/kT \sim 10^{-14}$с ($T$ — температура) по порядку величины больше типичного времени жизни возбужденного состояния.

Поучительным примером является случай, в котором система A представляет собой один гармонический осциллятор $(\mathbf{H}_A = \omega \mathbf{a}^\dagger \mathbf{a})$, а резервуар R состоит из множества гармонических осцилляторов $(\mathbf{H}_R = \sum_i \omega_i \mathbf{b}_i^\dagger \mathbf{b}_i)$, слабо связанных с рассматриваемой системой возмущением

$$\mathbf{H}' = \sum_i \lambda_i (\mathbf{a}\mathbf{b}_i^\dagger + \mathbf{a}^\dagger \mathbf{b}_i). \tag{3.143}$$

Гамильтониан резервуара может, например, представлять (свободное) электромагнитное поле, тогда $\mathbf{H}'$ в низшем нетривиальном порядке теории возмущений индуцирует переходы, в которых осциллятор излучает или поглощает один фотон, при этом уменьшая или соответственно увеличивая свое число заполнения $\mathbf{n} = \mathbf{a}^\dagger \mathbf{a}$.

Мы могли бы подойти к основному уравнению, анализируя систему с помощью зависящей от времени теории возмущений, аккуратно вводя конечную обрезающую частоту. Детали этого анализа можно найти в книге Говарда Кармайкла[1]. Однако здесь я хотел бы обойтись без него и перепрыгнуть к основному уравнению более эвристическим путем.

[1] Howard Carmichael, *Open Systems Approach to Quantum Optics*, Springer Verlag, Berlin et al 1993. На русском языке см. Ю. Л. Климонтович *Статистическая теория открытых систем*, тт. 1–3, Янус-К М., 1995–2001; Ю. Л. Климонтович *Введение в физику открытых систем*, Янус-К М., 2002. — *Прим. ред.*

3.5.2. Линдбладиан

При унитарной эволюции изменение матрицы плотности во времени управляется уравнением Шредингера[1]

$$\dot{\boldsymbol{\rho}} = -i\big[\mathbf{H}, \boldsymbol{\rho}\big], \tag{3.144}$$

которое, при не зависящем от времени $\mathbf{H}$, можно формально решить и найти

$$\boldsymbol{\rho}(t) = e^{-i\mathbf{H}t}\boldsymbol{\rho}(0)e^{i\mathbf{H}t}. \tag{3.145}$$

Нашей целью является обобщение этого уравнения на случай марковской, но не унитарной, эволюции, в котором мы будем иметь

$$\dot{\boldsymbol{\rho}} = \mathcal{L}[\boldsymbol{\rho}]. \tag{3.146}$$

Линейный оператор $\mathcal{L}$, порождающий конечный супероператор в том же смысле, в каком гамильтониан $\mathbf{H}$ порождает унитарную эволюцию во времени, будет называться *линдбладианом*. Если $\mathcal{L}$ не зависит от времени, то формальное решение уравнения (3.146) имеет вид

$$\boldsymbol{\rho}(t) = e^{\mathcal{L}t}\boldsymbol{\rho}(0). \tag{3.147}$$

Чтобы вычислить линдбладиан, мы начинаем с уравнения Шредингера для системы, связанной с резервуаром

$$\dot{\boldsymbol{\rho}}_A = \mathrm{tr}_R\left(\dot{\boldsymbol{\rho}}_{AR}\right) = -i\,\mathrm{tr}_R\left(\big[\mathbf{H}_{AR}, \boldsymbol{\rho}_{AR}\big]\right), \tag{3.148}$$

но, как уже отмечалось, мы не ожидаем, что эта формула для $\dot{\boldsymbol{\rho}}_A$ может быть выражена лишь через $\boldsymbol{\rho}_A$. Чтобы найти линдбладиан, необходимо явно воспользоваться марковским приближением (как это делает Кармайкл). С другой стороны, предположим, что марковское приближение применимо. Мы уже знаем, что *наиболее общий* супероператор можно записать в представлении Крауса:

$$\boldsymbol{\rho}_A(t) = \$_t\big[\boldsymbol{\rho}(0)\big] = \sum_{\mu} \mathbf{M}_\mu(t)\boldsymbol{\rho}(0)\mathbf{M}_\mu^\dagger(t), \tag{3.149}$$

причем $\$_{t=0} = \mathbf{1}$. Если пролетевшее время является инфинитезимальным интервалом dt и

$$\boldsymbol{\rho}(dt) = \boldsymbol{\rho}(0) + O(dt), \tag{3.150}$$

[1] В статистической физике это уравнение принято называть квантовым уравнением Лиувилля, хотя, конечно, оно непосредственно выводится из уравнения Шредингера. — *Прим. ред.*

тогда одним из операторов Крауса будет $\mathbf{M}_0 = \mathbf{1} + O(dt)$, а все остальные будут иметь порядок $\sqrt{dt}$. Операторы $\mathbf{M}_\mu$ $(\mu > 0)$ описывают «квантовые скачки», которые с вероятностью порядка dt может совершать система. Следовательно, мы можем записать

$$\begin{aligned} \mathbf{M}_\mu &= \sqrt{dt}\mathbf{L}_\mu, \qquad \mu = 1, 2, 3, \ldots, \\ \mathbf{M}_0 &= \mathbf{1} + (-i\mathbf{H} + \mathbf{K})dt, \end{aligned} \tag{3.151}$$

где $\mathbf{H}$ и $\mathbf{K}$ эрмитовы, причем $\mathbf{L}_\mu$, $\mathbf{H}$ и $\mathbf{K}$ имеют нулевой порядок по dt. Фактически, оператор $\mathbf{K}$ можно определить, используя условие нормировки Крауса

$$\mathbf{1} = \sum_\mu \mathbf{M}_\mu^\dagger \mathbf{M}_\mu = \mathbf{1} + dt\left(2\mathbf{K} + \sum_{\mu>0} \mathbf{L}_\mu^\dagger \mathbf{L}_\mu\right), \tag{3.152}$$

или

$$\mathbf{K} = -\frac{1}{2}\sum_{\mu>0} \mathbf{L}_\mu^\dagger \mathbf{L}_\mu. \tag{3.153}$$

Подставляя это в уравнение (3.149), выражая $\boldsymbol{\rho}(dt) = \boldsymbol{\rho}(0) + \dot{\boldsymbol{\rho}}(0)dt$ и сравнивая слагаемые порядка dt, получим уравнение Линдблада[1]:

$$\dot{\boldsymbol{\rho}} - \mathcal{L}[\boldsymbol{\rho}] = -i[\mathbf{H}, \boldsymbol{\rho}] + \sum_{\mu>0}\left(\mathbf{L}_\mu \boldsymbol{\rho} \mathbf{L}_\mu^\dagger - \frac{1}{2}\mathbf{L}_\mu^\dagger \mathbf{L}_\mu \boldsymbol{\rho} - \frac{1}{2}\boldsymbol{\rho}\mathbf{L}_\mu^\dagger \mathbf{L}_\mu\right). \tag{3.154}$$

Первый член в $\mathcal{L}[\boldsymbol{\rho}]$ представляет собой обычное слагаемое Шредингера, генерирующее унитарную эволюцию. Остальные слагаемые описывают возможные переходы, которые может испытывать система, вследствие ее взаимодействия с резервуаром. Операторы $\mathbf{L}_\mu$ называются *операторами Линдблада* или *операторами квантовых скачков*. Каждое слагаемое $\mathbf{L}_\mu \boldsymbol{\rho} \mathbf{L}_\mu^\dagger$ индуцирует один из возможных квантовых скачков, тогда как слагаемые $-\frac{1}{2}\mathbf{L}_\mu^\dagger \mathbf{L}_\mu \boldsymbol{\rho} - \frac{1}{2}\boldsymbol{\rho}\mathbf{L}_\mu^\dagger \mathbf{L}_\mu$ необходимы для корректного описания тех случаев, когда скачки не возникают.

Уравнение Линдблада (3.154) и есть то, что мы искали, общая форма (вполне положительной) марковской эволюции матрицы плотности: то есть основное уравнение. Из представления Крауса, с которого мы начинали, следует, что уравнение Линдблада сохраняет матрицу плотности: $\boldsymbol{\rho}(t+dt)$ — матрица плотности, если таковой является $\boldsymbol{\rho}(t)$. Действительно, используя уравнение (3.154), можно непосредственно проверить,

[1] Уравнение Линдблада, описывающее марковскую эволюцию матрицы плотности открытой системы, получено в работе G. Lindblad, *On the Generators of Quantum Dynamical Semigroups*, Commun. Math. Phys., **48**, 119–130 (1976). – *Прим. ред.*

что $\dot{\rho}$ эрмитов, а $\mathrm{tr}\,\dot{\rho} = 0$. То, что $\mathcal{L}[\rho]$ сохраняет положительность, несколько менее очевидно, но, как уже отмечалось, следует из представления Крауса.

Если мы вспомним связь между представлением Крауса и унитарным представлением супероператора, то интерпретацию основного уравнения можно сделать более прозрачной. Представим, что мы непрерывно контролируем резервуар, проецируя его в каждый момент времени на базис $|\mu\rangle_R$. С вероятностью $1 - O(dt)$ резервуар остается в состоянии $|0\rangle_R$, а с вероятностью порядка dt он совершает скачок в одно из состояний $|\mu\rangle_R$ ($\mu > 0$). Говоря, что резервуар «забыл» информацию, полученную от системы (так что применимо марковское приближение), мы считаем, что эти переходы происходят с вероятностями, линейно растущими со временем. Напомним, что это *не следует* автоматически из зависящей от времени теории возмущений. На малых временах t вероятности отдельных переходов пропорциональны t^2; мы получаем темп (дифференцируя «золотое правило Ферми») только после суммирования по непрерывному континууму возможных конечных состояний. Поскольку количество доступных состояний в действительности убывает как $1/t$, просуммированная по конечным состояниям вероятность перехода пропорциональна t. Используя марковское описание динамики, мы явно предполагали, что масштаб времени $(\Delta t)_{\mathrm{coarse}}$ настолько велик, что мы можем приписать частоты различным возможным переходам, которые могут быть обнаружены, пока мы контролируем окружение системы (резервуар). В действительности это следует из требования $(\Delta t)_{\mathrm{coarse}} \gg (\Delta t)_{\mathrm{res}}$.

3.5.3. Затухающий гармонический осциллятор

В качестве примера, иллюстрирующего основное уравнение, рассмотрим взаимодействующий с электромагнитным полем гармонический осциллятор

$$\mathbf{H}' = \sum_i \lambda_i (\mathbf{a}\mathbf{b}_i^\dagger + \mathbf{a}^\dagger \mathbf{b}_i). \tag{3.155}$$

Предположим, что температура резервуара равна нулю; тогда будет наблюдаться падение уровня возбуждения осциллятора, сопровождающееся последовательным излучением фотонов, но поглощения фотонов происходить не будет. Следовательно, имеется только один оператор скачка:

$$\mathbf{L}_1 = \sqrt{\Gamma}\mathbf{a}. \tag{3.156}$$

Здесь Γ представляет собой темп распада первого возбужденного ($n = 1$) состояния осциллятора в основное ($n = 0$) состояние; в соответствии со

структурой гамильтониана $\mathbf{H}'$ темп затухания в результате перехода с n-го уровня на $(n-1)$-й равен $n\Gamma$.[1] Основное уравнение в форме Линдблада приобретает вид

$$\dot{\boldsymbol{\rho}} = -i[\mathbf{H}_0, \boldsymbol{\rho}] + \Gamma\left(\mathbf{a}\boldsymbol{\rho}\mathbf{a}^\dagger - \frac{1}{2}\mathbf{a}^\dagger\mathbf{a}\boldsymbol{\rho} - \frac{1}{2}\boldsymbol{\rho}\mathbf{a}^\dagger\mathbf{a}\right), \tag{3.157}$$

где $\mathbf{H}_0 = \omega\mathbf{a}^\dagger\mathbf{a}$ — гамильтониан осциллятора. Это то же самое уравнение, что и полученное Кармайклом с помощью более изощренного анализа. (Мы не учли здесь только *лэмбовский сдвиг*, или радиационную перенормировку частоты осциллятора, имеющую тот же порядок, что и слагаемые скачков в $\mathcal{L}[\boldsymbol{\rho}]$.)

Слагаемые скачков в основном уравнении описывают *затухание* осциллятора вследствие излучения им фотонов[2]. Чтобы исследовать влияние скачков, удобно перейти к *представлению взаимодействия*; определим операторы $\boldsymbol{\rho}_I$ и $\mathbf{a}_I$ в представлении взаимодействия

$$\begin{aligned}\boldsymbol{\rho}(t) &= e^{-i\mathbf{H}_0 t}\boldsymbol{\rho}_I(t)e^{i\mathbf{H}_0 t},\\ \mathbf{a}(t) &= e^{-i\mathbf{H}_0 t}\mathbf{a}_I(t)e^{i\mathbf{H}_0 t},\end{aligned} \tag{3.158}$$

так что

$$\dot{\boldsymbol{\rho}}_I = \Gamma\left(\mathbf{a}_I\boldsymbol{\rho}_I\mathbf{a}_I^\dagger - \frac{1}{2}\mathbf{a}_I^\dagger\mathbf{a}_I\boldsymbol{\rho}_I - \frac{1}{2}\boldsymbol{\rho}_I\mathbf{a}_I^\dagger\mathbf{a}_I\right), \tag{3.159}$$

где фактически $\mathbf{a}_I(t) = \mathbf{a}e^{-i\omega t}$, следовательно, в правой части уравнения (3.159) можно заменить $\mathbf{a}_I$ на $\mathbf{a}$. В отсутствии затухания переменная $\tilde{\mathbf{a}} = e^{-i\mathbf{H}_0 t}\mathbf{a}e^{i\mathbf{H}_0 t} = \mathbf{a}e^{i\omega t}$ остается постоянной. При наличии затухания $\tilde{\mathbf{a}}$ изменяется в соответствии с уравнением

$$\frac{d}{dt}\langle\tilde{\mathbf{a}}\rangle = \frac{d}{dt}\operatorname{tr}(\mathbf{a}\boldsymbol{\rho}_I) = \operatorname{tr}\mathbf{a}\dot{\boldsymbol{\rho}}, \tag{3.160}$$

а из (3.159) мы имеем

$$\begin{aligned}\operatorname{tr}\mathbf{a}\dot{\boldsymbol{\rho}} &= \Gamma\operatorname{tr}\left(\mathbf{a}^2\boldsymbol{\rho}_I\mathbf{a}^\dagger - \frac{1}{2}\mathbf{a}\mathbf{a}^\dagger\mathbf{a}\boldsymbol{\rho}_I - \frac{1}{2}\mathbf{a}\boldsymbol{\rho}_I\mathbf{a}^\dagger\mathbf{a}\right) =\\ &= \Gamma\operatorname{tr}\left(\frac{1}{2}[\mathbf{a}^\dagger, \mathbf{a}]\mathbf{a}\boldsymbol{\rho}_I\right) = -\frac{\Gamma}{2}\operatorname{tr}(\mathbf{a}\boldsymbol{\rho}_I) = -\frac{\Gamma}{2}\langle\tilde{\mathbf{a}}\rangle.\end{aligned} \tag{3.161}$$

[1] n-е возбужденное состояние осциллятора может интерпретироваться как состояние n невзаимодействующих частиц; его темп затухания равен $n\Gamma$, поскольку исчезнуть при этом может любая из n частиц (квантов).

[2] Эта модель распространяет наше обсуждение канала затухания амплитуды, скорее на затухающий осциллятор, а не на затухающий кубит.

Интегрируя это уравнение, получим

$$\langle \tilde{\mathbf{a}}(t) \rangle = e^{-\Gamma t/2} \langle \tilde{\mathbf{a}}(0) \rangle. \tag{3.162}$$

Аналогично, число заполнения осциллятора $\mathbf{n} = \mathbf{a}^\dagger \mathbf{a} = \tilde{\mathbf{a}}^\dagger \tilde{\mathbf{a}}$ затухает согласно

$$\begin{aligned} \frac{d}{dt}\langle \mathbf{n} \rangle = \frac{d}{dt}\langle \mathbf{a}^\dagger \mathbf{a} \rangle = \operatorname{tr}(\mathbf{a}^\dagger \mathbf{a} \boldsymbol{\rho}_I) = \\ = \Gamma \operatorname{tr} \left(\mathbf{a}^\dagger \mathbf{a}^2 \boldsymbol{\rho}_I \mathbf{a}^\dagger - \frac{1}{2} \mathbf{a}^\dagger \mathbf{a} \mathbf{a}^\dagger \mathbf{a} \boldsymbol{\rho}_I - \frac{1}{2} \mathbf{a}^\dagger \mathbf{a} \boldsymbol{\rho}_I \mathbf{a}^\dagger \mathbf{a} \right) = \\ = \Gamma \operatorname{tr} \mathbf{a}^\dagger [\mathbf{a}^\dagger, \mathbf{a}] \mathbf{a} \boldsymbol{\rho}_I = -\Gamma \operatorname{tr} \mathbf{a}^\dagger \mathbf{a} \boldsymbol{\rho}_I = -\Gamma \langle \mathbf{n} \rangle, \end{aligned} \tag{3.163}$$

что после интегрирования дает

$$\langle \tilde{\mathbf{n}}(t) \rangle = e^{-\Gamma t} \langle \tilde{\mathbf{n}}(0) \rangle. \tag{3.164}$$

Таким образом, Γ представляет собой темп затухания осциллятора. Мы можем интерпретировать n-е возбужденное состояние осциллятора как состояние n невзаимодействующих частиц, каждая из которых распадается с отнесенной к единице времени вероятностью Γ; следовательно, уравнение (3.164) и есть тот самый экспоненциальный закон, которому удовлетворяет численность популяции распадающихся частиц.

Более интересно то, что говорит основное уравнение о декогерентизации. Детали этого анализа будут в домашнем задании. А здесь мы проанализируем более простую задачу — осциллятор, испытывающий затухание фазы.

3.5.4. Затухание фазы

Чтобы смоделировать затухание фазы гармонического осциллятора, возьмем другую связь осциллятора с резервуаром:

$$\mathcal{H}' = \left(\sum_i \lambda_i \mathbf{b}_i^\dagger \mathbf{b}_i \right) \mathbf{a}^\dagger \mathbf{a}. \tag{3.165}$$

Таким образом, существует только один оператор Линдблада, а основное уравнение в представлении взаимодействия имеет вид

$$\dot{\boldsymbol{\rho}}_I = \Gamma \left(\mathbf{a}^\dagger \mathbf{a} \boldsymbol{\rho}_I \mathbf{a}^\dagger \mathbf{a} - \frac{1}{2} (\mathbf{a}^\dagger \mathbf{a})^2 \boldsymbol{\rho}_I - \frac{1}{2} \boldsymbol{\rho}_I (\mathbf{a}^\dagger \mathbf{a})^2 \right). \tag{3.166}$$

Здесь Γ может интерпретироваться как частота (отнесенная к единице времени вероятность), с которой фотоны резервуара *рассеиваются* осциллятором, находящимся в первом возбужденном состоянии. Если число заполнения равно n, то частота рассеяния становится равной $n^2\Gamma$. Причина появления множителя n^2 состоит в том, что все вклады в амплитуду рассеяния от каждой из n осцилляторных «частиц» складываются когерентно; амплитуда пропорциональна n, а частота (темп) – n^2.

Уравнение для $\boldsymbol{\rho}_I$ (3.166) легко решить в базисе чисел заполнения. Разлагая

$$\boldsymbol{\rho}_I = \sum_{n,m} \rho_{nm} |n\rangle\langle m| \tag{3.167}$$

(где $\mathbf{a}^\dagger \mathbf{a}|n\rangle = n|n\rangle$), запишем основное уравнение в виде

$$\dot{\rho}_{nm} = \Gamma \left(nm - \frac{1}{2}n^2 - \frac{1}{2}m^2 \right) \rho_{nm} = -\frac{\Gamma}{2}(n-m)^2 \rho_{nm}. \tag{3.168}$$

Его интегрирование дает

$$\rho_{nm}(t) = \rho_{nm}(0) \exp\left[-\frac{1}{2}\Gamma t (n-m)^2 \right]. \tag{3.169}$$

Если мы приготовим подобную «кот-состоянию» суперпозицию собственных состояний оператора чисел заполнения с большой разницей значений n

$$|\mathrm{cat}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |n\rangle \right), \qquad n \gg 1, \tag{3.170}$$

то недиагональные элементы матрицы плотности будут затухать как $\exp\left(-\Gamma n^2 t/2\right)$. Фактически это точно такой же тип поведения, что и обнаруженный нами при анализе затухания фазы одного кубита. Темп декогерентизации равен $n^2\Gamma$, так как он равен частоте рассеяния фотонов резервуара осциллятором, возбужденным в состояние $|n\rangle$. Как и ранее, мы видим также, что декогерентизация фазы выбирает предпочтительный базис. Она возникает в базисе собственных состояний оператора чисел заполнения, поскольку это именно тот оператор, который входит в связь осциллятора с резервуаром $\mathbf{H}'$.

Вернемся к затуханию амплитуды. Поскольку в нашей модели затухания амплитуды в связь осциллятора с резервуаром $\mathbf{H}'$ входит оператор уничтожения $\mathbf{a}$ (и сопряженный ему оператор рождения $\mathbf{a}^\dagger$), то можно предположить, что декогерентизация возникает в базисе собственных состояний

оператора $\mathbf{a}$. *Когерентное состояние*

$$|\alpha\rangle = e^{-|\alpha|^2/2} e^{\alpha \mathbf{a}^\dagger} |0\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{3.171}$$

представляет собой собственное состояние оператора $\mathbf{a}$, отвечающее собственному значению α. Два когерентных состояния с разными собственными значениями α_1 и α_2 не ортогональны друг другу:

$$|\langle \alpha_1 | \alpha_2 \rangle|^2 = e^{-|\alpha_1|^2} e^{-|\alpha|_2^2} e^{2\mathrm{Re}(\alpha_1^* \alpha_2)} = \exp\left(-|\alpha_1 - \alpha_2|^2\right), \tag{3.172}$$

следовательно, перекрытие очень мало при большой величине $|\alpha_1 - \alpha_2|^2$.

Представим, что мы приготовили «кот-состояние»

$$|\mathrm{cat}\rangle = \frac{1}{\sqrt{2}} \left(|\alpha_1\rangle + |\alpha_2\rangle \right), \tag{3.173}$$

суперпозицию когерентных состояний с $|\alpha_1 - \alpha_2| \gg 1$. Вы покажете (в домашнем упражнении), что недиагональные элементы матрицы плотности $\boldsymbol{\rho}$ затухают как

$$\exp\left(-\frac{\Gamma t}{2} |\alpha_1 - \alpha_2|^2 \right) \tag{3.174}$$

(при $\Gamma t \ll 1$). Таким образом, темп декогерентизации

$$\Gamma_{\mathrm{dec}} = \frac{1}{2} |\alpha_1 - \alpha_2|^2 \Gamma_{\mathrm{damp}} \tag{3.175}$$

огромен по сравнению с темпом затухания. Такое поведение также легко интерпретируется. В когерентном состоянии ожидаемое значение числа заполнения равно $\langle \alpha | \mathbf{a}^\dagger \mathbf{a} | \alpha \rangle = |\alpha|^2$. Следовательно, если $\alpha_{1,2}$ сравнимы по модулю, но имеют существенно разные фазы (как в суперпозиции волновых пакетов с минимальной неопределенностью, центрированных в точках $+x$ и $-x$), темп декогерентизации имеет порядок темпа эмиссии *одного* фотона. Он очень велик по сравнению с темпом диссипации значительной части энергии осциллятора.

Аналогично можно рассмотреть осциллятор, связанный с резервуаром, находящимся при конечной температуре. Вновь темп декогерентизации будет иметь порядок частоты излучения или поглощения одного фотона, но теперь она гораздо выше, чем при нулевой температуре. Поскольку фотонные моды с частотой, сравнимой с частотой осциллятора ω, имеют термически равновесное число заполнения

$$n_\gamma = \frac{T}{\hbar \omega} \tag{3.176}$$

(при $T \gg \hbar\omega$), то интенсивность взаимодействия увеличивается множителем n_γ. Тогда мы имеем

$$\frac{\Gamma_{\text{dec}}}{\Gamma_{\text{damp}}} \sim n_{\text{osc}} n_\gamma \sim \frac{E}{\hbar\omega}\frac{T}{\hbar\omega} \sim \frac{m\omega_2 x^2}{\hbar\omega}\frac{T}{\hbar\omega} \sim x^2\frac{mT}{\hbar^2} \sim \frac{x^2}{\lambda_T^2}, \qquad (3.177)$$

где x – амплитуда осцилляций, а λ_T – тепловая длина волны де Бройля. Декогерентизация протекает *очень быстро*.

3.6. В чем проблема? (Здесь есть проблема?)

Наш обзор оснований квантовой теории почти завершен. Но прежде чем мы займемся своим главным делом, кратко проанализируем состояние этих оснований. Находится ли квантовая теория в «хорошей форме» или в ее корнях имеются фундаментальные проблемы, до сих пор требующие своего решения?

Одной такой потенциально серьезной проблемой, впервые упомянутой в § 2.1, является *проблема измерения*. Мы отмечали странный дуализм, присущий аксиомам квантовой теории. Существует два способа изменения квантового состояния: *детерминистская унитарная эволюция* и *вероятностное измерение*. Но почему измерение должно принципиально отличаться от любого другого физического процесса? Этот дуализм вселяет в некоторых людей подозрение, что современная формулировка квантовой теории все еще не полна.

В этой главе мы многое узнали об измерениях. В § 3.1.1 мы обсудили, как унитарная эволюция может привести к появлению корреляций (запутывания) между системой и «переменной-указателем» измерительного прибора. Таким образом, чистое состояние системы может эволюционировать в смешанное (после взятия следа по состояниям «указателя»), которое допускает интерпретацию как *ансамбля* взаимно ортогональных чистых состояний (собственных состояний оператора плотности рассматриваемой системы), каждое из которых возникает с вполне определенной вероятностью. Таким образом, уже в этом простом высказывании заложены зерна более глубокого понимания того, как исключительно в рамках унитарной эволюции может возникнуть «коллапс» (редукция) вектора состояния. С другой стороны, в § 2.5 мы говорили о том, что интерпретация матрицы плотности как ансамбля неоднозначна. В § 2.5.5 мы особенно ясно видели, что если мы способны измерить «указатель» в любом понравившемся нам базисе, то мы можем приготовить систему в любом из множества «экзотических» состояний, суперпозиций собственных состояний системы ρ (теоре-

ма ЖХЙВ). Следовательно, редукция (*разрушающая* относительные фазы состояний в суперпозиции) не может быть объяснена одним только запутыванием.

В § 3.4 и § 3.5 мы изучали другой важный аспект процесса измерения — *декогерентизацию*. Главная идея состоит в том, что в случае макроскопических систем мы не можем надеяться уследить за всеми микроскопическими степенями свободы. Нам приходится довольствоваться *сглаженным* (*coarse-grained*) описанием, получающимся в результате взятия следа по множеству ненаблюдаемых переменных. В случае макроскопического измерительного прибора мы должны взять след по степеням свободы окружения, с которым прибор неизбежно взаимодействует. Тогда мы обнаружим, что прибор исключительно быстро релаксирует в некоторый предпочтительный базис, определяемый природой связи прибора с его окружением. Похоже, что особенностью гамильтониана Вселенной является то, что фундаментальные взаимодействия хорошо локализованы в пространстве, следовательно, избираемый в процессе декогерентизации базис также хорошо локализован в пространстве. Кот или жив или мертв, а не в суперпозиции состояний $1/\sqrt{2}(|\text{alive}\rangle + |\text{dead}\rangle)$.

Вычисляя след по степеням свободы окружения, мы получаем более полную картину процесса измерения, «редукции». Наша система запутывается с прибором, который, в свою очередь, запутан с окружением. Если мы рассматриваем микросостояние окружения, как недоступное в любой момент времени, то мы вправе говорить, что измерение состоялось. Относительные фазы базисных состояний системы безвозвратно потеряны — ее вектор состояния коллапсировал.

Конечно, с принципиальной точки зрения никакой реальной потери информации о фазах нет. Эволюция системы+прибора+окружения является унитарной и детерминистской. В принципе мы, вероятно, могли бы выполнить в высшей степени нелокальное измерение окружающей среды и восстановить якобы разрушенную фазовую информацию о системе. В этом смысле наше объяснение коллапса, по выражению Белла, годится только «для всех практических целей» (FAPP: «for all practical purposes»). После «измерения» когерентность системы базисных состояний в принципе могла бы быть восстановлена (мы могли бы обратить измерение с помощью «квантового удаления»), но осуществление такого измерения в высшей степени невероятно. В самом деле, коллапс имеет место только «для всех практических целей» (хотя, вероятно, мы могли бы доказать в космологическом смысле, что некоторые измерения действительно принципиально необратимы), но существует ли то, что достойно быть не «для всех практических целей»?

Нашей целью в физике является объяснение наблюдаемых явлений на основе как можно более простых моделей. Не нужно постулировать два фундаментальных процесса (унитарная эволюция *и* измерение), если существенным является только один из них (унитарная эволюция). Тогда примем, по крайней мере временно, такую гипотезу:

Эволюция *замкнутой* квантовой системы *всегда унитарна*.

Конечно, мы видели, что не все супероператоры унитарны. Суть гипотезы в том, что неунитарная эволюция *открытой* системы, в том числе и происходящая в процессе измерения редукция, всегда возникает в результате игнорирования некоторых степеней свободы большей системы. Эта точка зрения была провозглашена Хьюго Эвереттом в 1957 г[1]. Согласно ей эволюция квантового состояния Вселенной является действительно детерминистской!

Но даже если мы согласимся с тем, что редукция объясняется декогерентизацией в системе, то есть на самом деле является детерминированной, мы не избавимся от всех загадок квантовой теории. Для волновой функции Вселенной фактически существует суперпозиция состояния, в котором кот мертв, и состояния, в котором кот жив. Несмотря на это, всякий раз, когда я наблюдаю за котом, он либо жив, либо мертв. Оба исхода возможны, но только один из них реализуется в действительности. Почему это так?

Ваш ответ на этот вопрос может зависеть от вашего понимания квантовой теории. Существует (по меньшей мере) два приемлемых направления рассуждений.

Платоник: Физика описывает *реальность*. В квантовой теории «волновая функция Вселенной» представляет полное описание физической реальности.

Позитивист: Физика описывает наши *ощущения*. Волновая функция кодирует состояние наших знаний, а задача квантовой теории – дать по возможности наилучшие предсказания относительно будущего на основе текущего уровня наших знаний.

Я верю в реальность. Я думаю, что мои доводы прагматичны. Как физик, я стремлюсь к наиболее экономичной модели, «объясняющей» то, что я воспринимаю. По крайней мере для физика, простейшим предположением является то, что мои (и ваши) ощущения скоррелированы с лежащей в их

[1] H. Everett, III *"Relative State" Formulation of Quantum Mechanics*, Rev. Mod. Phys., **29**, 454–462 (1957). — *Прим. ред.*

основе внешней по отношению ко мне реальностью. Серьезному философу эта онтология может показаться безнадежно наивной. Однако я предпочитаю верить в реальность, поскольку это предположение выглядит простейшим из тех, что могли бы успешно объяснить мои ощущения. (В подобном же духе я предпочитаю верить, что наука представляет нечто большее, чем просто консенсус. Я верю, что наука способствует прогрессу и приближает нас к удовлетворительному пониманию Природы — законы физики открыты, а не придуманы. Я верю в это, потому что это наиболее простое объяснение того, почему ученые так легко приходят к взаимопониманию.)

Те, кто придерживается другой точки зрения (даже если существует объективная реальность, вектор состояния описывает не ее, а всего лишь уровень наших знаний о ней), склонны считать, что современная формулировка квантовой теории не вполне удовлетворительна, что существует более глубокое описание, все еще ждущее своего открытия. Пока вы не сможете убедить меня в обратном, мне представляется более разумным предполагать, что волновая функция дает описание реальности.

Если мы полагаем, что волновая функция описывает реальность, и *если* принимаем точку зрения Эверетта, что вся эволюция является унитарной, *то* мы обязаны признать, что все возможные исходы измерения имеют одинаковое право быть «реальными». Как тогда понять, почему в эксперименте реализуется только *один* результат — кот или жив или мертв.

На самом деле здесь нет никакого парадокса, если только мы (в духе интерпретации Эверетта) готовы включить и себя в квантовую систему, описываемую волновой функцией. Эта волновая функция описывает все возможные корреляции между подсистемами, в том числе между котом и состоянием моего сознания. Если мы приготовили «кот-состояние», а затем наблюдаем за ним, то оператор плотности (после взятия следа по всем внешним степеням свободы) приобретает вид

$$\begin{aligned} &|\text{decay}\rangle_{\text{atom}}|\text{dead}\rangle_{\text{cat}}|\text{know it's dead}\rangle_{\text{me}}, \qquad \left(\text{Prob} = \frac{1}{2}\right), \\ &|\text{no decay}\rangle_{\text{atom}}|\text{alive}\rangle_{\text{cat}}|\text{know it's alive}\rangle_{\text{me}}, \qquad \left(\text{Prob} = \frac{1}{2}\right). \end{aligned} \tag{3.178}$$

Эта матрица плотности ρ описывает две альтернативы, но в обоих случаях я имею точное знание о состоянии здоровья кота. Я *никогда* не вижу его полуживым-полумертвым. (В соответствии с опытом, я нахожусь в собственном состоянии «оператора определенности».)

Допуская, что волновая функция описывает реальность и что вся эволюция является унитарной, мы приходим к интерпретации квантовой теории на основе «множественности миров». В этой картине всякий раз, когда

совершается «измерение», волновая функция Вселенной «расщепляется» на две ветви, соответствующие двум возможным исходам. После множества измерений существует множество ветвей (множество миров), каждая из которых с одинаковым правом может описывать реальность. Это размножение миров выглядит насмешкой над нашим намерением разработать наиболее экономичное описание. Но мы следуем одной конкретной ветви и для предсказания того, что мы увидим в следующий момент, множество других миров не имеет значения. Размножение миров ничего нам не стоит. «Множественность миров» может показаться странной, но стоит ли удивляться тому, что *полное* описание реальности — нечто находящегося полностью за пределами нашего опыта, кажется нам странным?

Включив себя в реальность, описываемую волновой функцией, мы поняли, почему мы воспринимаем определенный результат измерения, но по-прежнему стоит следующий вопрос: «Каким образом в этот (детерминистский) формализм входит понятие *вероятности*?» Этот вопрос продолжает беспокоить, для ответа на него мы должны быть готовы точно сформулировать, что значит «с вероятностью»?

Слово «вероятность» используется в двух различных смыслах. Иногда *вероятность* означает *частоту*. Мы говорим, что вероятность того, что монета выпадет орлом вверх равна $1/2$, если мы ожидаем, что при многократном подбрасывании монеты число выпадений орла, деленное на полное число подбрасываний, сходится к $1/2$. (Это, однако, ненадежное понятие; даже если вероятность равна $1/2$, монета все равно *может* выпасть орлом триллион раз подряд.) При строгом математическом обсуждении теория вероятностей часто формулируется как раздел теории меры — она занимается свойствами бесконечных последовательностей.

Но в повседневной практике, а также в квантовой теории, вероятности обычно *не* являются частотами. Когда мы выполняем измерение, мы не можем повторить его бесконечное число раз на идентично приготовленных системах. С точки зрения Эверетта, или с космологической, существует только одна Вселенная, а не множество одинаково приготовленных ее копий.

Так что же такое вероятность? На практике это число, которое дает количественное определение достоверности утверждения при данном состоянии знаний. Возможно, это удивительно, что такое представление можно положить в основу хорошо определенной математической теории, иногда называемой «бейесовским» подходом к вероятности. Термин «бейесовский» отражает теоретико-вероятностный метод, обычно используемый (в науке и в повседневной практике) для проверки гипотезы при наличии некоторых результатов наблюдения. Проверка гипотезы выполняется с ис-

пользованием правила Бейеса для условной вероятности:

$$P(A_0|B) = P(B|A_0)P(A_0)/P(B). \tag{3.179}$$

Предположим, например, что A_0 — приготовление частного квантового состояния, а B — частный результат измерения состояния. Мы выполнили измерение (получение B) и теперь хотим сделать заключение о том, какое состояние было приготовлено (вычислить $P(A_0|B)$). Квантовая механика позволяет нам вычислить $P(B|A_0)$, но она ничего не говорит о $P(A_0)$ (или $P(B)$). Мы делаем предположение относительно $P(A_0)$, что возможно, если принять «принцип безразличия»: если неизвестно, что более или менее вероятно, A_i или A_j, то предполагается, что $P(A_i) = P(A_j)$. Как только множество приготовлений определено, мы можем вычислить

$$P(B) = \sum_i P(B|A_i)P(A_i) \tag{3.180}$$

и, следовательно, применяя правило Бейеса, получить $P(A_0|B)$.

Но если мы будем считать, что теория вероятностей дает количественное определение достоверности при данном состоянии знаний, то мы обязаны спросить «при состоянии чьих знаний»? Чтобы построить объективную теорию, мы должны интерпретировать вероятность в квантовой теории не как предсказание, основанное на нашем *текущем* состоянии знаний, а скорее как предсказание, основанное на самом полном *возможном* знании о квантовом состоянии. Если мы готовим состояние $|\uparrow_x\rangle$, а измеряем σ_3, то мы говорим, что с вероятностью $1/2$ результатом является $|\uparrow_z\rangle$, не потому, что это лучшее предсказание, которое можно сделать, опираясь на то, что нам известно, а потому, что это лучшее предсказание, которое *кто-либо* может сделать, независимо от того, как много он знает. В этом смысле результат истинно *случайный*; его невозможно предсказать с уверенностью, даже если наше знание является полным (в противоположность псевдослучайности, возникающей в классической физике вследствие неполноты наших знаний).

Как же теперь нам извлечь вероятности из детерминистской Вселенной Эверетта? Вероятности возникают, потому что мы (часть системы) не можем с уверенностью предсказать наше будущее. Я знаю формализм, мне известны гамильтониан и волновая функция Вселенной, я знаю свою ветвь волновой функции. Теперь я собираюсь следить за котом. Мгновение спустя я буду определенно знать, что кот мертв, или я буду уверен в том, что он жив. Даже со всеми своими знаниями я не могу предсказать будущее. Даже имея полное знание о настоящем, невозможно сказать, каким будет состояние моего знания после того, как я понаблюдаю за котом. Самое лучшее,

что я могу, это приписать вероятности результатам. Итак, несмотря на то, что волновая функция Вселенной детерминистская, я, как часть системы, не способен на большее, чем делать вероятностные предсказания.

Конечно, главным героем этой истории является *декогерентизация*. Мы можем последовательно приписать вероятности альтернативам Dead и Alive, если только интерференция между ними невозможна (или по крайней мере пренебрежима). Вероятности имеют смысл, только когда мы можем исчерпывающим образом идентифицировать множества взаимно исключающих альтернатив. Поскольку это сложный вопрос, реально ли возникновение интерференции в более позднее время, мы не можем решить, приемлема ли теория вероятностей, рассматривая квантовое состояние в данный момент времени; мы должны проверить множество взаимно исключающих (сглаженных) историй или последовательностей событий. Существует утонченная техника («функционалы декогерентизации») определения того, являются ли различные истории в достаточной степени некогерентными, чтобы им можно было корректно приписать вероятности.

Итак, позицию Эверетта *можно* примирить с наблюдаемым квантовым индетерминизмом, однако, насколько я понимаю, в этой картине остается тревожащее белое пятно. Я собираюсь наблюдать за котом, и я знаю, что мгновение спустя матрица плотности примет вид

$$\begin{aligned} &|\text{dead}\rangle_{\text{cat}}|\text{know it's dead}\rangle_{\text{me}}, \qquad \text{Prob} = p_{\text{dead}}, \\ &|\text{alive}\rangle_{\text{cat}}|\text{know it's alive}\rangle_{\text{me}}, \qquad \text{Prob} = p_{\text{alive}}. \end{aligned} \tag{3.181}$$

Но как я узнаю, что p_{dead} и p_{alive} действительно являются теми вероятностями, которые я (в моей бейесовской картине) могу приписать своим будущим ощущениям? Мне *по-прежнему* необходимо правило преобразования этого оператора плотности в приписываемые альтернативам вероятности. *Предположение* о таком правиле выглядит противоречащим философии Эверетта; мы предпочли бы сказать, что единственным правилом, необходимым для формулировки теории, является уравнение Шредингера (и, возможно, предписание, указывающее начальную волновую функцию). Постулирование формулы вероятности находится в опасной близости к допущению, что, в конце концов, существует недетерминированный процесс измерения. Это сложная, касающаяся фундамента теории, проблема, полностью удовлетворительного решения которой я не знаю.

Поскольку, касаясь природы вероятности в квантовой теории, мы не в состоянии полностью избавиться от замешательства, может быть, полезно прокомментировать интересное предложение Хартла. Чтобы осуществить его предложение, мы должны вернуться (возможно, с сожалением) к ча-

стотной интерпретации вероятности. Идея Хартла состоит в том, что нам не нужно считать интерпретацию вероятности как часть постулата об измерении. На самом деле достаточно сделать более слабое предположение:

> Если мы готовим квантовое состояние $|a\rangle$ такое, что $\mathbf{A}|a\rangle = a|a\rangle$, и сразу вслед за этим измеряем $\mathbf{A}$, то результатом измерения является a.

Это выглядит как предположение о том, что во Вселенной Эверетта действует бейесовский подход. Я собираюсь измерить наблюдаемую, и волновая функция будет ветвиться, но если наблюдаемая имеет *одно и то же* значение в каждой ветви, то я *могу* предсказать результат.

Чтобы реализовать частотную интерпретацию вероятности, нам следует, строго говоря, рассмотреть бесконечное множество испытаний. Допустим, мы хотим сделать утверждение относительно вероятности получения результата $|\uparrow_z\rangle$ при измерении $\boldsymbol{\sigma}_3$ в состоянии

$$|\psi\rangle = a|\uparrow_z\rangle + b|\downarrow_z\rangle. \tag{3.182}$$

Тогда мы должны представить, что приготовлено бесконечное число копий, то есть состояние имеет вид

$$|\psi^{(\infty)}\rangle \equiv \big(|\psi\rangle\big)^{\infty} = |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle \otimes \cdots, \tag{3.183}$$

и мы мысленно представляем измерение $\boldsymbol{\sigma}_3$ в каждой из копий. Формально случай бесконечного числа испытаний можно сформулировать как N испытаний в пределе $N \to \infty$.

Идея Хартла состоит в том, чтобы рассматривать оператор «среднего спина»

$$\bar{\boldsymbol{\sigma}}_3 = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\sigma}_3^{(i)}, \tag{3.184}$$

и доказывать, что $\big(|\psi\rangle\big)^N$ при $N \to \infty$ стремится к *собственному состоянию* оператора $\bar{\boldsymbol{\sigma}}_3$ с собственным значением $|a|^2 - |b|^2$. Тогда мы можем, ссылаясь на слабый постулат измерения, сделать вывод, что измерение $\bar{\boldsymbol{\sigma}}_3$ наверняка даст результат $|a|^2 - |b|^2$, и, следовательно, $|a|^2$ равно той части наших спинов, которые ориентированы вверх. В этом смысле, $|a|^2$ является вероятностью того, что измерение $\boldsymbol{\sigma}_3$ дает результат $|\uparrow_z\rangle$.

Рассмотрим в качестве примера частный случай

$$|\psi_x^{(N)}\rangle \equiv \big(|\uparrow_x\rangle\big)^N = \left[\frac{1}{\sqrt{2}}\big(|\uparrow_z\rangle + |\downarrow_z\rangle\big)\right]^N. \tag{3.185}$$

Мы можем вычислить

$$\langle\psi_x^{(N)}|\bar{\sigma}_3|\psi_x^{(N)}\rangle = 0,$$

$$\langle\psi_x^{(N)}|\bar{\sigma}_3^2|\psi_x^{(N)}\rangle = \frac{1}{N^2}\left\langle\psi_x^{(N)}\left|\sum_{ij}\sigma_3^{(i)}\sigma_3^{(j)}\right|\psi_x^{(N)}\right\rangle =$$

$$= \frac{1}{N^2}\sum_{ij}\delta_{ij} = \frac{N}{N^2} = \frac{1}{N}. \tag{3.186}$$

Формально переходя к пределу при $N \to \infty$, мы приходим к выводу, что $\bar{\sigma}_3$ имеет исчезающую дисперсию вокруг его среднего значения $\langle\bar{\sigma}_3\rangle = 0$, следовательно, по крайней мере в этом смысле, $|\psi_x^{(\infty)}\rangle$ является «собственным состоянием» оператора $\bar{\sigma}_3$ с нулевым собственным значением.

Коулмен и Лесниевски отметили, что в доказательстве Хартла можно пойти дальше и показать, что результат измерения $|\uparrow_z\rangle$ не только появляется с правильной частотой, но что результаты $|\uparrow_z\rangle$ *случайным образом распределены*. Чтобы придать смысл этому утверждению, мы должны сформулировать определение случайности. Мы говорим, что бесконечная последовательность битов случайна, если она *несжимаема*; нет проще способа генерировать первые N битов, чем просто выписать их. Мы формализуем эту идею, рассматривая длину кратчайшей компьютерной программы (для некоторого компьютера), генерирующей первые N битов последовательности. Тогда для случайного ряда

$$\text{длина кратчайшей программы} > N - \text{const}, \tag{3.187}$$

где константа может зависеть от конкретного используемого компьютера или от конкретной последовательности, но не от N.

Коулмен и Лесниевски рассмотрели ортогональный проекционный оператор $\mathbf{E}_{\text{random}}$, действие которого на $|\psi\rangle$ — собственное состояние оператора $\sigma_3^{(i)}$ удовлетворяет условиям

$$\mathbf{E}_{\text{random}}|\psi\rangle = |\psi\rangle, \tag{3.188}$$

если последовательность собственных значений $\sigma_3^{(i)}$ случайна, и

$$\mathbf{E}_{\text{random}}|\psi\rangle = 0, \tag{3.189}$$

если эта последовательность не случайна. Одного этого свойства недостаточно для определения того, как $\mathbf{E}_{\text{random}}$ действует на всем пространстве $(\mathcal{H}_2)^\infty$, но с учетом дополнительного, имеющего технический характер, предположения они нашли, что $\mathbf{E}_{\text{random}}$ существует, единственный

и обладает свойством

$$\mathbf{E}_{\text{random}}|\psi_x^{(\infty)}\rangle = |\psi_x^{(\infty)}\rangle. \tag{3.190}$$

Таким образом, мы «также можем сказать», что $|\psi_x^{(\infty)}\rangle$ является случайным, что касается измерений σ_3, — процедура, отличающая случайные состояния от неслучайных, которая прекрасно работает для последовательности собственных значений оператора σ_3, столь же надежно будет идентифицировать $|\psi_x^{(\infty)}\rangle$ как случайный вектор.

Эти аргументы интересны, но они не приносят мне полного удовлетворения. Больше всего беспокоит необходимость рассматривать бесконечные последовательности (общая черта любой частотной интерпретации вероятности). При любом конечном N мы не можем применить ослабленный постулат измерения Хартла, и даже в пределе $N \to \infty$ применение этого постулата содержит некоторые тонкости. Желательно было бы иметь *усиленный* слабый постулат измерения, применимый к конечному N, но я не знаю, как сформулировать такой постулат или как его объяснить.

В заключение: Физика должна описывать объективный физический мир, и лучшим из известных нам представлений физической реальности является квантово-механическая волновая функция. Физика должна стремиться объяснять все наблюдаемые явления как можно более экономично, в частности, не апеллируя к постулату, что процесс измерения управляется иными динамическими принципами, нежели другие процессы. К счастью, все, что мы знаем о физике, совместимо с гипотезой о том, что все физические процессы (в том числе измерения) могут быть точно смоделированы унитарной эволюцией волновой функции (или матрицы плотности). Если микроскопическая квантовая система взаимодействует с макроскопическим прибором, то «для всех практических целей» декогерентизация вызывает «коллапс» волновой функции.

Если мы избегаем рассматривать измерение как некий таинственный процесс и принимаем волновую функцию в качестве описания физической реальности, то это ведет нас к Эверетту или интерпретации квантовой теории с позиции «множественности миров». Согласно этой точке зрения все возможные исходы любого «измерения» рассматриваются как «реальные» — но я воспринимаю только один результат, поскольку состояние моего мозга (как части квантовой системы) сильно скоррелировано с ним.

Несмотря на то, что эволюция волновой функции в интерпретации Эверетта является детерминистской, у меня нет возможности однозначно предсказать результат выполняемого в будущем эксперимента — я не знаю, в какой ветви волновой функции окажусь после него, следовательно, я не в состоянии предсказать будущее состояние моего разума. Таким образом,

хотя «глобальная» картина в известной степени детерминистская, из моего собственного локального вида изнутри системы я ощущаю квантово-механическую случайность.

Мой личный взгляд состоит в том, что эверетовская интерпретация квантовой теории дает удовлетворительное объяснение измерения и природы случайности, но все еще не дает полного объяснения квантово-механических правил вычисления вероятностей. Для полного объяснения следует выйти за рамки частотной интерпретации вероятности — в идеале хотелось бы поставить бейесовский взгляд на вероятность на надежное объективное основание.

3.7. Резюме

ПОЗМ. Если мы ограничиваем наше внимание подпространством более широкого гильбертова пространства, то ортогональное измерение (измерение фон Неймана), выполненное в более широком пространстве, вообще говоря, не может быть описано как ортогональное измерение в подпространстве. Это скорее *обобщенное измерение* или *ПОЗМ*, результат которого появляется с вероятностью

$$\mathrm{Prob}(a) = \mathrm{tr}\,(\mathbf{F}_a \rho), \tag{3.191}$$

где ρ — матрица плотности подсистемы, $\mathbf{F}_a$ — положительные эрмитовы операторы, удовлетворяющие условию

$$\sum_a \mathbf{F}_a = \mathbf{1}. \tag{3.192}$$

ПОЗМ в $\mathcal{H}_A$ может быть реализована как унитарное преобразование на тензорном произведении $\mathcal{H}_A \otimes \mathcal{H}_B$ после ортогонального измерения в $\mathcal{H}_B$.

Супероператор. Унитарная в $\mathcal{H}_A \otimes \mathcal{H}_B$ эволюция в общем случае не будет выглядеть унитарной, если мы ограничим свое внимание только пространством $\mathcal{H}_A$. Скорее эволюция в $\mathcal{H}_A$ будет описываться *супероператором* (который обратим только тогда, когда он унитарен). Произвольный супероператор \$ имеет представление операторной суммы (представление Крауса)

$$\$: \rho \to \$(\rho) = \sum_\mu \mathbf{M}_\mu \rho \mathbf{M}_\mu^\dagger, \tag{3.193}$$

где

$$\sum_\mu \mathbf{M}_\mu^\dagger \mathbf{M}_\mu = \mathbf{1}. \tag{3.194}$$

Фактически любое разумное (линейное и вполне положительное) отображение матриц плотности в матрицы плотности имеет унитарное представление и представление операторной суммы.

Декогерентизация. Декогерентизация – разрушение квантовой информации вследствие взаимодействия системы с ее окружением – может быть описана супероператором. Если окружение часто «рассеивает» систему и его состояние не контролируется, тогда в некотором выделенном базисе (обычно, в соответствии с природой связи системы с окружением, выбирается пространственно-локализованный базис) недиагональные элементы матрицы плотности системы быстро затухают. Временной масштаб декогерентизации определяется частотой рассеяния, которая может быть гораздо больше темпа затухания состояния.

Основное уравнение. Когда соответствующий динамический временной масштаб открытой квантовой системы велик по сравнению со временем, в течение которого окружение «забывает» квантовую информацию, эволюция системы эффективно локальна во времени (марковское приближение). Подобно тому как общая унитарная эволюция генерируется гамильтонианом, общая марковская эволюция генерируется *супероператором Линдблада* $\mathcal{L}$, как это описывается *основным уравнением*

$$\dot{\rho} = \mathcal{L}\rho = -i[\mathbf{H}, \rho] + \sum_{\mu} \left(\mathbf{L}_{\mu} \rho \mathbf{L}_{\mu}^{\dagger} - \frac{1}{2} \mathbf{L}_{\mu}^{\dagger} \mathbf{L}_{\mu} \rho - \frac{1}{2} \rho \mathbf{L}_{\mu}^{\dagger} \mathbf{L}_{\mu} \right). \qquad (3.195)$$

Здесь каждый *оператор Линдблада* (или *оператор квантового скачка*) представляет «квантовый скачок», который в принципе можно регистрировать, если достаточно тщательно контролировать окружение. Решая основное уравнение, мы можем вычислить темп декогерентизации открытой системы.

3.8. Упражнения

3.1. Реализация ПОЗМ. Рассмотрим ПОЗМ, определенную четырьмя положительными операторами

$$\begin{aligned} P_1 &= \tfrac{1}{2} |\uparrow_z\rangle\langle\uparrow_z|, & P_2 &= \tfrac{1}{2} |\downarrow_z\rangle\langle\downarrow_z|, \\ P_3 &= \tfrac{1}{2} |\uparrow_x\rangle\langle\uparrow_x|, & P_4 &= \tfrac{1}{2} |\downarrow_x\rangle\langle\downarrow_x|. \end{aligned} \qquad (3.196)$$

Покажите, каким образом эту ПОЗМ можно реализовать как ортогональное измерение в двухкубитовом гильбертовом пространстве, если введен один вспомогательный (ancilla) спин.

3.2. Обратимость супероператоров. Цель этого упражнения — показать, что супероператор обратим только тогда, когда он унитарен. Напомним, что любой супероператор может быыть представлен в виде операторной суммы; он действует на чистое состояние как

$$\mathcal{M}(|\psi\rangle\langle\psi|) = \sum_{\mu} \mathbf{M}_{\mu}|\psi\rangle\langle\psi|\mathbf{M}_{\mu}^{\dagger}, \tag{3.197}$$

где $\sum_{\mu} \mathbf{M}_{\mu}^{\dagger}\mathbf{M}_{\mu} = \mathbf{1}$. Другой супероператор $\mathcal{N}$ называется обратным по отношению к $\mathcal{M}$, если $\mathcal{N} \circ \mathcal{M} = 1$, или

$$\sum_{\mu,a} \mathbf{N}_{a}\mathbf{M}_{\mu}|\psi\rangle\langle\psi|\mathbf{M}_{\mu}^{\dagger}\mathbf{N}_{a}^{\dagger} = |\psi\rangle\langle\psi| \tag{3.198}$$

для любого $|\psi\rangle$. Отсюда следует, что

$$\sum_{\mu,a} \left|\langle\psi|\mathbf{N}_{a}\mathbf{M}_{\mu}|\psi\rangle\right|^{2} = 1. \tag{3.199}$$

a) Используя условия нормировки, которым удовлетворяют $\mathbf{N}_a$ и $\mathbf{M}_\mu$, покажите, что $\mathcal{N} \circ \mathcal{M} = 1$ влечет за собой

$$\mathbf{N}_{a}\mathbf{M}_{\mu} = \lambda_{a\mu}\mathbf{1} \tag{3.200}$$

для *всех* a и μ, другими словами, каждое произведение $\mathbf{N}_a\mathbf{M}_\mu$ пропорционально тождественному (единичному) оператору.

b) Используя результат (**a**), покажите, что для всех μ и ν $\mathbf{M}_{\nu}^{\dagger}\mathbf{M}_{\mu}$ пропорционально тождественному оператору.

c) Покажите, что из (**b**) следует унитарность $\mathcal{M}$.

3.3. Как много супероператоров? Сколько вещественных параметров необходимо для параметризации супероператора общего вида

$$\$: \rho \rightarrow \rho', \tag{3.201}$$

если ρ — оператор плотности в N-мерном гильбертовом пространстве? [**Указание**: Сколько вещественных чисел параметризует эрмитову $N \times N$-матрицу? Как много линейных отображений эрмитовых матриц на эрмитовы матрицы? Как много *сохраняющих след* отображений эрмитовых матриц на эрмитовы матрицы?]

3.4. Насколько быстра декогерентизация? Очень хороший маятник с массой $m = 1\,\text{г}$ и круговой частотой $\omega = 1\,\text{с}^{-1}$ имеет добротность $Q = 10^9$. Маятник приготовлен в состоянии «кот-суперпозиции»

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}\big(|x\rangle + |-x\rangle\big) \tag{3.202}$$

волновых пакетов с минимальной неопределенностью, первоначально покоящихся в положениях $\pm x$, где $x = 1\,\text{см}$. Оценить по порядку величины, как быстро произойдет декогерентизация этого «кот-состояния», если окружение находится

a) при нулевой температуре;

b) при комнатной температуре.

3.5. Затухание фазы. На лекции мы получили представление операторной суммы канала затухания фазы для одного кубита с операторами Крауса

$$\mathbf{M}_0 = \sqrt{1-p}\,\mathbf{1},\ \mathbf{M}_1 = \sqrt{p}\,\frac{1}{2}(\mathbf{1}+\boldsymbol{\sigma}_3),\ \mathbf{M}_2 = \sqrt{p}\,\frac{1}{2}(\mathbf{1}-\boldsymbol{\sigma}_3). \tag{3.203}$$

a) Найдите альтернативное представление, используя только два оператора Крауса $\mathbf{N}_0, \mathbf{N}_1$.

b) Найдите унитарную 3×3-матрицу $U_{\mu a}$ такую, что полученные вами в (a) операторы Крауса (дополненные третьим $\mathbf{N}_2 = 0$) связаны с $\mathbf{M}_{0,1,2}$ соотношением

$$\mathbf{M}_\mu = U_{\mu a}\mathbf{N}_a. \tag{3.204}$$

c) Рассмотрите унитарное представление однокубитового канала

$$\begin{aligned} |0\rangle_A|0\rangle_E &\to \sqrt{1-p}\,|0\rangle_A|0\rangle_E + \sqrt{p}\,|0\rangle_A|\gamma_0\rangle_E, \\ |1\rangle_A|0\rangle_E &\to \sqrt{1-p}\,|1\rangle_A|0\rangle_E + \sqrt{p}\,|1\rangle_A|\gamma_1\rangle_E, \end{aligned} \tag{3.205}$$

где $|\gamma_0\rangle_E$ и $|\gamma_1\rangle_E$ — ортогональные $|0\rangle_E$ нормированные состояния, удовлетворяющие условию

$${}_E\langle\gamma_0|\gamma_1\rangle_E = 1-\varepsilon, \qquad 0<\varepsilon<1. \tag{3.206}$$

Покажите, что это тоже канал затухания фазы, и найдите его представление операторной суммы с двумя операторами Крауса.

d) Допустим, что канал из (c) описывает то, что происходит с кубитом, когда на нем рассеивается один фотон. Выразите темп декогерентизации Γ_{decoh} через темп рассеяния Γ_{scatt}.

3.6. Декогерентизация на сфере Блоха. Параметризуйте матрицу плотности одного кубита следующим образом:

$$\boldsymbol{\rho} = \frac{1}{2}\left(\mathbf{1} + \vec{P}\cdot\vec{\boldsymbol{\sigma}}\right). \tag{3.207}$$

a) Опишите, что происходит с $\vec{P}$ под действием канала затухания фазы.

b) Опишите, что происходит с $\vec{P}$ под действием канала затухания амплитуды, определяемого операторами Крауса

$$\mathbf{M}_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}, \qquad \mathbf{M}_1 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}. \tag{3.208}$$

c) Проделайте то же самое для «двойного канала Паули»:

$$\mathbf{M}_0 = \sqrt{1-p}\,\mathbf{1}, \quad \mathbf{M}_1 = \sqrt{\frac{p}{2}}\,\boldsymbol{\sigma}_1, \quad \mathbf{M}_2 = \sqrt{\frac{p}{2}}\,\boldsymbol{\sigma}_3. \tag{3.209}$$

3.7. Декогерентизация затухающего осциллятора. На лекции мы говорили, что в представлении взаимодействия матрица плотности $\boldsymbol{\rho}_I(t)$ осциллятора, который может излучать кванты в находящийся при нулевой температуре резервуар, подчиняется основному уравнению

$$\dot{\boldsymbol{\rho}}_I = \Gamma\left(\mathbf{a}\boldsymbol{\rho}_I\mathbf{a}^\dagger - \frac{1}{2}\mathbf{a}^\dagger\mathbf{a}\boldsymbol{\rho}_I - \frac{1}{2}\boldsymbol{\rho}_I\mathbf{a}^\dagger\mathbf{a}\right), \tag{3.210}$$

где $\mathbf{a}$ — осцилляторный оператор уничтожения.

a) Рассмотрите величину

$$X(\lambda, t) = \mathrm{tr}\left[\boldsymbol{\rho}_I(t)e^{\lambda\mathbf{a}^\dagger}e^{-\lambda^*\mathbf{a}}\right], \tag{3.211}$$

где λ — комплексное число. Используя основное уравнение, выведите и решите дифференциальное уравнение для $X(\lambda, t)$. Найдите

$$X(\lambda, t) = X(\lambda', 0), \tag{3.212}$$

где λ' является функцией от λ, Γ и t. Что это за функция $\lambda'(\lambda, \Gamma, t)$?

b) Предположим, что при $t = 0$ приготовлено «кот-состояние» осциллятора

$$|\mathrm{cat}\rangle = \frac{1}{\sqrt{2}}\left(|\alpha_1\rangle + |\alpha_2\rangle\right), \tag{3.213}$$

где $|\alpha\rangle$ обозначает когерентное состояние

$$|\alpha\rangle = e^{-|\alpha|^2/2}e^{\alpha\mathbf{a}^\dagger}|0\rangle. \tag{3.214}$$

Используйте результат (**a**), чтобы получить матрицу плотности в более поздний момент времени t. Каков темп затухания недиагональных элементов $\boldsymbol{\rho}$ (в этом когерентном базисе) при $\Gamma t \ll 1$?

Глава 4

Квантовое запутывание

4.1. Несепарабельность ЭПР-пар

4.1.1. Скрытая квантовая информация

Глубокие аспекты, отличающие квантовую информацию от классической, включают в себя свойства, привлечение и использование *квантового запутывания*. Вспомним, что, согласно § 2.4.1, бинарное состояние *запутано*, если его число Шмидта больше единицы. Запутанные состояния интересны тем, что в них проявляются не имеющие классических аналогов корреляции.

В качестве примера напомним определенное в § 3.4.1 *максимально запутанное* состояние двух кубитов (или *ЭПР-пара*[1]):

$$|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}\big(|00\rangle_{AB} + |11\rangle_{AB}\big). \tag{4.1}$$

«Максимально запутанный» означает, что если мы вычислим след по состояниям кубита B, чтобы найти оператор плотности ρ_A кубита A, то получим оператор, пропорциональный единичному:

$$\rho_A = \mathrm{tr}_B\big(|\phi^+\rangle_{AB\,AB}\langle\phi^+|\big) = \frac{1}{2}\mathbf{1}_A \tag{4.2}$$

(и аналогично $\rho_B = \frac{1}{2}\mathbf{1}_B$). Это значит, что результат измерения спина A вдоль *любой* оси будет полностью случайным: с вероятностью $1/2$ мы найдем его ориентированным вверх, и с вероятностью $1/2$ – вниз. Следовательно, если мы выполним любое локальное измерение A или B, то не получим никакой информации о приготовленном состоянии, лишь

[1] ЭПР – Эйнштейн, Подольский, Розен. – *Прим. перев.*

породив вместо этого случайный бит. Эта ситуация резко контрастирует со случаем одного кубита в чистом состоянии. Приготовив, скажем, $|\uparrow_{\hat{n}}\rangle$ или $|\downarrow_{\hat{n}}\rangle$, мы можем хранить в этом состоянии один бит и достоверно извлекать его, выполняя измерение вдоль оси $\hat{n}$. В случае двух кубитов нам следовало бы уметь хранить два бита, но в состоянии $|\phi^+\rangle_{AB}$ эта информация *скрыта*; по крайней мере, мы не можем извлечь ее, измеряя A или B.

Фактически $|\phi^+\rangle_{AB}$ является одним из представителей введенного в §3.4.1 базиса четырех взаимно ортогональных состояний двух кубитов, каждый из которых также максимально запутан:

$$\begin{aligned} |\phi^{\pm}\rangle_{AB} &= \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} \pm |11\rangle_{AB}\right), \\ |\psi^{\pm}\rangle_{AB} &= \frac{1}{\sqrt{2}}\left(|01\rangle_{AB} \pm |10\rangle_{AB}\right). \end{aligned} \tag{4.3}$$

Представим, что Алиса и Боб играют с Чарли. Чарли готовит одно из этих четырех состояний, кодируя таким образом два бита в состоянии двухкубитовой системы. Один из них представляет собой бит *четности* ($|\phi\rangle$ или $|\psi\rangle$): параллельны или антипараллельны состояния двух спинов? Другой — бит *фазы* (+ или −): какой четности выбрана суперпозиция двух состояний? Затем Чарли посылает кубит A Алисе, а кубит B Бобу. Чтобы выиграть, Алиса (или Боб) должна определить, какое из четырех состояний приготовил Чарли.

Конечно, если бы Алиса и Боб соединили свои кубиты вместе, то они смогли бы идентифицировать состояние, выполняя ортогональное измерение, проецирующее на базис $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$. Но представим, что они находятся в разных городах и вообще не могут связаться друг с другом. Действуя локально, ни Алиса, ни Боб не могут извлечь никакой информации о состоянии.

Все, что они могут делать локально, это *манипулировать* этой информацией. Алиса может применить преобразование σ_3 к своему кубиту A, изменяя относительную фазу $|0\rangle_A$ и $|1\rangle_A$. Это действие обращает бит фазы, хранящийся в запутанном состоянии:

$$\begin{aligned} |\phi^{\pm}\rangle &\leftrightarrow |\phi^{\mp}\rangle, \\ |\psi^{+}\rangle &\leftrightarrow |\psi^{-}\rangle. \end{aligned} \tag{4.4}$$

С другой стороны, она может применить преобразование σ_1, которое опрокидывает ее спин ($|0\rangle_A \leftrightarrow |1\rangle_A$) и таким образом инвертирует бит четности

запутанного состояния:

$$\begin{aligned} |\phi^+\rangle &\leftrightarrow |\psi^+\rangle, \\ |\phi^-\rangle &\leftrightarrow -|\psi^-\rangle. \end{aligned} \tag{4.5}$$

Аналогично и Боб может манипулировать запутанным состоянием. Фактически, как мы обсуждали в §2.4, или Алиса, или Боб могут выполнить локальное унитарное преобразование, заменяющее одно максимально запутанное состояние на любое другое максимально запутанное состояние[1]. Поскольку их локальные унитарные преобразования *не могут* изменить $\rho_A = \rho_B = \frac{1}{2}\mathbf{1}$, информация, которой они манипулируют, ни одним из них не может быть прочитана.

Предположим теперь, что Алиса и Боб могут обмениваться (классическими) сообщениями о результатах своих измерений; тогда вместе они могут узнать о том, как скоррелированы их измерения. Запутанные состояния базиса удобно характеризовать одновременными собственными значениями двух коммутирующих наблюдаемых

$$\begin{aligned} &\sigma_1^A \otimes \sigma_1^B, \\ &\sigma_3^A \otimes \sigma_3^B; \end{aligned} \tag{4.6}$$

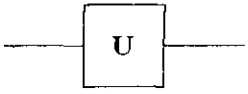
собственное значение оператора $\sigma_3^A \otimes \sigma_3^B$ является битом четности, а собственное значение $\sigma_1^A \otimes \sigma_1^B$ — битом фазы. Так как эти операторы коммутируют, они в принципе могут быть измерены одновременно. Но это невозможно, пока Алиса и Боб выполняют локальные измерения. Они могли бы оба решить измерить свои спины вдоль оси $\hat{z}$, приготовив одновременно собственные состояния операторов σ_3^A и σ_3^B. Поскольку σ_3^A и σ_3^B коммутируют с оператором четности $\sigma_3^A \otimes \sigma_3^B$, их ортогональные измерения не возмущают бит четности, а их результаты можно скомбинировать так, чтобы получить значение этого бита. Однако операторы σ_3^A и σ_3^B *не* коммутируют с оператором $\sigma_1^A \otimes \sigma_1^B$, поэтому выполненное таким способом измерение бита четности возмущает бит фазы. С другой стороны, Алиса и Боб могли бы решить измерить свои спины вдоль оси $\hat{x}$; тогда они могли бы узнать бит фазы ценой возмущения бита четности. Но они не могут одновременно выполнить оба этих измерения. Чтобы можно было надеяться получить бит четности без возмущения бита фазы, Алисе и Бобу нужно получить информацию о произведении $\sigma_3^A \otimes \sigma_3^B$, не изме-

[1] Но, конечно, этого недостаточно для того, чтобы выполнить произвольное унитарное преобразование в четырехмерном пространстве $\mathcal{H}_A \otimes \mathcal{H}_B$, содержащее также и не максимально запутанные состояния. Максимально запутанные состояния *не* образуют подпространства — их суперпозиция обычно *не* является максимально запутанной.

ряя отдельно ни σ_3^A, ни σ_3^B, что не может быть выполнено локальным образом.

Пусть теперь Алиса и Боб соберутся вместе, так чтобы они могли оперировать своими кубитами сообща. Как они могут узнать бит четности и бит фазы их пары? Применив подходящее унитарное преобразование, они могут повернуть запутанный базис $\{|\phi^+\rangle,\ |\psi^-\rangle\}$ таким образом, что он перейдет в незапутанный базис $\{|00\rangle,\ |01\rangle,\ |10\rangle,\ |11\rangle\}$. Тогда Алиса и Боб могут измерить кубиты A и B, чтобы получить искомые ими биты. Как строится это преобразование?

Воспользуемся удобным моментом, чтобы ввести обозначение, которое будет широко использоваться далее в этом курсе, обозначение квантовой схемы. Кубиты изображаются горизонтальными линиями, а однокубитовое унитарное преобразование $\mathbf{U}$ —



В частности, ниже нам очень пригодится однокубитовое унитарное *преобразование Адамара*

$$\mathbf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_3), \tag{4.7}$$

которое обладает свойствами

$$\mathbf{H}^2 = \mathbf{1}, \tag{4.8}$$

и

$$\begin{aligned} \mathbf{H}\boldsymbol{\sigma}_1\mathbf{H} &= \boldsymbol{\sigma}_3, \\ \mathbf{H}\boldsymbol{\sigma}_3\mathbf{H} &= \boldsymbol{\sigma}_1. \end{aligned} \tag{4.9}$$

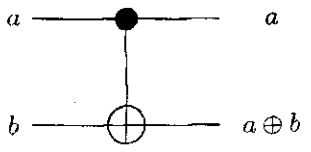
[Мы можем рассматривать $\mathbf{H}$ (с точностью до общей фазы) как поворот на угол $\theta = \pi$ вокруг оси $\hat{n} = \frac{1}{\sqrt{2}}(\hat{n}_1 + \hat{n}_3)$, который переводит друг в друга оси $\hat{x}$ и $\hat{z}$; мы имеем

$$\mathbf{U}(\hat{n}, \theta) = \mathbf{1}\cos\frac{\theta}{2} + i\hat{n}\cdot\vec{\boldsymbol{\sigma}}\sin\frac{\theta}{2} = i\frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_3) = i\mathbf{H}.] \tag{4.10}$$

Также полезно двухкубитовое преобразование, известное как обратимое XOR или контролируемое НЕ (CNOT) преобразование; оно действует как

$$\mathbf{CNOT}\ :\ |a, b\rangle \rightarrow |a, a \oplus b\rangle \tag{4.11}$$

на базисных состояниях $a, b = 0, 1$, где $a \oplus b$ обозначает сумму по модулю два. **CNOT** изображается диаграммой

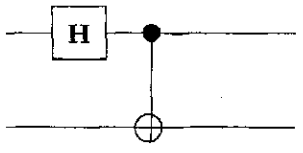


Таким образом, это преобразование инвертирует второй бит, если первый имеет значение 1, и действует тривиально, если первый бит имеет значение 0; оно обладает свойством

$$(\mathbf{CNOT})^2 = \mathbf{1} \otimes \mathbf{1}. \tag{4.12}$$

Мы называем a *контролирующим* битом (или *источником*) операции **CNOT**, а b - *контролируемым* битом (или *целью*).

Комбинируя эти «примитивные» преобразования, или квантовые *вентили*, мы можем построить другие унитарные преобразования. Например, «схема» (читается слева направо)



представляет произведение примененной к первому кубиту операции **H** и следующей за ней **CNOT** с первым кубитом в качестве контролирующего и вторым — в качестве контролируемого. Непосредственно видно, что эта схема преобразует стандартный базис в запутанный:

$$\begin{aligned}
|00\rangle &\to \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)|0\rangle \to |\phi^+\rangle,\\
|01\rangle &\to \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big)|1\rangle \to |\psi^+\rangle,\\
|10\rangle &\to \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big)|0\rangle \to |\phi^-\rangle,\\
|11\rangle &\to \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big)|1\rangle \to |\psi^-\rangle,
\end{aligned} \tag{4.13}$$

так что первый бит становится битом фазы в запутанном базисе, а второй — битом четности.

Аналогично мы можем обратить преобразование, проходя схему в обратном направлении (поскольку и **CNOT**, и **H** идемпотентны); если мы применим обращенную схему к запутанному состоянию, а затем измерим оба бита, то мы узнаем значения бита фазы и бита четности.

Конечно, **H** действует только на один из кубитов; «нелокальной» частью нашей схемы является операция контролируемого НЕ (**CNOT**) – это операция, устанавливающая или устраняющая запутывание. Если бы только мы могли выполнить «межзвездную **CNOT**», то были бы в состоянии запутывать пространственно-разделенные пары или извлекать закодированную в них информацию. Однако мы не можем этого сделать. Чтобы выполнить эту работу, вентиль **CNOT** должен действовать на цель, не открывая значения источника. Локальных операций и классической связи для этого недостаточно.

4.1.2. Эйнштейновская локальность и скрытые переменные

Эйнштейна смущало квантовое запутывание. В конце концов он совместно с Подольским и Розеном (ЭПР) выразил это беспокойство в том, что они рассматривали как парадокс[1]. Согласно более поздней интерпретации Бома, описанная ими ситуация в действительности та же самая, что и обсуждавшаяся в § 2.5.3. Данное максимально запутанное состояние двух кубитов поделено между Алисой и Бобом, Алиса может выбрать одно из нескольких возможных измерений, чтобы выполнить его на своем спине, реализуя тем самым различные возможные интерпретации ансамблей матрицы плотности Боба; например, она может приготовить собственные состояния: или σ_1, или σ_3.

Мы видели, что Алиса и Боб не могут использовать это явление для сверхсветовой связи. Эйнштейн знал это, но оставался неудовлетворенным. Он считал, что теория, дающая *полное* описание физической реальности, должна удовлетворять более строгому критерию, который можно назвать *эйнштейновской локальностью* (иногда известный как *локальный реализм*).

> Предположим, что A и B – разделенные пространственно-подобным интервалом системы. Тогда в *полном* описании физической реальности действие, совершенное над системой A, не должно изменять описание системы B.

[1] A. Einstein, B. Podolsky, N. Rosen, *Can Quantum-mechanical Description of Physical Reality Be Considered Complete?*, Phys. Rev., 47, 777–780 (1935); современная интерпретация мысленного эксперимента ЭПР, использующая максимально запутанное состояние спинов, предложена Д. Бомом: D. Bohm, *Quantum Theory*, Prentice-Hall, Englewood Cliffs< New Jersey (1951); перевод: Д. Бом, *Квантовая теория*, М., Наука (1965). – *Прим. ред.*

Но если A и B запутаны, измерение A выполнено и *конкретный* полученный результат известен, то матрица плотности B обязательно *изменится*. Следовательно, согласно критерию Эйнштейна описание квантовой системы с помощью волновой функции или оператора плотности не может считаться полным.

Эйнштейн пытался представить более полное описание, которое устранило бы индетерминизм квантовой механики. Теории такого рода называются *теориями локальных скрытых переменных*. В теории скрытых переменных измерение в действительности является детерминистским, но выглядит вероятностным, поскольку некоторые степени свободы точно неизвестны. Например, возможно, что для описания приготовленного спинового состояния, которое в квантовой теории рассматривается как чистое состояние $|\uparrow_{\hat{z}}\rangle$, в действительности существует более глубокая теория, в которой оно параметризуется как $(\hat{z}, \lambda)$, где λ $(0 \leqslant \lambda \leqslant 1)$ — скрытая переменная. Допустим, что при современном уровне экспериментальной техники мы не контролируем λ, следовательно, когда мы готовим спиновое состояние, λ может принять любое значение — распределение вероятностей, управляющее ее значениями, является однородным на единичном интервале.

Теперь предположим, что при измерении спина вдоль оси $\hat{n}$, повернутой на угол θ относительно оси $\hat{z}$, будет получен результат

$$\begin{aligned} |\uparrow_{\hat{z}}\rangle \quad &\text{при} \quad 0 \leqslant \lambda \leqslant \cos^2\frac{\theta}{2}, \\ |\downarrow_{\hat{z}}\rangle \quad &\text{при} \quad \cos^2\frac{\theta}{2} < \lambda \leqslant 1. \end{aligned} \tag{4.14}$$

Если мы знаем λ, то результат является детерминистским, но если λ полностью неизвестна, тогда управляющее измерением распределение вероятностей будет согласоваться с предсказаниями квантовой теории. В теории скрытых переменных случайность результата измерения не является ее внутренним свойством; скорее она является следствием невежества — наше описание системы не является максимально возможным полным описанием.

Теперь как насчет запутанных состояний? Когда мы говорим, что теория скрытых переменных является *локальной*, мы подразумеваем, что она удовлетворяет эйнштейновскому требованию локальности. Измерение A не изменяет значения переменных, определяющих измерения B. Когда Алиса измеряет свою половину запутанного состояния, которое она делит с Бобом, она получает информацию о значениях скрытых переменных, увеличивая возможность предсказать, что получит Боб после измерения своей половины. Это кажется похожим на то, что имел в виду Эйнштейн, говоря о более полном описании.

4.2. Неравенство Белла

4.2.1. Три квантовые монеты

Является ли теория скрытых переменных просто переформулировкой квантовой теории, или она представляет собой допускающую проверку гипотезу? Плодотворная идея Джона Белла состояла в том, чтобы проверить эйнштейновскую локальность, рассматривая количественные свойства корреляций между результатами, полученными двумя экспериментаторами, Алисой и Бобом, измерявшими разные части системы, находящейся в запутанном состоянии. Рассмотрим пример корреляций, которые Алиса и Боб хотели бы объяснить.

Изучаемая Алисой и Бобом система могла бы быть описана следующим образом: Алиса в Пасадене имеет в своем распоряжении три выложенные на стол монеты, помеченные как 1, 2, 3. Каждая монета выпадает «орлом» (О) или «решкой» (Р), но они закрыты непрозрачными крышками, так что Алиса не может сказать, что на них выпало. Она может открыть любую одну из трех монет и таким образом узнать ее значение (О или Р). Но как только одна монета оказывается открытой, две другие закрытые монеты мгновенно исчезают облачком дыма и Алиса уже не имеет возможности открыть их. В ее распоряжении множество копий трехмонетного набора и в конце концов она понимает, что, независимо от того, какая монета открывается, вероятности обнаружить О или Р одинаковы. Боб в Чикаго имеет аналогичный набор монет, также помеченных 1, 2 и 3. Он тоже обнаруживает, что каждая из его монет, когда открывается, с одинаковой вероятностью показывает или О, или Р.

Фактически Алиса и Боб имеют множество идентичных копий поделенных между ними наборов монет; они проводят обширную серию экспериментов, чтобы исследовать, как их наборы монет коррелируют между собой. Они быстро делают замечательное открытие: всякий раз, когда Алиса и Боб открывают монеты с одинаковыми метками (1,2 или 3), они *всегда* находят их в одинаковом состоянии - обе показывают или О, или Р. Они проводят миллион испытаний, чтобы быть точно уверенными, что это происходит всегда! Их наборы монет идеально скоррелированы.

Алиса и Боб догадываются, что обнаружили нечто важное, и часто разговаривают по телефону, обсуждая внезапные идеи о смысле своих результатов. Однажды Алиса находилась в особенно задумчивом настроении.

Алиса: Ты знаешь, Боб, мне иногда трудно решить, какую из трех монет открыть. Я знаю, что если я открою, скажем, монету 1, то монеты 2 и 3 исчезнут, и у меня не будет возможности узнать, что выпало на

них. Хотя бы раз мне удалось открыть две из трех монет и узнать, что на них выпало: «орел» (О) или «решка» (Р). Я пыталась, но это действительно невозможно — нет способа увидеть одну монету и не дать исчезнуть другим!

Боб: [*Долгая пауза*]. Эй! ... подожди минутку, Алиса, у меня появилась идея.... Смотри, я думаю, что у тебя *есть* способ в конце концов узнать, что выпало на двух твоих монетах! Допустим, ты хотела бы открыть монеты 1 и 2. Ну, тогда я открою свою монету 2 здесь, в Чикаго, и сообщу тебе, что я обнаружил, пусть, к примеру, на ней выпало О. Тогда мы знаем, что если ты тоже откроешь монету 2, то наверняка найдешь О. В этом нет никаких сомнений, так как мы проверяли это миллион раз. Верно?

Алиса: Верно ....

Боб: Но теперь тебе ни к чему открывать твою монету 2; ты же точно знаешь, что на ней обнаружишь. Вместо этого ты можешь открыть монету 1. Тогда ты узнаешь, что выпало на обеих монетах.

Алиса: Гмм ..., да, может быть. Да, я имею в виду, что раньше, когда мы открывали одни и те же монеты, это всегда срабатывало, но теперь ты открыл свою монету 2 и твои монеты 1 и 3 исчезли, а я открыла свою монету 1, и мои монеты 2 и 3 исчезли. Нет возможности даже попытаться еще раз проверить, что случилось бы, если бы мы оба открыли монету 2.

Боб: Не надо проверять это еще раз, Алиса; мы уже миллион раз проверяли это. Смотри, твои монеты в Пасадене, а мои — в Чикаго. Очевидно, что просто нет способа, которым мое решение открыть мою монету 2 может *повлиять* на то, что ты обнаружишь, когда откроешь свою монету 2. То есть это невозможно. Просто когда я открываю мою монету 2, мы получаем информацию, необходимую нам, чтобы с уверенностью предсказать, что произойдет, когда ты откроешь свою монету 2. Так как мы уже уверены в этом, зачем заботиться о проверке!

Алиса: Хорошо, Боб, я понимаю, что ты имеешь в виду. Почему мы не можем выполнить эксперимент, чтобы увидеть, что действительно происходит, когда ты и я открываем разные монеты?

Боб: Я не знаю, Алиса. Было бы невероятно получить хоть какое-то финансирование такого «глупого» эксперимента. Я имею в виду, интересует

ли кого-нибудь на самом деле, что происходит, когда я открываю монету 2, а ты — монету 1?

Алиса: Я не уверена. Но я слышала о теоретике по имени Белл. Говорят, что у него интересные идеи относительно монет. Возможно, у него есть теория, которая делает предсказание относительно того, что мы обнаружим. Может быть, нам стоит поговорить с ним?

Боб: Хорошая мысль! И даже неважно, имеет его теория смысл или нет. Мы можем тем не менее предложить эксперимент, чтобы проверить его предсказание, и тогда нас, возможно, спонсируют.

Итак, Алиса и Боб отправляются в ЦЕРН[1], чтобы побеседовать с Беллом. Они рассказывают ему об эксперименте, который они предлагают выполнить. Белл внимательно слушает их, но некоторое время с отрешенным видом хранит молчание. Алису и Боба не очень это беспокоит, так как они не много понимают из того, что говорят теоретики. Но наконец Белл говорит:

Белл: Я думаю, что у меня есть идея .... Когда Боб открывает свою монету в Чикаго, он не может оказать никакого *влияния* на монету Алисы в Пасадене. Вместо этого то, что обнаруживает Боб, открывая свою монету, дает некоторую *информацию* о том, что случится, когда Алиса откроет свою монету.

Боб: Ну, то есть что я и говорил ....

Белл: Правильно. Звучит разумно. Итак, допустим, что Боб в этом прав. Теперь Боб может открыть любую одну из его монет и узнать наверняка, что найдет Алиса, когда она откроет соответствующую монету. Он никак не *затронул* ее монеты; он просто получил информацию о ней. Нам придется сделать вывод, что должны существовать некоторые *скрытые переменные*, которые определяют состояние монет Алисы. И если эти переменные полностью известны, тогда состояние каждой из монет Алисы может быть однозначно предсказано.

Боб: [*Раздраженный всей этой абстрактной чепухой*]. Да, ну и что?

Белл: Когда ваши коррелированные наборы монет были приготовлены, значения скрытых переменных не были полностью определены, вот

[1]CERN — Европейский центр ядерных исследований. — *Прим. перев.*

почему на любой одной монете с равной вероятностью может выпасть О, а может и Р. Однако должно существовать некоторое распределение вероятностей $P(x, y, z)$ (с $x, y, z \in \{\text{O,P}\}$), которое характеризует приготовление и управляет тремя монетами Алисы. Эти вероятности должны быть неотрицательны и в сумме равны единице:

$$\sum_{x,y,z\in\{\text{O,P}\}} P(x, y, z) = 1. \tag{4.15}$$

Алиса не может открыть все три ее монеты, следовательно, она не может непосредственно измерить $P(x, y, z)$. Однако с помощью Боба она в действительности может открыть любые две монеты из своего набора. Обозначим как $P_{\text{same}}(i, j)$ вероятность того, что монеты i и j $(i, j = 1, 2, 3)$ показывают одно и то же: или обе О, или обе Р. Тогда мы видим, что

$$\begin{aligned} P_{\text{same}}(1,2) &= P(\text{OOO}) + P(\text{OOP}) + P(\text{PPO}) + P(\text{PPP}), \\ P_{\text{same}}(2,3) &= P(\text{OOO}) + P(\text{POO}) + P(\text{OPP}) + P(\text{PPP}), \\ P_{\text{same}}(1,3) &= P(\text{OOO}) + P(\text{OPO}) + P(\text{POP}) + P(\text{PPP}). \end{aligned} \tag{4.16}$$

Из уравнения (4.15) непосредственно следует, что

$$\begin{aligned} P_{\text{same}}(1,2) + P_{\text{same}}(2,3) + P_{\text{same}}(1,3) = \\ = 1 + 2P(\text{OOO}) + 2P(\text{PPP}) \geqslant 1. \end{aligned} \tag{4.17}$$

Это и есть мое предсказание: P_{same} должны подчиняться неравенству[1]

$$P_{\text{same}}(1,2) + P_{\text{same}}(2,3) + P_{\text{same}}(1,3) \geqslant 1. \tag{4.18}$$

Вы можете проверить мой вывод в вашем эксперименте, в котором «открываются» две монеты сразу.

Боб: Ну, я допускаю, что математика выглядит правильной. Но на самом деле я не понимаю этого. Почему это работает?

Алиса: Мне кажется, что я поняла ... Белл говорит, что если на столе лежат три монеты и на каждой из них либо О, либо Р, тогда по крайней мере на двух из трех выпадает *одно и то же*: или на обеих О, или на обеих Р. Не так ли, Белл?

[1] Неравенства таково типа получены в работе J.S. Bell, *On the Einstein–Podolsky–Rosen Paradox*, Physics, **1**, 195–200 (1964); см. также J.S. Bell, *On the Problem of Hidden Variables in Quantum Mechanics*, Rev. Mod. Phys., **38**, 447–452 (1966). – *Прим. ред.*

Белл с изумлением смотрит на Алису. Его глаза блестят, на некоторое время он теряет дар речи. Наконец он говорит:

Белл: Да.

Итак, Алиса и Боб были счастливы узнать, что Белл, как редкий зверь, — теоретик, от которого есть толк. Благодаря Беллу, их предложение получило одобрение и они выполняют эксперимент, получая обескураживающий результат. После множества тщательных проверок они с очень высокой статистической надежностью делают вывод, что

$$P_{\text{same}}(1,2) \simeq P_{\text{same}}(2,3) \simeq P_{\text{same}}(1,3) \simeq \frac{1}{4} \tag{4.19}$$

и, следовательно,

$$P_{\text{same}}(1,2) + P_{\text{same}}(2,3) + P_{\text{same}}(1,3) \simeq 3 \cdot \frac{1}{4} = \frac{3}{4} < 1. \tag{4.20}$$

Обнаруженные Алисой и Бобом корреляции вопиюще нарушают неравенство Белла!

Алиса и Боб — хорошие экспериментаторы, но они не решаются публиковать столь возмутительный результат до тех пор, пока не смогут найти ему правдоподобное теоретическое истолкование. Наконец, дойдя до полного отчаяния, они идут в библиотеку, чтобы узнать, может ли принести хоть какое-то утешение квантовая механика ....

4.2.2. Квантовое запутывание против эйнштейновской локальности

Там Алиса и Боб читают о квантовом запутывании. В конце концов, они узнают, что их волшебные монеты управляются максимально запутанным состоянием двух кубитов. Алиса и Боб в действительности делят множество копий состояния $|\psi^-\rangle$.[1] Когда Алиса открывает монету, она измеряет свой кубит вдоль одной из трех возможных осей, не перпендикулярных между собой. Поскольку измерения не коммутируют, Алиса может открыть только одну из ее трех монет. Аналогично, когда Боб открывает свою монету, он измеряет свою часть запутанной пары вдоль любой одной из трех осей, следовательно, он тоже имеет возможность открыть только одну из

[1] Судя по тому, что, открывая монеты с одинаковыми номерами, Алиса и Боб *всегда* обнаруживали их в одинаковом состоянии (см. § 4.2.1), их трех-монетные наборы должны управляться максимально запутанным состоянием ЭПР-типа $|\phi^+\rangle_{AB}$ (см. уравнение (4.3)). Однако дальнейшие вычисления в этом параграфе выполняются для состояния $|\psi^-\rangle$. *–Прим. ред.*

его трех монет. Но поскольку измерения Алисы коммутируют с измерениями Боба, каждый из них может открыть одну монету и исследовать, как их монеты коррелируют между собой.

Чтобы помочь Алисе и Бобу интерпретировать их эксперимент, посмотрим, что говорит квантовая механика об этих корреляциях. Состояние $|\psi^-\rangle$ обладает полезным свойством: оно остается неизменным, если Алиса и Боб применяют одно и то же унитарное преобразование (2.27)

$$\mathbf{U} \otimes \mathbf{U}|\psi^-\rangle = |\psi^-\rangle. \tag{4.21}$$

В случае бесконечно малого унитарного преобразования оно превращается в свойство

$$\left(\vec{\sigma}^A + \vec{\sigma}^B\right)|\psi^-\rangle = 0 \tag{4.22}$$

(состояние с равным нулю полным моментом импульса, в чем вы можете легко убедиться с помощью явных вычислений). Рассмотрим ожидаемое значение

$$\langle\psi^-|\left(\vec{\sigma}^A \cdot \hat{a}\right)\left(\vec{\sigma}^B \cdot \hat{b}\right)|\psi^-\rangle, \tag{4.23}$$

где $\hat{a}$ и $\hat{b}$ — единичные трехмерные векторы. Действуя на $|\psi^-\rangle$, мы можем заменить $\vec{\sigma}^B$ на $-\vec{\sigma}^A$; следовательно, ожидаемое значение (4.23) может быть представлено как свойство системы Алисы, которая имеет оператор плотности $\boldsymbol{\rho}_A = \frac{1}{2}\mathbf{1}$:

$$\begin{aligned} -\langle\psi^-|\left(\vec{\sigma}^A \cdot \hat{a}\right)\left(\vec{\sigma}^A \cdot \hat{b}\right)|\psi^-\rangle = \\ = -a_i b_j \operatorname{tr}\left(\boldsymbol{\rho}_A \boldsymbol{\sigma}_i^A \boldsymbol{\sigma}_j^A\right) = -a_i b_j \delta_{ij} = -\hat{a}\cdot\hat{b} = -\cos\theta, \end{aligned} \tag{4.24}$$

где θ — угол между осями $\hat{a}$ и $\hat{b}$. Таким образом, мы нашли, что результаты измерения всегда идеально антикоррелированы, когда оба спина измеряются вдоль одной и той же оси $\hat{a}$, но мы также получили и более общий результат, применимый к случаю, когда две оси различны.

Проекционный оператор на состояние спин-вверх (спин-вниз) вдоль оси $\hat{n}$ имеет вид $\mathbf{E}(\hat{n}, \pm) = \frac{1}{2}(\mathbf{1} \pm \hat{n}\cdot\vec{\sigma})$; следовательно, мы получаем

$$\begin{aligned} P(++) &= \langle\psi^-|\mathbf{E}^A(\hat{a},+)\mathbf{E}^B(\hat{b},+)|\psi^-\rangle = \frac{1}{4}(1-\cos\theta), \\ P(--) &= \langle\psi^-|\mathbf{E}^A(\hat{a},-)\mathbf{E}^B(\hat{b},-)|\psi^-\rangle = \frac{1}{4}(1-\cos\theta), \\ P(+-) &= \langle\psi^-|\mathbf{E}^A(\hat{a},+)\mathbf{E}^B(\hat{b},-)|\psi^-\rangle = \frac{1}{4}(1+\cos\theta), \\ P(-+) &= \langle\psi^-|\mathbf{E}^A(\hat{a},-)\mathbf{E}^B(\hat{b},+)|\psi^-\rangle = \frac{1}{4}(1+\cos\theta); \end{aligned} \tag{4.25}$$

здесь $P(++)$ — вероятность того, что Алиса и Боб, оба получают в результате спин-вверх, когда Алиса выполняет измерение вдоль оси $\hat{a}$, а Боб — вдоль оси $\hat{b}$, и так далее. Вероятность того, что их результаты совпадают, равна

$$P_{\text{same}} = P(++) + P(--) = \frac{1}{2}(1 - \cos\theta), \tag{4.26}$$

а вероятность того, что их результаты различны, —

$$P_{\text{opposite}} = P(+-) + P(-+) = \frac{1}{2}(1 + \cos\theta). \tag{4.27}$$

Теперь предположим, что Алиса измеряет свои спины вдоль одной из трех, симметрично ориентированных в плоскости OXZ, осей

$$\hat{a}_1 = (0, 0, 1), \quad \hat{a}_2 = \left(\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right), \quad \hat{a}_3 = \left(-\frac{\sqrt{3}}{2}, 0, -\frac{1}{2}\right), \tag{4.28}$$

так что

$$\hat{a}_1 \cdot \hat{a}_2 = \hat{a}_2 \cdot \hat{a}_3 = \hat{a}_3 \cdot \hat{a}_1 = -\frac{1}{2}. \tag{4.29}$$

Предположим также, что Боб выполняет измерение вдоль одной из трех осей, диаметрально противоположных осям Алисы:

$$\hat{b}_1 = -\hat{a}_1, \qquad \hat{b}_2 = -\hat{a}_2, \qquad \hat{b}_3 = -\hat{a}_3. \tag{4.30}$$

Если Алиса и Боб выбирают противоположные оси, то $\theta = 180^\circ$ и $P_{\text{same}} = 1$. В противном случае $\theta = \pm 60^\circ$, так что $\cos\theta = 1/2$ и $P_{\text{same}} = 1/4$. Это именно то нарушающее предсказание Белла поведение, которое Алиса и Боб обнаружили в своем эксперименте.

Логика Белла выглядит безупречной, но кое-что встало с ног на голову, поэтому мы вынуждены пересмотреть молчаливо подразумеваемые им предположения. Во-первых, Белл предполагает, что существует совместное распределение вероятностей, управляющее возможными исходами всех измерений, которые могут выполнить Алиса и Боб. Это является гипотезой о скрытых переменных. Белл представляет, что если значения скрытых переменных точно известны, то можно с уверенностью предсказать результат любого измерения — результаты измерения описываются вероятностным образом, поскольку значения скрытых переменных извлекаются из некоторого ансамбля возможных значений. Во-вторых, Белл полагает, что решение Боба, какое выполнять измерение в Чикаго, не влияет на скрытые

переменные, управляющие измерением Алисы в Пасадене. Это представляет собой предположение о локальности скрытых переменных. Если мы принимаем эти два предположения, то с неизбежностью приходим к выводу Белла. Мы обнаружили, что корреляции, предсказываемые квантовой теорией, несовместимы с этими предположениями.

Что отсюда следует? Вероятно, урок этой истории в том, что может быть опасно рассуждать о том, что могло бы случиться, но на самом деле не происходит — что иногда называют *контрфактом*. Конечно, в нашей повседневной жизни мы постоянно этим занимаемся и обычно выходим сухими из воды; рассуждения о контрфактах выглядят приемлемыми в классическом мире, но в квантовом мире с ними иногда можно попасть впросак. Мы утверждали, что, поскольку Боб выполнил измерение вдоль оси $\hat{a}_1$, Алиса знала, что произошло бы, если бы она провела измерение вдоль оси $\hat{a}_1$, и сколько бы мы ни проверяли, их результаты всегда идеально скоррелированы. Однако Алиса *не стала* измерять вдоль $\hat{a}_1$; вместо этого она выполнила измерение вдоль $\hat{a}_2$. Мы столкнулись с трудностями, пытаясь приписать вероятности результатам измерений вдоль $\hat{a}_1$, $\hat{a}_2$ и $\hat{a}_3$, несмотря на то, что Алиса может выполнить только одно из них. Предположение о существовании распределения вероятностей, управляющего исходами всех трех измерений, каждое из которых, но только одно, Алиса могла бы выполнить, в квантовой теории ведет к математическим противоречиям, так что нам лучше его не делать. Мы подтвердили принцип *дополнительности* Бора — запрещено одновременно рассматривать исходы двух взаимно исключающих экспериментов.

Тот, кто отвергает принцип дополнительности, может, предпочтет сказать, что (экспериментально подтвержденные) нарушения неравенств Белла продемонстрировали существенную нелокальность, присущую квантовому описанию Природы. *Если* мы действительно настаиваем на законности обсуждения результатов взаимно исключающих экспериментов, *то* неизбежно приходим к выводу, что выбор измерения Боба действительно оказывает тонкое *влияние* на результат измерения Алисы. Таким образом, сторонники этой точки зрения говорят о «квантовой нелокальности».

Исключив локальные скрытые переменные, Белл разбил мечту Эйнштейна о том, что индетерминизм квантовой теории мог бы быть устранен более полным, но все же локальным, описанием Природы. Если мы принимаем локальность как нерушимый принцип, мы вынуждены принять случайность не как следствие неполного знания, а как неизбежное внутреннее свойство квантового измерения.

Некоторые считают, что раскрытые неравенствами Белла специфические корреляции требуют более глубокого объяснения, чем способна дать

квантовая механика. Они рассматривают явление ЭПР как предтечу ожидающей своего открытия новой физики. Но они могут и ошибаться. После ЭПР мы ждали больше 65-ти лет, а новой физики так и нет.

Похоже, человеческий разум плохо подготовлен к тому, чтобы постичь корреляции, демонстрируемые запутанными квантовыми состояниями, и поэтому мы говорим о таинственности квантовой теории. Но какой бы ни была ваша позиция, эксперимент вынуждает вас согласиться с наличием странных корреляций между результатами измерений. Нет большой тайны в том, как эти корреляции были установлены – мы видели, что Алисе и Бобу было необходимо вместе в некоторой точке пространства создать запутывание между их кубитами. Необычность состоит в том, что даже когда A и B пространственно разделены, мы не можем строго рассматривать A и B как два отдельных кубита и использовать классическую информацию для характеристики того, как они коррелируют. Они более, чем просто коррелированы, они представляют собой нечто *единое и неделимое*. Они *запутаны*.

4.3. Еще неравенства Белла

4.3.1. Неравенство КГШХ

Экспериментальные проверки эйнштейновской локальности обычно основываются на другой форме неравенства Белла, применяемого к ситуации, в которой Алиса может измерить одну из двух наблюдаемых $\mathbf{a}$ и $\mathbf{a}'$, в то время как Боб может измерить или $\mathbf{b}$, или $\mathbf{b}'$. Предположим, что наблюдаемые $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$ принимают значения $\{\pm 1\}$ и являются функциями скрытых случайных переменных.

Если $\mathbf{a}, \mathbf{a}' = \pm 1$, то отсюда следует, что или $\mathbf{a} + \mathbf{a}' = 0$, тогда $\mathbf{a} - \mathbf{a}' = \pm 2$, или же $\mathbf{a} - \mathbf{a}' = 0$, тогда $\mathbf{a} + \mathbf{a}' = \pm 2$; следовательно:

$$\mathbf{C} = (\mathbf{a} + \mathbf{a}')\mathbf{b} + (\mathbf{a} - \mathbf{a}')\mathbf{b}' = \pm 2. \tag{4.31}$$

(Здесь тайком введено предположение о локальных скрытых переменных – мы представили, что значения $\{\pm 1\}$ могут быть приписаны одновременно всем четырем наблюдаемым, даже если невозможно одновременное измерение $\mathbf{a}$ и $\mathbf{a}'$ или $\mathbf{b}$ и $\mathbf{b}'$.) Очевидно

$$|\langle \mathbf{C} \rangle| \leqslant \langle |\mathbf{C}| \rangle = 2, \tag{4.32}$$

так что

$$|\langle \mathbf{ab} \rangle + \langle \mathbf{a}'\mathbf{b} \rangle + \langle \mathbf{ab}' \rangle - \langle \mathbf{a}'\mathbf{b}' \rangle| \leqslant 2. \tag{4.33}$$

Этот результат называется *неравенством КГШХ* (Клаузер - Горн - Шимони - Хольт). Оно справедливо для любых случайных переменных $\mathrm{a}, \mathrm{a}', \mathrm{b}, \mathrm{b}'$, принимающих значения $\{\pm 1\}$, которые управляются совместным распределением вероятностей.

Чтобы увидеть, что квантовая механика нарушает неравенство КГШХ, допустим, что a, a' обозначают эрмитовы операторы

$$\mathbf{a} = \vec{\sigma}^{(A)} \cdot \hat{a}, \qquad \mathbf{a}' = \vec{\sigma}^{(A)} \cdot \hat{a}', \tag{4.34}$$

действующие на кубит Алисы, где $\hat{a}$, $\hat{a}'$ — трехмерные единичные векторы. Аналогично b, b' обозначают операторы

$$\mathbf{b} = \vec{\sigma}^{(B)} \cdot \hat{b}, \qquad \mathbf{b}' = \vec{\sigma}^{(B)} \cdot \hat{b}', \tag{4.35}$$

действующие на кубит Боба. Каждая наблюдаемая имеет собственные значения ± 1, то есть результатами их измерения являются значения ± 1.

Напомним, что если Алиса и Боб делят максимально запутанное состояние $|\psi^-\rangle$, то

$$\langle \psi^- | (\vec{\sigma}^A \cdot \hat{a})(\vec{\sigma}^B \cdot \hat{b}) | \psi^- \rangle = -\hat{a} \cdot \hat{b} = -\cos\theta, \tag{4.36}$$

где θ — угол между $\hat{a}$ и $\hat{b}$. Рассмотрим случай, когда $\hat{a}'$, $\hat{b}$, $\hat{a}$, $\hat{b}'$ компланарны и располагаются последовательно через 45°, так что квантовая механика предсказывает:

$$\begin{aligned} \langle \mathbf{ab} \rangle = \langle \mathbf{a}'\mathbf{b} \rangle = \langle \mathbf{ab}' \rangle &= -\cos\frac{\pi}{4} = -\frac{1}{\sqrt{2}}, \\ \langle \mathbf{a}'\mathbf{b}' \rangle &= -\cos\frac{3\pi}{4} = \frac{1}{\sqrt{2}}. \end{aligned} \tag{4.37}$$

Тогда неравенство КГШХ

$$4 \cdot \frac{1}{\sqrt{2}} = 2\sqrt{2} \leqslant 2 \tag{4.38}$$

очевидно нарушается предсказанием квантовой механики.

4.3.2. Максимальное нарушение

Фактически, как мы увидим из следующих аргументов, только что рассмотренный случай представляет максимально возможное квантово-механическое нарушение неравенства КГШХ. Предположим, что $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$

эрмитовы операторы с собственными значениями ± 1, так что

$$\mathbf{a}^2 = \mathbf{a}'^2 = \mathbf{b}^2 = \mathbf{b}'^2 = \mathbf{1}; \tag{4.39}$$

допустим также, что «наблюдаемые Алисы» $\mathbf{a}, \mathbf{a}'$ коммутируют с «наблюдаемыми Боба» $\mathbf{b}, \mathbf{b}'$:

$$[\mathbf{a}, \mathbf{b}] = [\mathbf{a}, \mathbf{b}'] = [\mathbf{a}', \mathbf{b}] = [\mathbf{a}', \mathbf{b}'] = 0. \tag{4.40}$$

Определяя

$$\mathbf{C} = \mathbf{ab} + \mathbf{a}'\mathbf{b} + \mathbf{ab}' - \mathbf{a}'\mathbf{b}' \tag{4.41}$$

и учитывая (4.39), вычислим

$$\mathbf{C}^2 = \begin{array}{cccc} \mathbf{1} & +\mathbf{aa}' & +\mathbf{bb}' & -\mathbf{aa}'\mathbf{bb}' \\ +\mathbf{a}'\mathbf{a} & +\mathbf{1} & +\mathbf{a}'\mathbf{abb}' & -\mathbf{bb}' \\ +\mathbf{b}'\mathbf{b} & +\mathbf{aa}'\mathbf{b}'\mathbf{b} & +\mathbf{1} & -\mathbf{aa}' \\ -\mathbf{a}'\mathbf{ab}'\mathbf{b} & -\mathbf{b}'\mathbf{b} & -\mathbf{a}'\mathbf{a} & +\mathbf{1} \end{array}. \tag{4.42}$$

Все квадратичные члены попарно сокращаются, так что мы остаемся с

$$\begin{aligned} \mathbf{C}^2 &= 4 \cdot \mathbf{1} - \mathbf{aa}'\mathbf{bb}' + \mathbf{a}'\mathbf{abb}' + \mathbf{aa}'\mathbf{b}'\mathbf{b} - \mathbf{a}'\mathbf{ab}'\mathbf{b} \\ &= 4 \cdot \mathbf{1} - [\mathbf{a}, \mathbf{a}'][\mathbf{b}, \mathbf{b}']. \end{aligned} \tag{4.43}$$

Теперь вспомним, что *норма* $\|\mathbf{M}\|$ ограниченного оператора $\mathbf{M}$ определяется как[1]

$$\|\mathbf{M}\| = \sup_{|\psi\rangle} \left(\frac{\|\mathbf{M}|\psi\rangle\|}{\||\psi\rangle\|} \right); \tag{4.44}$$

то есть нормой $\mathbf{M}$ является максимальное собственное значение оператора $\sqrt{\mathbf{M}^\dagger \mathbf{M}}$. Нетрудно проверить, что норма оператора обладает свойствами

$$\begin{aligned} \|\mathbf{MN}\| &\leqslant \|\mathbf{M}\| \cdot \|\mathbf{N}\|, \\ \|\mathbf{M} + \mathbf{N}\| &\leqslant \|\mathbf{M}\| + \|\mathbf{N}\|. \end{aligned} \tag{4.45}$$

[1] В оригинале используется обозначение $\|\cdot\|_{\sup}$ и термин *sup norm*, который можно было бы перевести как *супремум-норма*, или *верхняя норма*. На самом деле (4.44) дает определение обычной нормы ограниченного оператора, которая в русской литературе обозначается как $\|\cdot\|$. См., например, М. Рид, Б. Саймон, *Методы современной математической физики*. Т. 1. *Функциональный анализ*, Мир, М., 1977, стр. 21. — *Прим. ред.*

Эрмитовский оператор с собственными значениями ± 1 имеет единичную норму, так что

$$\|\mathbf{C}^2\| \leqslant 4 + 4\|\mathbf{a}\| \cdot \|\mathbf{a}'\| \cdot \|\mathbf{b}\| \cdot \|\mathbf{b}'\| = 8. \tag{4.46}$$

Поскольку оператор $\mathbf{C}$ эрмитов,

$$\|\mathbf{C}^2\| = \|\mathbf{C}\|^2 \tag{4.47}$$

и, следовательно,

$$\|\mathbf{C}\| \leqslant 2\sqrt{2}, \tag{4.48}$$

что известно как неравенство Цирельсона.

Неравенство КГШХ утверждает, что $|\langle \mathbf{C} \rangle| \leqslant 2$. В квантовой механике абсолютная величина ожидаемого значения эрмитовского оператора $\mathbf{C}$ не может быть больше его максимального собственного значения

$$|\langle \mathbf{C} \rangle| \leqslant \|\mathbf{C}\| \leqslant 2\sqrt{2}. \tag{4.49}$$

Мы видим, что верхняя грань достигается в случае, когда $\hat{a}'$, $\hat{b}$, $\hat{a}$, $\hat{b}'$ компланарны и располагаются последовательно через углы $45°$. Таким образом, найденное нами нарушение неравенства КГШХ является наибольшим допустимым в квантовой теории.

4.3.3. Квантовые стратегии действуют лучше классических

Неравенство КГШХ представляет собой ограничение на величину корреляций между двумя частями бинарной классической системы, а неравенство Цирельсона — ограничение на величину корреляций между двумя частями бинарной квантовой системы. Мы можем углубить наше понимание того, чем квантовые корреляции отличаются от классических, рассматрев игру, в которой квантовые стратегии работают лучше классических.

Алиса и Боб играют с Чарли. Чарли готовит два бита $x, y \in \{0, 1\}$; затем он посылает x Алисе, а y Бобу. Получив входящий бит x, Алиса производит выходящий бит $a \in \{0, 1\}$, точно так же, получив y, Боб производит выходящий бит $b \in \{0, 1\}$. Но им запрещено общаться друг с другом, так что Алиса не знает y, а Боб не знает x.

Алиса и Боб побеждают в игре, если их выходящие биты окажутся связанными с входящими соотношением

$$a \oplus b = x \wedge y, \tag{4.50}$$

где $\oplus$ обозначает сложение по модулю два (вентиль XOR), а $\wedge$ обозначает произведение (вентиль AND). Могут ли Алиса и Боб найти стратегию, позволяющую им всегда выигрывать, независимо от того, какие входящие биты выбирает Чарли?

Нет, очевидно, что такой стратегии здесь нет. Пусть a_0, a_1 обозначают значения выходящих битов Алисы, если входящими были $x = 0, 1$, и пусть b_0, b_1 — выходящие биты Боба, соответствующие его входящим битам $y = 0, 1$. Чтобы Алиса и Боб выиграли при всех возможных входах, их выходящие биты должны удовлетворять

$$a_0 \oplus b_0 = 0, \quad a_0 \oplus b_1 = 0, \quad a_1 \oplus b_0 = 0, \quad a_1 \oplus b_1 = 1. \tag{4.51}$$

Однако это невозможно, так как, складывая эти четыре равенства, мы получим $0 = 1$.

Предположим, что Чарли генерирует входящие биты случайным образом. Тогда существует очень простая стратегия, позволяющая Алисе и Бобу выигрывать в трех случаях из четырех: они всегда выбирают выходящие биты $a = b = 0$, так что они проигрывают, если только входящие биты $x = y = 1$. Неравенство КГШХ может рассматриваться как утверждение того, что если Алиса и Боб делят не квантовое запутанное состояние, то лучшей стратегии нет.

Чтобы связать это утверждение с нашей предыдущей формулировкой неравенства КГШХ, определим случайные переменные, принимающие значения ± 1:

$$\begin{aligned} \mathbf{a} &= (-1)^{a_0}, \qquad \mathbf{a}' = (-1)^{a_1}, \\ \mathbf{b} &= (-1)^{b_0}, \qquad \mathbf{b}' = (-1)^{b_1}. \end{aligned} \tag{4.52}$$

Тогда неравенство КГШХ говорит, что при любом совместном распределении вероятностей, управляющем переменными $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}' \in \{0, 1\}$, ожидаемые значения удовлетворяют неравенству

$$\langle \mathbf{ab} \rangle + \langle \mathbf{ab}' \rangle + \langle \mathbf{a}'\mathbf{b} \rangle - \langle \mathbf{a}'\mathbf{b}' \rangle \leqslant 2. \tag{4.53}$$

Более того, если мы обозначим p_{xy} вероятность того, что уравнения (4.51) удовлетворяются, когда входящие биты равны (x, y), то

$$\begin{aligned} \langle \mathbf{ab} \rangle &= 2p_{00} - 1, \qquad \langle \mathbf{ab}' \rangle = 2p_{01} - 1, \\ \langle \mathbf{a}'\mathbf{b} \rangle &= 2p_{10} - 1, \qquad \langle \mathbf{a}'\mathbf{b}' \rangle = 1 - 2p_{11}; \end{aligned} \tag{4.54}$$

Например, $\langle \mathbf{ab} \rangle = p_{00} - (1 - p_{00}) = 2p_{00} - 1$, поскольку значение $\mathbf{ab}$ равно $+1$, когда Алиса и Боб выигрывают, и -1, когда они проигрывают.

Неравенство КГШХ (4.53) приобретает вид

$$2(p_{00} + p_{01} + p_{10} + p_{11}) - 4 \leqslant 2 \tag{4.55}$$

или

$$\langle p \rangle \equiv \frac{1}{4}(p_{00} + p_{01} + p_{10} + p_{11}) \leqslant \frac{3}{4}, \tag{4.56}$$

где $\langle p \rangle$ обозначает вероятность выигрыша, усредненную по однородному ансамблю входящих битов. Таким образом, если входящие биты случайны, то Алиса и Боб не могут достичь вероятности выигрыша, превосходящей $3/4$.

Имеет смысл рассмотреть, как предположение о том, что Алиса и Боб действуют под управлением «*локальных* скрытых переменных», ограничивает их успех в игре. Несмотря на то, что Алиса и Боб не делят квантовое запутанное состояние, им разрешено разделить таблицу случайных чисел, в соответствии с которой они могут генерировать их выходящие биты. Таким образом, мы можем представить, что Алиса и Боб принимают коррелированные решения, руководствуясь скрытыми переменными, извлекаемыми из ансамбля возможных значений. Эти корреляции ограничены локальностью — Алиса не знает входящих битов Боба, а Боб — входящих битов Алисы.

Но если Алиса и Боб делят квантовое запутанное состояние, то они могут изобрести стратегию получше. В зависимости от значения своего входящего бита, Алиса решает измерить одну из двух эрмитовых наблюдаемых с собственными значениями ± 1: $\mathbf{a}$, если $x = 0$, и $\mathbf{a}'$, если $x = 1$. Аналогично, Боб измеряет $\mathbf{b}$, если $y = 0$, и $\mathbf{b}'$, если $y = 1$. Тогда квантово-механические ожидаемые значения этих наблюдаемых удовлетворяют неравенству Цирельсона

$$\langle \mathbf{ab} \rangle + \langle \mathbf{ab}' \rangle + \langle \mathbf{a}'\mathbf{b} \rangle - \langle \mathbf{a}'\mathbf{b}' \rangle \leqslant 2\sqrt{2}, \tag{4.57}$$

а вероятность того, что Алиса и Боб выиграют гейм, ограничена условием

$$2(p_{00} + p_{01} + p_{10} + p_{11}) - 4 \leqslant 2\sqrt{2} \tag{4.58}$$

или

$$\langle p \rangle \equiv \frac{1}{4}(p_{00} + p_{01} + p_{10} + p_{11}) \leqslant \frac{1}{2} + \frac{1}{2\sqrt{2}} \approx 0{,}853. \tag{4.59}$$

Более того, мы видели, что это неравенство может перейти в равенство, если Алиса и Боб делят максимально запутанное состояние двух кубитов, а наблюдаемые $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$ выбраны подходящим образом.

Итак, мы обнаружили, что Алиса и Боб могут играть более успешно при наличии квантового запутывания, чем в его отсутствие. По крайней мере для этих целей, разделенное квантовое запутывание является более мощным средством, чем разделенная классическая случайность. Но даже ресурс квантового запутывания имеет свои пределы, устанавливаемые неравенством Цирельсона.

4.3.4. Все запутанные чистые состояния нарушают неравенства Белла

Сепарабельные состояния не нарушают неравества Белла. Например, если $\mathbf{a}$ является наблюдаемой, действующей на кубит Алисы, а $\mathbf{b}$ — наблюдаемой, действующей на кубит Боба, то в случае сепарабельного *чистого* состояния

$$\langle \mathbf{ab} \rangle = \langle \mathbf{a} \rangle \langle \mathbf{b} \rangle. \tag{4.60}$$

Никакого нарушения неравеств Белла не может быть, поскольку, как мы уже видели, действительно *существует* (локальная) теория скрытых переменных, которая корректно воспроизводит предсказания квантовой теории для чистого состояния одного кубита. Общее сепарабельное состояние представляет просто вероятностную смесь сепарабельных чистых состояний, так что корреляции между подсистемами являются полностью классическими и неравенства Белла применимы.

С другой стороны, мы видели, что максимально запутанное состояние, такое как $|\psi^-\rangle$, *нарушает* неравенства Белла. Но что можно сказать относительно чистого состояния, запутанного лишь частично, такого как

$$|\phi\rangle = \alpha|00\rangle + \beta|11\rangle\,? \tag{4.61}$$

Любое чистое состояние двух кубитов может быть выражено таким способом в базисе Шмидта; при подходящем соглашении относительно фаз α и β вещественные и неотрицательные.

Предположим, что Алиса и Боб выполняют измерение вдоль оси, лежащей в плоскости OXZ, так что их наблюдаемыми являются

$$\begin{aligned} \mathbf{a} &= \boldsymbol{\sigma}_3^{(A)} \cos\theta_A + \boldsymbol{\sigma}_1^{(A)} \sin\theta_A, \\ \mathbf{b} &= \boldsymbol{\sigma}_3^{(B)} \cos\theta_B + \boldsymbol{\sigma}_1^{(B)} \sin\theta_B. \end{aligned} \tag{4.62}$$

Состояние $|\phi\rangle$ обладает свойствами

$$\begin{gathered} \langle\phi|\boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_3|\phi\rangle = 1, \qquad \langle\phi|\boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1|\phi\rangle = 2\alpha\beta, \\ \langle\phi|\boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_1|\phi\rangle = \langle\phi|\boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_3|\phi\rangle = 0, \end{gathered} \tag{4.63}$$

так что квантово-механическое ожидаемое значение переменной $\mathbf{ab}$ равно

$$\langle \mathbf{ab} \rangle = \langle \phi | \mathbf{ab} | \phi \rangle = \cos\theta_A \cos\theta_B + 2\alpha\beta \sin\theta_A \sin\theta_B \tag{4.64}$$

[и мы воспроизводим $\cos(\theta_A - \theta_B)$ в максимально запутанном случае $\alpha = \beta = 1/\sqrt{2}$]. Теперь для простоты рассмотрим частный (не оптимальный!) случай

$$\theta_A = 0, \quad \theta'_A = \frac{\pi}{2}, \quad \theta'_B = -\theta_B, \tag{4.65}$$

так что квантовые предсказания равны

$$\begin{aligned} \langle \mathbf{ab} \rangle &= \cos\theta_B = \langle \mathbf{ab}' \rangle, \\ \langle \mathbf{a}'\mathbf{b} \rangle &= 2\alpha\beta \sin\theta_B = -\langle \mathbf{a}'\mathbf{b}' \rangle. \end{aligned} \tag{4.66}$$

Подставляя в неравенство КГШХ, получаем

$$|\cos\theta_B - 2\alpha\beta \sin\theta_B| \leqslant 1, \tag{4.67}$$

что очевидно нарушается при значениях θ_B, близких к 0 и π. Разлагая левую часть в линейном порядке по θ_B, имеем

$$\simeq 1 - 2\alpha\beta\theta_B, \tag{4.68}$$

что, конечно же, превосходит единицу при $\alpha\beta > 0$ и малом отрицательном θ_B.

Мы показали, что *любое* запутанное чистое состояние двух кубитов нарушает некоторое неравенство Белла. Это доказательство нетрудно обобщить на произвольное бинарное чистое состояние. То есть для бинарных чистых состояний «запутывание» эквивалентно «нарушению неравенства Белла». Однако, как мы увидим ниже, для бинарных смешанных состояний ситуация более тонкая.

4.3.5. Фотоны

Эксперименты по проверке неравенства Белла обычно выполняются на запутанных фотонах, а не на объектах со спином-1/2. Каковы квантово-механические предсказания для фотонов?

Вспомним из §2.2.2, что в случае фотонов, распространяющихся в направлении $\hat{z}$, мы используем обозначения $|x\rangle$, $|y\rangle$ для состояний, линейно поляризованных вдоль осей OX и OY соответственно. На языке этих базисных состояний, поляризованных вдоль «горизонтальной» и «вертикальной»

осей, состояния, повернутые на угол θ относительно OX и OY осей, могут быть выражены как

$$|H(\theta)\rangle = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \qquad |V(\theta)\rangle = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}. \tag{4.69}$$

Мы можем построить 2×2-матрицу, собственными состояниями которой являются $|H(\theta)\rangle$ и $|V(\theta)\rangle$ с соответствующими собственными значениями ± 1; она имеет вид

$$\tau(\theta) \equiv |H(\theta)\rangle\langle H(\theta)| - |V(\theta)\rangle\langle V(\theta)| = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}. \tag{4.70}$$

Генератором поворотов вокруг оси $\hat{z}$ является $\mathbf{J} = \sigma_2$, а собственными состояниями оператора $\mathbf{J}$ с собственными значениями ± 1 — циркулярно поляризованные состояния

$$|+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \qquad |-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} i \\ 1 \end{pmatrix}. \tag{4.71}$$

Предположим, что возбужденный атом излучает два фотона, которые вылетают в противоположных направлениях в состоянии с равным нулю суммарным угловым моментом и положительной четностью. Двухфотонные состояния

$$|+\rangle_A|-\rangle_B, \qquad |-\rangle_A|+\rangle_B \tag{4.72}$$

инвариантны относительно поворотов вокруг оси $\hat{z}$. Фотоны имеют противоположные значения J_z, но одинаковые *спиральности* (угловые моменты вдоль направления распространения), так как они распространяются в противоположных направлениях. При отражении в плоскости OYZ поляризованные состояния преобразуются согласно

$$|x\rangle \to -|x\rangle, \qquad |y\rangle \to |y\rangle \tag{4.73}$$

или

$$|+\rangle \to +i|-\rangle, \qquad |-\rangle \to -i|+\rangle; \tag{4.74}$$

следовательно, собственными состояниями четности являются запутанные состояния

$$\frac{1}{\sqrt{2}}\left(|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B\right). \tag{4.75}$$

Тогда состояние с $J_z = 0$ и положительной четностью, выраженное через линейно поляризованные состояния, имеет вид

$$-\frac{i}{\sqrt{2}}\left(|+-\rangle_{AB} + |-+\rangle_{AB}\right) = \frac{1}{\sqrt{2}}\left(|xx\rangle_{AB} + |yy\rangle_{AB}\right) \equiv |\phi^+\rangle_{AB}. \tag{4.76}$$

Вследствие инвариантности относительно поворотов вокруг оси $\hat{z}$, оно имеет такой вид независимо от того, как мы ориентируем OX и OY оси.

Алиса и Боб могут использовать анализатор поляризации, чтобы спроецировать состояния поляризации фотона на базис $\{|H(\theta)\rangle, |V(\theta)\rangle\}$ и, следовательно, измерить $\tau(\theta)$. Для двух фотонов в состоянии $|\phi^+\rangle$, если Алиса ориентирует свой анализатор под углом θ_A, а Боб — под углом θ_B, тогда корреляции результатов их измерений закодированы в ожидаемом значении

$$\langle\phi^+|\tau^{(A)}(\theta_A)\tau^{(B)}(\theta_B)|\phi^+\rangle. \tag{4.77}$$

С учетом вращательной симметрии:

$$\begin{aligned} &= \langle\phi^+|\tau^{(A)}(0)\tau^{(B)}(\theta_B-\theta_A)|\phi^+\rangle = \\ &= \frac{1}{2}\langle x|\tau^{(B)}(\theta_B-\theta_A)|x\rangle - \frac{1}{2}\langle y|\tau^{(B)}(\theta_B-\theta_A)|y\rangle = \\ &= \cos 2(\theta_B-\theta_A). \end{aligned} \tag{4.78}$$

Напомним, что в случае измерения кубитов на сфере Блоха мы находили подобное выражение $\cos\theta$, где θ — угол между направлениями поляризации у Алисы и Боба. Здесь вместо этого мы имеем $\cos 2\theta$, поскольку фотоны имеют спин-1, а не спин-1/2.

Если Алиса измеряет одну из двух наблюдаемых $\mathbf{a} = \tau^{(A)}(\theta_A)$ или $\mathbf{a}' = \tau^{(A)}(\theta'_A)$, а Боб измеряет или $\mathbf{b} = \tau^{(B)}(\theta_B)$, или $\mathbf{b}' = \tau^{(B)}(\theta'_B)$, то в предположении о существовании локальных скрытых переменных применимо неравенство КГШХ. Если мы подставляем квантовые предсказания для ожидаемых значений, то получим

$$|\cos 2(\theta_B-\theta_A)+\cos 2(\theta_B-\theta'_A)+\cos 2(\theta'_B-\theta_A)-\cos 2(\theta'_B-\theta'_A)| \leqslant 2. \tag{4.79}$$

Максимальное нарушение этого неравенства, при котором неравенство Цирельсона превращается в равенство, — левая часть равна $2\sqrt{2}$ — возникает, когда θ'_A, θ_B, θ_A и θ'_B последовательно разделены углами $22\frac{1}{2}^\circ$, так что

$$\begin{aligned} \frac{1}{\sqrt{2}} &= \cos 2(\theta_B-\theta_A) = \cos 2(\theta_B-\theta'_A) = \\ &= \cos 2(\theta'_B-\theta_A) = -\cos 2(\theta'_B-\theta'_A). \end{aligned} \tag{4.80}$$

4.3.6. Эксперименты и лазейки

Лазейка локальности. Экспериментами с запутанными парами фотонов было проверено неравенство КГШХ в форме (4.79). Эксперименты

подтвердили квантовые предсказания и убедительно продемонстрировали, что неравенство КГШХ нарушается. Следовательно, эти эксперименты, по всей видимости, показывают, что Природа не может корректно описываться теорией локальных скрытых переменных.

Но так ли это? Скептик может выдвинуть возражения. Например, при выводе неравенства КГШХ мы предполагали, что после того как Алиса решит, что измерять: a или a′, Боб не получает информации о ее решении прежде, чем он выполнит свои измерения (а также, если первым измерения выполняет Боб, то мы предполагаем, что информация о его решении не доходит до Алисы прежде, чем она выполнит свои измерения). С другой стороны, маргинальное распределение вероятностей для результатов измерений Боба может быть дополнено после измерений Алисы, но до измерений Боба, так что неравенство КГШХ становится неприменимым. Предположение о невозможности такого дополнения подтверждается релятивистской причинностью, если решение и измерение Алисы, как события, отделены от решения и измерения Боба пространственно-подобными интервалами. Скептик упорно настаивает, чтобы эксперимент удовлетворял этому условию, которое называется *лазейкой локальности*.

В 1982 г. Аспек с сотрудниками выполнили эксперимент с целью проверки лазейки локальности. Два запутанных фотона рождаются в результате распада возбужденного состояния атома кальция и поляризация каждого фотона ориентируется включением одного из двух псевдо-случайно выбранного анализатора поляризации. Фотоны регистрируются на удалении около 12 м от источника, что соответствует времени распространения света около 40 нс. Это время гораздо больше времени включения или разности времен прибытия обоих фотонов. Следовательно, «решение» о том, какую наблюдаемую измерять, принимается, когда фотоны уже находятся в полете, а события, состоящие в выборе осей для измерения поляризации фотонов A и B, разделены пространственно-подобным интервалом. Результаты согласуются с квантовыми предсказаниями и нарушают неравенство КГШХ на пять стандартных отклонений. После Аспека этот результат был подтвержден в других экспериментах, включая те, в которых детекторы A и B были удалены на километры.

Лазейка детектирования. Другое возражение, которое может выдвинуть скептик, называется *лазейкой детектирования*. В экспериментах с фотонами эффективность детектирования исключительно низкая. Большинство запутанных фотонных пар не регистрируются обоими детекторами A и B. Среди событий, ведущих к ошибке: фотон может быть поглощен, прежде чем он достигнет детектора, фотон может пролететь мимо детектора, или фотон может достичь детектора, но не быть им зарегистрирован-

ным. В эксперименте принимаются только те данные, которые получены при совпадении регистрации двух фотонов, поскольку, проверяя неравенство КГШХ, мы должны предполагать, что полученные данные представляют объективную выборку из всех запутанных пар.

Но что если локальные скрытые переменные управляют не только тем, *какое* состояние поляризации детектируется, но также и тем, *сработает ли вообще* детектор? Тогда полученные нами данные могут быть необъективной (смещенной) выборкой, а неравенство КГШХ — неприменимым.

В упражнении 4.2 мы покажем, что лазейку детектирования можно закрыть, если фотоны регистрируются с эффективностью около 82,84%. Современные эксперименты с фотонами далеки от требуемой эффективности. В экспериментах с ионными ловушками неравенство КГШХ было проверено с эффективностью детектирования, близкой к 100%, однако в этих экспериментах открыта (для скрытых переменных (перев.)) лазейка локальности. До сих пор не поставлено эксперимента, в котором одновременно были бы закрыты обе лазейки — локальности и детектирования.

Лазейка свободы воли. Предположим, что выполнен эксперимент, в котором фотоны регистрируются с идеальной эффективностью, а решения, принимаемые Алисой и Бобом, выглядят разделенными пространственно-подобным интервалом. Но скептик может продолжать сопротивляться выводу о том, что теории локальных скрытых переменных исключены, обращаясь к *лазейке свободы воли*. Предполагается, что принимаемые Алисой и Бобом решения о том, что измерять, сами управляются скрытыми переменными. Тогда их решения могут коррелировать со значениями скрытых переменных, которые определяют результаты измерения, следовательно, они не в состоянии получить объективную выборку из распределения скрытых переменных, а неравенство КГШХ может быть нарушено.

Каждый из нас сам решает для себя, насколько серьезно относиться к этому возражению.

4.4. Использование запутывания

После работы Белла квантовое запутывание стало предметом интенсивных исследований среди тех, кто интересуется основаниями квантовой теории. Постепенно сформировалась новая точка зрения: запутывание не только уникальный инструмент для демонстрации странностей квантовой механики, но и потенциально полезный *ресурс*. Используя запутывание квантовых состояний, мы можем решить задачи, сложные или неразрешимые при других подходах.

4.4.1. Плотное кодирование

Нашим первым примером является использование запутывания для связи. Алиса хочет послать сообщение Бобу. Она может послать классические биты (типа точек и тире азбуки Морзе), но предположим, что Алиса и Боб связаны *квантовым* каналом связи. Например, Алиса может приготовить кубиты (фотоны) в любом состоянии поляризации, в каком пожелает, и послать их Бобу, который измеряет поляризацию вдоль выбранной им оси. Существует ли какое-нибудь преимущество в отправлении кубитов вместо классических битов?

В принципе, если их квантовый канал имеет идеальную точность воспроизведения, а Алиса и Боб выполняют приготовление и измерение с идеальной эффективностью, тогда они *не будут испытывать затруднений*, используя кубиты вместо классических битов. Скажем, Алиса может приготовить или $|\uparrow_z\rangle$, или $|\downarrow_z\rangle$, а Боб может измерить вдоль $\hat{z}$, чтобы определить сделанный ей выбор. Таким образом, с каждым кубитом Алиса может послать один классический бит. Но фактически это максимум того, что она может сделать. Посылая по одному кубиту независимо от того, как она их готовит, и независимо от того, как Боб их измеряет, с каждым кубитом можно передать не более одного классического бита (даже если кубиты запутаны между собой). Это утверждение, частный случай предела Холево способности квантового канала пропускать классическую информацию, будет доказано в главе 5.

Теперь немного изменим правила — предположим, что Алиса и Боб делят запутанную пару кубитов в состоянии $|\phi^+\rangle_{AB}$. Пара была приготовлена в прошлом году; один кубит был отправлен Алисе, а другой — Бобу в надежде, что разделенное запутывание однажды пригодится. Использование квантового канала весьма дорого, так что Алиса может позволить себе послать Бобу только один кубит. Тем не менее для нее крайне важно сообщить Бобу *два* классических бита информации.

К счастью, Алиса помнит о запутанном состоянии $|\phi^+\rangle_{AB}$, которое она делит с Бобом, и выполняет протокол, который они с Бобом приготовили как раз для такого случая. На своей части запутанной пары она может выполнить одно из четырех возможных унитарных преобразований:

1) $\mathbf{1}$ (она ничего не делает),

2) σ_1 (поворот на 180° вокруг оси $\hat{x}$),

3) σ_2 (поворот на 180° вокруг оси $\hat{y}$),

4) σ_3 (поворот на 180° вокруг оси $\hat{z}$).

Как мы видели, делая это, она преобразует $|\phi^+\rangle_{AB}$ к одному из четырех взаимно ортогональных состояний:

1) $|\phi^+\rangle_{AB}$,

2) $|\psi^+\rangle_{AB}$,

3) $|\psi^-\rangle_{AB}$ (с точностью до фазы),

4) $|\phi^-\rangle_{AB}$.

Теперь она посылает свой кубит Бобу, который получает его и выполняет ортогональное коллективное измерение на паре, проецируя ее на максимально запутанный базис. Результат измерения недвусмысленно различает четыре возможных действия, которые Алиса могла выполнить. Следовательно, один кубит, посланный Алисой Бобу, успешно переносит два бита классической информации! Поэтому такая процедра называется «плотным кодированием».

Приятной особенностью этого протокола является то, что если сообщение строго конфиденциальное, то Алиса может не беспокоиться о том, что пересылаемый кубит перехватят враги и расшифруют ее сообщение. Перехваченный кубит имеет матрицу плотности $\boldsymbol{\rho}_A = \frac{1}{2}\mathbf{1}_A$ и не несет информации вообще. Вся информация в корреляциях между кубитами A и B, а она недоступна, до тех пор пока враг не заполучит обе части запутанной пары. (Но, конечно, он может «перекрыть» канал, препятствуя получению информации Бобом.)

С одной точки зрения Алисе и Бобу в действительности *нужно* дважды воспользоваться каналом для обмена двумя битами информации. Например, мы можем представить, что Алиса сама приготовила состояние $|\phi^-\rangle_{AB}$. В прошлом году она послала Бобу половину состояния, а теперь посылает вторую. То есть на самом деле Алиса посылала два кубита Бобу в одном из четырех взаимно ортогональных состояний, чтобы передать ему два классических бита информации, что допускает предел Холево.

Плотное кодирование является странным по ряду причин. Во-первых, Алиса послала Бобу первый кубит задолго до того, как узнала, каким будет ее сообщение. Во-вторых, каждый кубит сам по себе не несет никакой информации; она целиком закодирована в корреляциях между кубитами. В-третьих, это сработало бы с тем же успехом, если бы запутанную пару приготовил Боб и половину ее послал Алисе; тогда два классических бита передаются от Алисы к Бобу путем пересылки одного кубита от Боба к Алисе и обратно.

Так или иначе, если бы возникла необходимость и понадобилось немедленно послать два бита, в то время как каналом связи можно воспользоваться только один раз, Алиса и Боб могли бы использовать предварительно приготовленное запутывание для более эффективной связи. Они использовали бы запутывание как ресурс.

4.4.2. Квантовая телепортация

В плотном кодировании квантовая информация может быть использована для увеличения передачи классической информации. В частности, если Алиса и Боб делят запутанное состояние, то для передачи двух классических битов достаточно послать один кубит. Интересно обратное утверждение. Если Алиса и Боб делят запутанное состояние, то достаточно ли послать два классических бита, чтобы передать один кубит?

Представим, что Чарли приготовил для Алисы кубит в состоянии $|\psi\rangle$, но Алиса ничего не знает о том, какое состояние приготовил Чарли. Бобу отчаянно нужен этот кубит, и Алиса хочет помочь ему. Но проклятый квантовый канал снова закрыт! Алиса может послать Бобу только *классическую* информацию.

Она могла бы попытаться измерить $\vec{\sigma} \cdot \hat{n}$, проецируя свой кубит или на $|\uparrow_{\hat{n}}\rangle$, или на $|\downarrow_{\hat{n}}\rangle$, и послать однобитовый результат измерения Бобу, который тогда мог бы приступить к приготовлению обнаруженного Алисой состояния. Но, как вы покажете в упражнении 4.7, состояние Боба не будет идеальной копией состояния Алисы; в среднем он будет соответствовать кубиту Алисы с точностью воспроизведения

$$F = |\langle\varphi|\psi\rangle|^2 = \frac{2}{3}. \tag{4.81}$$

Эта точность воспроизведения выше той, которой Боб мог бы добиться просто случайным образом выбирая состояние $\left(F = \frac{1}{2}\right)$, но она далека от той, что ему требуется. Более того, как мы увидим в главе 5, не существует алгоритма, позволяющего таким способом (Алиса измеряет кубит и посылает классическую информацию Бобу) достичь точности воспроизведения выше, чем $2/3$.

К счастью, Алиса и Боб помнят, что они делят максимально запутанное состояние $|\phi^+\rangle_{AB}$, которое они приготовили в прошлом году. Почему бы им не использовать запутывание как *ресурс*? Если они готовы израсходовать разделенное запутанное состояние и общаться классическим образом, то может ли Алиса послать свой кубит Бобу с точностью воспроизведения выше, чем $2/3$?

На самом деле они могут добиться точности воспроизведения $F = 1$, выполняя следующий протокол: Алиса соединяет неизвестный кубит $|\psi\rangle_C$, который она хочет послать Бобу, с ее половиной $|\phi^+\rangle_{AB}$-пары, которую она делит с Бобом. Она измеряет две коммутирующие наблюдаемые

$$\boldsymbol{\sigma}_1^{(C)} \otimes \boldsymbol{\sigma}_1^{(A)}, \qquad \boldsymbol{\sigma}_3^{(C)} \otimes \boldsymbol{\sigma}_3^{(A)}, \tag{4.82}$$

выполняя таким образом *измерение Белла* – проекцию двух кубитов на одно из четырех максимально запутанных состояний $|\phi^\pm\rangle_{CA}$, $|\psi^\pm\rangle_{CA}$. Затем она посылает результаты своих измерений (два бита классической информации) Бобу по классическому каналу. Получив эту информацию, Боб выполняет одну из четырех операций над своим кубитом:

$$\begin{aligned}
&\text{Алиса измеряет } |\phi^+\rangle_{CA} \to \text{Боб применяет } \mathbf{1}^{(B)},\\
&\text{Алиса измеряет } |\psi^+\rangle_{CA} \to \text{Боб применяет } \boldsymbol{\sigma}_1^{(B)},\\
&\text{Алиса измеряет } |\psi^-\rangle_{CA} \to \text{Боб применяет } \boldsymbol{\sigma}_2^{(B)},\\
&\text{Алиса измеряет } |\phi^-\rangle_{CA} \to \text{Боб применяет } \boldsymbol{\sigma}_3^{(B)}.
\end{aligned} \tag{4.83}$$

Это действие преобразует кубит Боба (его часть запутанной пары, предварительно поделенной с Алисой) в идеальную копию $|\psi\rangle_C$. Этот магический трюк называется *квантовой телепортацией*.

Как она работает? Заметим, что для $|\psi\rangle = a|0\rangle + b|1\rangle$ мы можем записать

$$\begin{aligned}
|\psi\rangle_C|\phi^+\rangle_{AB} &= (a|0\rangle_C + b|1\rangle_C)\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) =\\
&= \frac{1}{\sqrt{2}}(a|000\rangle_{CAB} + a|011\rangle_{CAB} + b|100\rangle_{CAB} + b|111\rangle_{CAB}) =\\
&= \frac{a}{2}(|\phi^+\rangle_{CA} + |\phi^-\rangle_{CA})|0\rangle_B + \frac{a}{2}(|\psi^+\rangle_{CA} + |\psi^-\rangle_{CA})|1\rangle_B +\\
&+ \frac{b}{2}(|\psi^+\rangle_{CA} - |\psi^-\rangle_{CA})|0\rangle_B + \frac{b}{2}(|\phi^+\rangle_{CA} - |\phi^-\rangle_{CA})|1\rangle_B =\\
&= \frac{1}{2}|\phi^+\rangle_{CA}(a|0\rangle_B + b|1\rangle_B) + \frac{1}{2}|\psi^+\rangle_{CA}(a|1\rangle_B + b|0\rangle_B) +\\
&+ \frac{1}{2}|\psi^-\rangle_{CA}(a|1\rangle_B - b|0\rangle_B) + \frac{1}{2}|\phi^-\rangle_{CA}(a|0\rangle_B - b|1\rangle_B) =\\
&= \frac{1}{2}|\phi^+\rangle_{CA}|\psi\rangle_B + \frac{1}{2}|\psi^+\rangle_{CA}\boldsymbol{\sigma}_1|\psi\rangle_B +\\
&+ \frac{1}{2}|\psi^-\rangle_{CA}(-i\boldsymbol{\sigma}_2)|\psi\rangle_B + \frac{1}{2}|\phi^-\rangle_{CA}\boldsymbol{\sigma}_3|\psi\rangle_B.
\end{aligned} \tag{4.84}$$

Таким образом, мы видим, что, когда Алиса выполняет измерение Белла на кубитах C и A, все четыре исхода равновероятны. Как только Боб узнает результат ее измерения, он получает в свое распоряжение чистое состояние $\sigma|\psi\rangle$, где σ — известный оператор Паули, один из $\{1, \sigma_1, \sigma_2, \sigma_3\}$. Действие, предписываемое уравнением (4.83), восстанавливает кубит Боба в начальном состоянии $|\psi\rangle$.

Квантовая телепортация — любопытная процедура. Первоначально кубит Боба полностью некоррелирован с неизвестным кубитом $|\psi\rangle_C$, но выполненное Алисой измерение Белла устанавливает корреляцию между A и C. Результат ее измерения фактически совершенно случаен, следовательно, выполняя это измерение, Алиса (и Боб) в действительности не получают никакой информации относительно $|\psi\rangle$. Это особенно приятно. Ведь как известно, если бы они получили любую информацию о состоянии, то неизбежно внесли бы в него возмущение.

Как же в таком случае квантовому состоянию удается перейти от Алисы к Бобу? Это довольно загадочно. С одной стороны, мы едва ли можем уверенно сказать, что два отправленых классических бита несли эту информацию, поскольку они были случайными. Следовательно, невольно хочется сказать, что разделенная запутанная пара сделала возможной телепортацию. Вспомним, однако, что запутанная пара в действительности была приготовлена еще в прошлом году, когда Алисе даже не снилось, что она будет посылать кубит Бобу ...

Следует также заметить, что процесс телепортации полностью согласуется с принципом невозможности клонирования. В самом деле, в руках Боба оказалась копия состояния $|\psi\rangle_B$. Но прежде, чем она могла возникнуть, оригинал $|\psi\rangle_C$ был разрушен измерением Алисы.

Наши сведения о плотном кодировании и квантовой телепортации можно подытожить как утверждения о том, как ресурс одного типа может моделировать другой. Введем термины *забит* для разделенной на две части запутанной пары кубитов[1] и *c-бит* для классического бита (*c* от слова *classical* — классический). Мы телепортируем один кубит от Алисы к Бобу, расходуя один забит и посылая два *c*-бита, а с помощью плотного кодирования мы посылаем два *c*-бита от Алисы к Бобу, расходуя один забит и транспортируя один кубит. Таким образом, можно сказать, что

$$\begin{aligned} &1 \text{ забит} + 2\,c\text{-бита} \rightarrow 1 \text{ кубит}, \\ &1 \text{ забит} + 1 \text{ кубит} \rightarrow 2\,c\text{-бита} \end{aligned} \tag{4.85}$$

означает, что ресурсов левых частей достаточно для копирования правых частей. В этих алгоритмах существенно запутывание. Без забитов кубит

[1] В оригинале *ebit* (*e* от слова *entangled* — *запутанный*). — *Прим. перев.*

стоит не больше одного c-бита и без них же «телепортируемый» кубит имеет точность воспроизведения $F \leqslant 2/3$.

4.4.3. Квантовая телепортация и максимальное запутывание

Идея телепортации выглядит довольно таинственно. Хотелось бы глубже разобраться, почему она работает. Полезным ключом к разгадке является то, что для телепортации с точностью воспроизведения $F = 1$ расходуемое запутанное состояние согласно протоколу должно быть *максимально* запутанным. А основной особенностью бинарных максимально запутанных состояний является то, что *или Алиса, или Боб* могут преобразовывать одно такое состояние в другое, применяя локальное унитарное преобразование.

Чтобы лучше увидеть, как работает квантовая телепортация, рассмотрим телепортирование N-мерной системы, используя максимально запутанное $N \times N$-состояние вида

$$|\Phi\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle \otimes |i\rangle. \tag{4.86}$$

Полезным свойством этого состояния является

$$\begin{aligned} {}_{CA}\langle\Phi|\Phi\rangle_{AB} &= \frac{1}{N} \sum_{i,j} \left({}_C\langle i| \otimes {}_A\langle j| \right) \left(|j\rangle_A \otimes |i\rangle_B \right) = \\ &= \frac{1}{N} \sum_i |i\rangle_B \, {}_C\langle i| = \frac{1}{N} \mathbf{T}_{BC}. \end{aligned} \tag{4.87}$$

Здесь мы определили *трансфер-оператор* (или *оператор переноса*) $\mathbf{T}_{BC}$, который обладает свойством

$$\mathbf{T}_{BC}|\varphi\rangle_C = \mathbf{T}_{BC} \left(\sum_i a_i |i\rangle_C \right) = \sum_i a_i |i\rangle_B = |\varphi\rangle_B ; \tag{4.88}$$

он отображает состояние из C на идентичное состояние в B. Это свойство не имеет инвариантного смысла, независимого от выбора базиса в B и C; скорее $\mathbf{T}_{BC}$ просто описывает произвольный способ связать ортонормированные базисы двух систем. Конечно, Алисе и Бобу придется определенным образом ориентировать свои базисы, чтобы проверить, что телепортация действительно состоялась.

Теперь вспомним, что любое другое максимально запутанное $N \times N$-состояние имеет разложение Шмидта вида

$$\frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i'\rangle \otimes |i\rangle \tag{4.89}$$

и может быть выражено как

$$|\Phi(\mathbf{U})\rangle \equiv \mathbf{U} \otimes \mathbf{1}|\Phi\rangle, \tag{4.90}$$

где

$$\mathbf{U}|i\rangle = |i'\rangle = \sum_j |j\rangle U_{ji}. \tag{4.91}$$

Записывая

$$|\Phi(\mathbf{U})\rangle_{AB} = \frac{1}{\sqrt{N}} \sum_{i,j} |j\rangle_A \otimes |i\rangle_B U_{ji}, \tag{4.92}$$

можно легко проверить, что

$${}_{CA}\langle\Phi(\mathbf{U})|\Phi(\mathbf{V}^T)\rangle_{AB} = \frac{1}{N}(\mathbf{V}\mathbf{U}^{-1})_B \mathbf{T}_{BC}, \tag{4.93}$$

где $\mathbf{V}^T$ обозначает транспонированную матрицу $\mathbf{V}$ в стандартном базисе ($V_{ij}^T = V_{ji}$); тогда, в частности, для любой унитарной матрицы $\mathbf{U}$ трансфер-оператор может быть представлен в виде

$$\frac{1}{N}\mathbf{T}_{BC} = {}_{CA}\langle\Phi(\mathbf{U})|\Phi(\mathbf{U}^T)\rangle_{AB}. \tag{4.94}$$

Предположим теперь, что Алиса и Боб делят $|\Phi\rangle_{AB}$, а Чарли приготовил состояние $|\psi\rangle_C$ и оставил его на хранение в лаборатории Алисы. Алиса выполняет измерение, которое проецирует CA на базис максимально запутанных состояний, получая результат $|\Phi(\mathbf{U}_a)\rangle_{CA}$ для некоторого унитарного преобразования $\mathbf{U}_a$. Тогда из уравнения (4.94) известно, что *если бы* Алиса и Боб вместо $|\Phi\rangle_{AB}$ поделили состояние $|\Phi(\mathbf{U}_a^T)\rangle_{CA}$, то измерение Алисы приготовило бы в лаборатории Боба идеальную копию (реплику) состояния $|\psi\rangle$. К сожалению, они не догадались с самого начала поделить подходящее состояние. Но еще не все потеряно! Боб понимает, что

$$|\Phi(\mathbf{U}_a^T)\rangle_{AB} = \mathbf{1}_A \otimes (\mathbf{U}_a)_B |\Phi\rangle_{AB}, \tag{4.95}$$

и, конечно, $(\mathbf{U}_a)_B$ коммутирует с измерением Алисы. Следовательно, когда Боб узнает у Алисы, что результатом ее измерения было $|\Phi(\mathbf{U}_a^T)\rangle_{AB}$, он применит $(\mathbf{U}_a)_B$ к своей половине поделенного с Алисой состояния. Тогда протокол станет эквивалентным тому, в котором они с самого начала делили именно то, какое нужно, максимально запутанное состояние, а состояние Боба преобразуется в $|\psi\rangle_B$!

Этот подход к телепортации имеет некоторые концептуальные преимущества. Во-первых, можно легко убедиться в том, что Алисе не требуется выполнять ортогональное измерение. Чтобы осуществить телепортацию с точностью воспроизведения $F = 1$, ей достаточно выполнить ПОЗМ с операторами $\mathbf{M}_a$, где каждый $\mathbf{M}_a$ обладает свойством

$$\mathbf{M}_a^\dagger \mathbf{M}_a \propto |\Phi(\mathbf{U}_a)\rangle\langle\Phi(\mathbf{U}_a)| \tag{4.96}$$

для некоторого унитарного преобразования $\mathbf{U}_a$. Так же легко можно увидеть, как должен быть модифицирован протокол телепортации, если начальным максимально запутанным состоянием, которое делят Алиса и Боб, является не $|\Phi\rangle_{AB}$, а

$$|\Phi(\mathbf{V}^T)\rangle_{AB} = \mathbf{1}_A \otimes \mathbf{V}_B |\Phi\rangle_{AB}. \tag{4.97}$$

Если результатом измерения Алисы является $|\Phi(\mathbf{U}_a)\rangle_{CA}$, то уравнение (4.93) говорит нам, что состояние Боба принимает вид

$$\mathbf{V}\mathbf{U}_a{}^{-1}|\psi\rangle_B. \tag{4.98}$$

Чтобы воспроизвести $|\psi\rangle_B$, Боб должен применить преобразование $\mathbf{U}_a\mathbf{V}^{-1}$.

Порядок следования операторов в уравнении (4.98) на первый взгляд может показаться интуитивно непонятным — он выглядит так, как если бы измерение Алисы $(\mathbf{U}_a)$ предшествовало приготовлению разделяемого запутанного состояния $(\mathbf{V})$. Однако это «обращение времени» имеет непосредственное толкование. Если результатом измерения Алисы является $|\Phi(\mathbf{U}_a)\rangle_{CA}$, то Боб получил бы идеальную копию $|\psi\rangle$, если бы начальным состоянием было $\mathbf{1}_A \otimes (\mathbf{U}_a)_B |\Phi\rangle_{AB}$. Чтобы смоделировать ситуацию, в которой сразу было подходящим образом выбрано запутанное состояние, Боб сначала применяет $\mathbf{V}^{-1}$, чтобы скомпенсировать «поворот» в $|\Phi(\mathbf{V}^T)\rangle_{AB}$ и восстановить $|\Phi\rangle_{AB}$, а затем применяет $\mathbf{U}_a$, чтобы преобразовать запутанное состояние к требуемому виду.

Существует более фантастическая интерпретация уравнения (4.98), которая хотя и необязательна, но тем не менее неопровержима. Мы можем «объяснить», как квантовая информация переносится от Алисы к Бобу, следуя движению кубита вдоль мировой линии в пространстве-времени. Кубит движется вперед во времени от его приготовления Чарли до измерения Алисой, затем – назад от измерения до первоначального приготовления запутанной пары и, наконец, снова вперед во времени от приготовления пары до лаборатории Боба. Поскольку эта мировая линия посещает измерение Алисы прежде чем добирается до приготовления запутанной пары, $\mathbf{U}_a^{-1}$ действует «первым», а $\mathbf{V}$ – «позднее».

4.4.4. Квантовый программный продукт

Телепортация имеет некоторые интересные приложения. Представим, например, что Алиса и Боб хотят применить «квантовый вентиль» $\mathbf{V}$ к неизвестному состоянию $|\psi\rangle_C$. Но применение $\mathbf{V}$ требует сложного оборудования, которое они себе не могут позволить.

Более экономичная альтернатива – *приобрести квантовый программный продукт* – бинарное состояние, которым, как уверяет продавец, является

$$|\Phi(\mathbf{V}^T)\rangle_{AB} = \mathbf{1}_A \otimes \mathbf{V}_B |\Phi\rangle_{AB}. \tag{4.99}$$

Аппаратное обеспечение Алисы достаточно мощное, чтобы выполнить измерение, проецирующее на базис $\{|\Phi(\mathbf{U}_a)\rangle_{CA}\}$; коль скоро результат a известен, состояние $\mathbf{V}\mathbf{U}_a^{-1}|\psi\rangle_B$ – приготовлено. Тогда Боб может завершить выполнение $\mathbf{V}$ на $|\psi\rangle_B$, применяя преобразование $\mathbf{V}\mathbf{U}_a\mathbf{V}^{-1}$.

Эта процедура может показаться неразумной – почему мы считаем, что Боб может применить преобразование $\mathbf{V}\mathbf{U}_a\mathbf{V}^{-1}$, но не способен применить $\mathbf{V}$? В действительности это не так глупо, а имеет важные применения к отказоустойчивым квантовым вычислениям, которые мы будем изучать позднее в главе 8[1]. В некоторых случаях *выполнение* $\mathbf{V}\mathbf{U}_a\mathbf{V}^{-1}$ в действительности несколько проще, нежели применение $\mathbf{V}$. Более того, вместо того, чтобы приобретать квантовое программное обеспечение, Алиса и Боб могут сами приготовить его, даже несмотря на то, что они не могут надежно применить $\mathbf{V}$. Это возможно, поскольку проще проверить, что было должным образом приготовлено *известное* квантовое состояние, чем проверять, что известное унитарное преобразование было успешно применено

[1] В это издание вошли первые шесть глав лекций Прескилла. Редакция РХД предполагает издание второй части, которая будет посвящена теории квантовых кодов, исправляющих ошибки, отказоустойчивых вычислений, топологических квантовых вычислений и другим вопросам. – *Прим. ред.*

к неизвестному состоянию. Если нельзя положиться на применяющее **V** аппаратное обеспечение, то мы предпочтем использовать его автономно для подготовки компьютерной программы, чтобы применить ее с гарантированной надежностью, нежели рисковать нанести неустранимое повреждение нашему неизвестному состоянию вследствие ошибочного выполнения **V**.

С каждым применением **V** расходуется одна копия квантового программного обеспечения. Таким образом, протокол выполнения преобразования **V** с его помощью использует запутывание как ресурс.

4.5. Квантовая криптография

4.5.1. Распределение квантового ЭПР-ключа

У каждого есть свои секреты, Алиса и Боб не исключение. Алисе нужно передать Бобу очень личное сообщение, но у них есть очень любопытная подружка, Ева, которая наверняка попытается их подслушать. Могут ли они связаться, будучи уверенными, что Еве это не удастся?

Очевидно, им нужно воспользоваться каким-то кодом. Но беда в том, что Ева не только очень любопытна, но и весьма ловка. Алиса и Боб не уверены, что им хватит ума придумать такой код, который Ева не сможет взломать, за исключением одной схемы кодирования, которая абсолютно надежна. Если Алиса и Боб поделят *тайный ключ*, известную только им случайную последовательность битов, тогда Алиса может конвертировать свое письмо в кодах ASCII (ряд битов не длиннее ключа), *сложив* (по модулю два) каждый бит своего сообщения с соответствующим битом ключа, и послать результат Бобу. Получив этот ряд, Боб может добавить ключ, чтобы извлечь сообщение Алисы.

Эта схема надежна, так как, даже если Ева перехватит сообщение, она ничего не сможет узнать, поскольку передаваемая последовательность сама по себе не несет никакой информации — сообщение закодировано в корреляции между передаваемой строкой и *ключом* (который Ева не знает).

Тем не менее проблема все еще остается, поскольку Алисе и Бобу необходимо установить общий случайный ключ и они должны быть уверены, что Ева не сможет его узнать. Они могли бы встретиться, чтобы обменяться ключом, но это может оказаться невыполнимо. Они могли бы доверить третьему лицу передать этот ключ, но что если оно состоит в тайном сговоре с Евой? Они могли бы использовать протоколы распределения «открытых ключей», но их надежность опирается на предположения относительно вычислительных ресурсов, доступных потенциальному противнику. Действительно, в главе 6 мы увидим, что протоколы открытых

ключей беззащитны перед атакой хакера, располагающего квантовым компьютером.

Могут ли Алиса и Боб использовать *квантовую* информацию (и особенно запутывание) для решения проблемы передачи ключа? Могут! Можно придумать протоколы *распределения квантовых ключей*, которые будут неуязвимы для любой, допустимой законами физики, атаки.

Предположим, что Алиса и Боб делят запас запутанных пар, приготовленных в состоянии $|\phi^+\rangle$. Чтобы приготовить известный только им тайный ключ, они должны выполнить следующий протокол.

Для каждого находящегося в их распоряжении кубита Алиса и Боб решают измерять *или* σ_1, *или* σ_3. Это решение является псевдо-случайным, каждый выбор реализуется с вероятностью $1/2$. Затем, после того как измерения выполнены, они открыто объявляют о том, какие наблюдаемые были измерены, но не открывают полученные ими результаты. В тех случаях (примерно в половине), в которых они измерили свои кубиты вдоль разных осей, их результаты отбрасываются (поскольку в них получены нескоррелированные результаты). В тех же случаях, в которых они выполнили измерения вдоль одних и тех же осей, их результаты хотя и случайны, но *идеально скоррелированы*. Следовательно, они установили между собой случайный ключ.

Но действительно ли этот протокол неуязвим перед коварной атакой Евы? В частности, еще раньше Ева могла тайком исказить пары. Тогда пары, которыми располагают Алиса и Боб, могут и не находиться в идеальных $|\phi^+\rangle$-состояниях, а скорее они будут запутаны с кубитами Евы (без ведома Алисы и Боба). Тогда Ева может подождать до тех пор, пока Алиса и Боб не сделают своего заявления, чтобы соответствующим образом выполнить измерение своих кубитов и получить максимальную информацию о полученных ими результатах. Алиса и Боб должны защититься от подобной атаки.

Если Ева действительно исказила пары Алисы и Боба, тогда наиболее общее возможное состояние AB-пары и множества E-кубитов имеет вид

$$|\Upsilon\rangle_{ABE} = |00\rangle_{AB}|e_{00}\rangle_E + |01\rangle_{AB}|e_{01}\rangle_E + \\ + |10\rangle_{AB}|e_{10}\rangle_E + |11\rangle_{AB}|e_{11}\rangle_E, \qquad (4.100)$$

где состояния кубитов Евы $|e_{ij}\rangle_E$ ни нормированы, ни взаимно ортогональны. Вспомним теперь, что определяющим свойством $|\phi^+\rangle$ является то, что оно представляет собой собственное состояние как $\sigma_1^{(A)} \otimes \sigma_1^{(B)}$, так и $\sigma_3^{(A)} \otimes \sigma_3^{(B)}$ с собственным значением $+1$. Предположим, что Алиса и Боб

могут проверить, обладают ли этим свойством имеющиеся у них кубиты. Чтобы удовлетворялось $\boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)} = 1$, мы должны иметь

$$|\Upsilon\rangle_{ABE} = |00\rangle_{AB}|e_{00}\rangle_E + |11\rangle_{AB}|e_{11}\rangle_E, \tag{4.101}$$

а чтобы выполнялось $\boldsymbol{\sigma}_1^{(A)} \otimes \boldsymbol{\sigma}_1^{(B)} = 1$, мы должны иметь

$$|\Upsilon\rangle_{ABE} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})|e\rangle_E = |\phi^+\rangle_{AB}|e\rangle_E. \tag{4.102}$$

Мы видим, что AB-пары могут быть собственными состояниями операторов $\boldsymbol{\sigma}_1^{(A)} \otimes \boldsymbol{\sigma}_1^{(B)}$ и $\boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)}$, если только они полностью незапутаны с кубитами Евы. Следовательно, измеряя свои кубиты, она не сможет что-либо узнать о результатах измерений Алисы и Боба. Случайный ключ надежен.

Чтобы проверить свойства $\boldsymbol{\sigma}_1^{(A)} \otimes \boldsymbol{\sigma}_1^{(B)} = 1 = \boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)}$, Алиса и Боб могут пожертвовать частью своего общего ключа и открыто сравнить результаты своих измерений. Они должны обнаружить, что их результаты действительно идеально скоррелированы. Если это так, то с высокой статистической надежностью они будут уверены в том, что Ева не в состоянии перехватить ключ. Если нет, то они зарегистрировали гнусную деятельность Евы. Тогда они могут выбросить этот ключ и сделать новую попытку установить надежный ключ.

Как я только что это представил, протокол распределения квантового ключа, казалось бы, требует наличия разделенных между Алисой и Бобом запутанных пар, но на самом деле это не так. Мы можем представить, что Алиса сама готовит пары $|\phi^+\rangle$, а затем измеряет один кубит в каждой паре, прежде чем послать другой Бобу. Это полностью эквивалентно схеме, в которой Алиса готовит одно из четырех состояний

$$|\uparrow_z\rangle, \quad |\downarrow_z\rangle, \quad |\uparrow_x\rangle, \quad |\downarrow_x\rangle \tag{4.103}$$

(выбираемое случайным образом, каждое из них возникает с вероятностью $1/4$) и посылает кубит Бобу. Тогда измерение Боба и проверка выполняются, как и раньше. Эта схема (известная как протокол распределения квантового ключа BB84 [1]) так же надежна, как и схема, основанная на запутывании[2].

[1] Предложен Беннетом и Брассаром в 1984 г.: C.H. Bennett, G. Brassard, in *Proc. IEEE Int. Conf. on Computers, Systems and Signal Processing*, IEEE, New York (1984). Экспериментально реализован в экспериментах с поляризованными фотонами. Детальное обсуждение можно найти в книге *Физика квантовой информации*, под ред. Д. Боумейстера, А. Экерта и А. Цайлингера, Постмаркет, М.: (2002). — *Прим. ред.*

[2] За исключением того, что в ЭПР-схеме Алиса и Боб могут подождать с созданием ключа до тех пор, пока им не понадобится поговорить, сокращая таким образом риск того, что в какой-то момент Ева может совершить взлом, чтобы узнать, какие состояния приготовила Алиса (и таким образом извлечь ключ).

Другой интригующий вариант называется «обращенной во времени ЭПР» схемой. Здесь и Алиса и Боб готовят по одному из четырех состояний (4.103) и отсылают свои кубиты Чарли. Тогда Чарли выполняет измерение Белла на паре, то есть он измеряет $\sigma_1^{(A)} \otimes \sigma_1^{(B)}$ и $\sigma_3^{(A)} \otimes \sigma_3^{(B)}$, совершая ортогональную проекцию на одно из состояний $|\phi^{\pm}\rangle$, $|\psi^{\pm}\rangle$, и открыто объявляет о результате. Поскольку все четыре из этих состояний одновременно являются собственными состояниями операторов $\sigma_1^{(A)} \otimes \sigma_1^{(B)}$ и $\sigma_3^{(A)} \otimes \sigma_3^{(B)}$, когда Алиса и Боб приготовили свои спины ориентированными вдоль одной и той же оси (что они делают примерно в половине случаев), они делят один бит[1]. Конечно, Чарли может оказаться в союзе с Евой, но, как и прежде, путем сравнения части своих кодов, Алиса и Боб могут проверить, что те не имели доступа к информации. Эта схема имеет то преимущество, что Чарли мог бы заведовать центральной коммутаторной станцией, храня кубиты, полученные от многих людей, и выполняя измерение Белла, когда двое из абонентов запросят установить безопасную связь. (Здесь мы предполагаем, что Чарли имеет устойчивую квантовую память, в которой кубит может храниться надежно и сколь угодно долго.) Безопасный ключ может быть установлен даже при временно закрытой линии квантовой связи, если оба абонента догадались послать свои кубиты Чарли раньше (когда квантовый канал был открыт).

До сих пор мы делали нереалистичное предположение о том, что квантовый канал связи идеален, но, конечно, в реальном мире будут возникать ошибки. Следовательно, даже если Ева не причинила никакого ущерба, Алиса и Боб иногда будут обнаруживать, что их проверочный тест терпит неудачу. Но как им отличить ошибки, возникающие из-за дефектов канала, от ошибок, возникающих в результате вторжения Евы?

Обращаясь к этой проблеме, Алиса и Боб могут усовершенствовать их протокол в двух отношениях. Во-первых, они могут осуществить (классическую) коррекцию ошибок, чтобы сократить эффективную частоту их появления. Например, чтобы установить каждый бит их общего ключа, они могут в действительности заменить его блоком трех случайных битов. Если среди трех битов не все одинаковые, то Алиса может сообщить Бобу, какой из них отличается от двух других; Боб может инвертировать этот бит в своем блоке, а *затем* использовать подсчет большинства голосов для определения значения бита для блока. Таким способом Алиса и Боб разделят одинаковый бит ключа, даже если для одного бита в блоке из трех возникла ошибка.

[1] Пока Чарли не выполнил свое измерение, состояния, приготовленные Алисой и Бобом, полностью некоррелированы. Определенная корреляция (или антикорреляция) устанавливается после того, как Чарли выполнит свое измерение.

Однако одной лишь коррекции ошибок недостаточно для уверенности в том, что Ева не получила информацию о ключе — коррекция ошибок должна быть дополнена (классическим) секретным усилением. Например, после выполнения коррекции ошибок, когда Алиса и Боб уже уверены, что располагают одинаковыми ключами, они могут выделить бит «суперключа», например, *четность* n битов ключа. Чтобы узнать *что-нибудь* о четности n битов, Еве нужно *хотя бы что-нибудь* узнать о каждом бите. Следовательно, бит четности в среднем существенно более надежен, чем каждый из отдельных битов ключа.

Если частота ошибок в канале достаточно низка, то можно показать, что распределение квантового ключа, дополненное коррекцией ошибок и секретным усилением, неуязвимо для любой атаки, которую может предпринять Ева (в том смысле, что можно гарантировать, что полученная ею информация будет сколь угодно мала). Мы вернемся к проблеме обеспечения безопасности распределения квантового ключа в главе 7.

4.5.2. Невозможность клонирования

Безопасность распределения квантового ключа основана на существенном различии между квантовой и классической информацией. Невозможно получить информацию, *определяющую различие* между неортогональными квантовыми состояниями, не *внося возмущение* в эти состояния.

Например, в протоколе BB84 Алиса посылает Бобу любое из четырех состояний, $|\uparrow_z\rangle$, $|\downarrow_z\rangle$, $|\uparrow_x\rangle$, $|\downarrow_x\rangle$: и они имеют возможность проверить, что ни одно из этих состояний не возмущено попыткой подслушивания со стороны Евы. В более общем виде предположим, что $|\varphi\rangle$ и $|\psi\rangle$ — два неортогональных состояния в $\mathcal{H}$ ($\langle\psi|\varphi\rangle \neq 0$) и что в $\mathcal{H}\otimes\mathcal{H}_E$ (где $\mathcal{H}_E$ — доступное Еве гильбертово пространство) применяется унитарное преобразование U, не возмущающее $|\varphi\rangle$ и $|\psi\rangle$. Тогда

$$U\colon \begin{cases} |\psi\rangle\otimes|0\rangle_E \to |\psi\rangle\otimes|e\rangle_E, \\ |\varphi\rangle\otimes|0\rangle_E \to |\varphi\rangle\otimes|f\rangle_E, \end{cases} \tag{4.104}$$

а унитарность предполагает, что

$$\begin{aligned} \langle\psi|\varphi\rangle &= \bigl({}_E\langle 0|\otimes\langle\psi|\bigr)\bigl(|\varphi\rangle\otimes|0\rangle_E\bigr) = \\ &= \bigl({}_E\langle e|\otimes\langle\psi|\bigr)\bigl(|\varphi\rangle\otimes|f\rangle_E\bigr) = \langle\psi|\varphi\rangle\langle e|f\rangle. \end{aligned} \tag{4.105}$$

Таким образом, при $\langle\psi|\varphi\rangle \neq 0$ мы имеем $\langle e|f\rangle = 1$ и, поскольку состояния нормированы, $|e\rangle = |f\rangle$. Это означает, что ни одно измерение в $\mathcal{H}_E$ не

может дать информацию, отличающую $|\psi\rangle$ от $|\varphi\rangle$. В случае BB84 это доказательство показывает, что если Ева не вносит возмущения в состояния, посланные Алисой, то состояния в $\mathcal{H}_E$ остаются неизменными, независимо от того, какое из четырех состояний $|\uparrow_z\rangle$, $|\downarrow_z\rangle$, $|\uparrow_x\rangle$, $|\downarrow_x\rangle$ было послано, и, следовательно, Ева ничего не узнает о разделенном Алисой и Бобом ключе. С другой стороны, если Алиса посылает Бобу одно из двух ортогональных состояний $|\uparrow_z\rangle$ или $|\downarrow_z\rangle$, то ничто не мешает Еве получить копию информации (как в случае с классическими битами).

Ранее мы отмечали, что если у нас есть множество идентичных копий кубита, то можно измерить средние значения некоммутирующих наблюдаемых типа σ_1, σ_2 и σ_3, чтобы полностью определить матрицу плотности кубита. Неотъемлемым в выводе о том, что неортогональные состояния нельзя различить не возмутив их, является неявное утверждение, что невозможно сделать идеальную копию кубита. (Если бы мы могли, мы сделали бы столько копий, сколько их необходимо для определения $\langle\sigma_1\rangle$, $\langle\sigma_2\rangle$ и $\langle\sigma_3\rangle$ с любой наперед заданной точностью.) Сформулируем это в явном виде: не существует квантового ксерокса.

Ортогональные квантовые состояния (подобные классической информации) *могут* надежно копироваться. Например, унитарное преобразование, действующее как

$$U:\ \begin{cases} |0\rangle_A|0\rangle_E \to |0\rangle_A|0\rangle_E, \\ |1\rangle_A|0\rangle_E \to |1\rangle_A|1\rangle_E, \end{cases} \tag{4.106}$$

копирует первый кубит на второй, если первый является одним из двух состояний: $|0\rangle_A$ или $|1\rangle_A$. Но если вместо этого первый кубит находится в состоянии $|\psi\rangle = a|0\rangle_A + b|1\rangle_A$, то

$$U:\ \big(a|0\rangle_A + b|1\rangle_A\big)|0\rangle_E \to a|0\rangle_A|0\rangle_E + b|1\rangle_A|1\rangle_E. \tag{4.107}$$

Это не состояние $|\psi\rangle \otimes |\psi\rangle$ (тензорное произведение исходного состояния и его копии); скорее это нечто совершенно отличное — запутанное состояние двух кубитов.

Чтобы рассмотреть наиболее общий квантовый ксерокс, допустим, что полное гильбертово пространство шире тензорного произведения исходного пространства и пространства копий. Тогда наиболее общее «копирующее» унитарное преобразование действует как

$$U:\ \begin{cases} |\psi\rangle_A|0\rangle_E|0\rangle_F \to |\psi\rangle_A|\psi\rangle_E|e\rangle_F, \\ |\varphi\rangle_A|0\rangle_E|0\rangle_F \to |\varphi\rangle_A|\varphi\rangle_E|f\rangle_F. \end{cases} \tag{4.108}$$

Тогда унитарность предполагает, что

$$ {}_A\langle\psi|\varphi\rangle_A = {}_A\langle\psi|\varphi\rangle_A \, {}_E\langle\psi|\varphi\rangle_E \, {}_F\langle e|f\rangle_F ; \tag{4.109} $$

следовательно, если ${}_A\langle\psi|\varphi\rangle_A \neq 0$, то

$$ 1 = {}_E\langle\psi|\varphi\rangle_E \, {}_F\langle e|f\rangle_F . \tag{4.110} $$

Поскольку состояния нормированы, мы приходим к выводу, что

$$ |\langle\psi|\varphi\rangle| = 1, \tag{4.111} $$

то есть $|\psi\rangle$ и $|\varphi\rangle$ в действительности представляют один и тот же луч. Ни одна унитарная машина не может сделать копии $|\varphi\rangle$ и $|\psi\rangle$, если они являются *различными неортогональными состояниями*. Этот результат называется теоремой о невозможности клонирования.

4.6. Многокомпонентное запутывание

4.6.1. Три квантовых ящика

После безумно успешного эксперимента с тремя монетами на столе Алиса и Боб стали всемирно известными. Они стали профессорами, Алиса в КАЛТЕХе, а Боб в Чикаго. Они слишком занятые, чтобы проводить много времени в лабораториях, но у них достаточно аспирантов и они продолжают активно заниматься наукой.

Их лучший студент, Чарли, который выполнял всю черновую работу в эксперименте с монетами, закончил образование и теперь он доцент в Принстоне. Алиса и Боб хотели бы содействовать карьере Чарли и помочь ему занять постоянную должность. Однажды они болтали по телефону.

Алиса: Знаешь, Боб, мы, конечно, должны помочь Чарли. Ты можешь придумать подходящий эксперимент, который мы можем выполнить втроем?

Боб: Ну, я не знаю, Алиса. Есть множество экспериментов, которые я хотел бы выполнить с нашими запутанными парами кубитов. Но в каждом эксперименте есть один кубит для меня, а другой — для тебя. Похоже, что Чарли — третий лишний.

Алиса: [*Длинная пауза*]. Боб... А ты когда-нибудь думал о постановке эксперимента с тремя кубитами?

У Боба отвисла челюсть и подскочил пульс. Во внезапном прозрении он словно увидел перед собой всю свою будущую карьеру. По правде говоря, Боб уже начинал задумываться о том, что их эксперименты с двумя кубитами порядком устарели. Теперь он знает, что в ближайшие пять лет он всецело посвятит себя выполнению полного трехкубитового эксперимента. К тому времени он, Алиса и Чарли обучат другого блестящего студента и будут готовы подступиться к четырем кубитам. Затем еще один студент и еще один кубит. И так до самой пенсии.

Вот как выглядит эксперимент с тремя кубитами, который Алиса и Боб решили попробовать осуществить: Алиса поручает сотруднику своей лаборатории в КАЛТЕХе тщательно приготовить состояние трех квантовых ящиков. (Но Алиса не знает точно, как он это сделает.) Она оставляет один ящик у себя, а два других отправляет срочной квантовой почтой Бобу и Чарли. В каждом ящике находится шар, который может быть или черным, или белым, но ящик плотно закрыт. Единственная возможность узнать, что находится внутри, — это открыть ящик, но открыть его можно двумя различными способами — ящик имеет две дверцы, промаркированные как X и Y. Определить цвет шара можно, когда открывается одна из двух дверок. Но невозможно открыть обе дверцы сразу.

Алиса, Боб и Чарли собираются исследовать, как скоррелированы ящики. Они проводят множество тщательно контролируемых испытаний. Каждый раз один из них, выбранный жребием, открывает дверцу X, тогда как двое других открывают дверцу Y. Удачливые, как всегда, Алиса, Боб и Чарли делают удивительное открытие. Они обнаруживают, что *всякий раз*, когда они открывают ящики в таком порядке, они находят нечетное количество черных шаров.

То есть Алиса, Боб и Чарли обнаружили, что когда они открывают дверцу X на одном ящике и дверцы Y на двух других, то гарантированно наблюдают одно из сочетаний цветов:

$$0_A0_B1_C, \quad 0_A1_B0_C, \quad 1_A0_B0_C, \quad 1_A1_B1_C \tag{4.112}$$

(0 обозначает белый, 1 — черный). Они ни разу не наблюдали ни одного сочетания из

$$1_A1_B0_C, \quad 1_A0_B1_C, \quad 0_A1_B1_C, \quad 0_A0_B0_C. \tag{4.113}$$

независимо от того, у какого из трех ящиков была открыта дверца X.

Некоторое время спустя Алиса, Боб и Чарли понимают, что после открытия двух ящиков они всегда могут предсказать, что произойдет, когда

будет открыт третий ящик. Если первые два шара одного цвета, то третий шар, разумеется, будет черным, а если первые два — разных цветов, то последний шар обязательно будет белым. Они проверяли это несметное количество раз, но так происходило всегда!

Даже после признания эксперимента с тремя монетами Алиса, Боб и Чарли не усомнились в своей приверженности к эйнштейновской локальности. Однажды между ними состоялся трехсторонний разговор.

Алиса: Знаете, парни, иногда я просто не могу решиться открыть ли дверцу X или дверцу Y своего ящика. Я знаю, я должна выбирать аккуратно . . . . Если я открою дверцу X, то я несомненно внесу возмущение в ящик; следовательно, я никогда не узнаю, что случилось бы, если бы вместо этого я открыла дверцу Y. А если я открою дверцу Y, я никогда не узнаю, что нашла бы, если бы открыла дверцу X. Это так огорчает!

Боб: Алиса, ты не права! Наш эксперимент показывает, что ты можешь знать оба эти случая. Неужели ты не видишь? Допустим, что ты хочешь знать, что произойдет, когда ты откроешь дверцу X. Тогда ты просто просишь Чарли и меня открыть дверцы Y наших ящиков и сообщить тебе, что мы обнаружили. Ты будешь знать абсолютно точно, без сомнения, что случится, если ты откроешь дверцу X. Мы проверяли это много раз, и это всегда работает. Так зачем же беспокоиться и открывать дверцу X? Ты можешь пойти дальше и вместо этого открыть дверцу Y и узнать, что ты найдешь. Таким образом ты реально узнаешь результаты открывания *обеих* дверок!

Чарли: Но как можно быть в этом уверенным? Если Алиса открывает дверцу Y, она теряет возможность открыть дверцу X. Она же не может реально получить оба этих случая. После того, как она открывает дверцу Y, мы не можем проверить, произойдет ли ожидаемый результат при открывании дверцы X.

Боб: Да ну, как может случиться другое? Смотри, ты же на самом деле не думаешь, что ты со своим ящиком в Принстоне, а я со своим — в Чикаго можем оказать какое-то влияние на то, что найдет Алиса, когда она откроет свой ящик в Пасадене, не так ли? Когда мы открываем наши ящики, мы ничего не можем изменить в ящике Алисы; мы только узнаем информацию, необходимую для того, чтобы с уверенностью предсказать, что обнаружит Алиса.

Чарли: Ну, может быть, нам следовало бы выполнить несколько больше экспериментов, чтобы выяснить, что вы в этом правы.

Действительно, открытие корреляции трех ящиков сделало Алису и Боба даже более знаменитыми, чем раньше, но Чарли еще не получил той репутации, которой заслуживает, – он все еще без должности. Не удивительно, что он хочет выполнить больше экспериментов! Он продолжает:

Чарли: Здесь есть нечто такое, что мы можем проверить. Во всех выполненных до сих пор экспериментах мы всегда открывали дверцу Y на двух ящиках и дверцу X на одном. Может быть, нам нужно проверить нечто другое. Например, может быть, нам стоит посмотреть, что произойдет, если мы откроем одни и те же дверцы на всех трех ящиках. Мы могли бы проверить открытие трех X-дверок.

Боб: Да ну! Мне надоели эти три ящика. Мы уже все о них знаем. Пора двигаться дальше, и, я думаю, Диана уже готова нам помогать. Давайте перейдем к четырем ящикам!

Алиса: Нет, я считаю, что Чарли прав. Мы действительно не можем сказать, что мы все знаем о трех ящиках, пока не поэкспериментируем с другими способами открывания дверок.

Боб: Забудьте об этом! Нас ни за что не финансируют. После того как мы вложили столько усилий в открывание двух Y-ов и одной X, мы скажем, что теперь мы хотим открывать три X. Нам скажут, что сначала вы занимались ерундой, а теперь вы предлагаете заняться чепухой. Да нас просто поднимут на смех.

Алиса: Боб прав. Я думаю, что только одним способом мы сможем получить финансирование этого эксперимента, если мы сможем сделать предсказание относительно его исхода. Тогда мы сможем сказать, что выполняем эксперимент для проверки предсказания. Я слышала о неких теоретиках, Гринбергере, Горне, Цайлингере и Мермине (ГГЦМ). Они много размышляли о наших экспериментах с тремя ящиками; может, они смогут что-нибудь предложить.

Боб: Ну, в этих ящиках вся моя жизнь, а они просто банда теоретиков. Сомневаюсь, что они скажут что-нибудь интересное или полезное. Но на самом деле не важно, имеет ли их теория какой-нибудь смысл, я поддерживаю это предложение. Если мы сможем проверить ее, то я даже соглашусь с тем, что есть смысл выполнить новый эксперимент с тремя ящиками.

Итак, Алиса, Боб и Чарли совершают путешествие, чтобы познакомиться с ГГЦМ. И, несмотря на глубокий скептицизм Боба, ГГЦМ действительно делают очень интересное предложение.

ГГЦМ: Боб говорит, что, открывая ящики в Принстоне и в Чикаго, никак не возможно повлиять на то, что происходит, когда Алиса открывает ящик в Пасадене. Ну, допустим, что он прав. Теперь вы, парни, отправляетесь выполнять эксперимент, в котором вы все открываете X-дверцы. Никто не может сказать, что из этого получится, но мы можем рассуждать следующим образом: предположим, что если бы вы открыли три Y-дверцы, то обнаружили бы три белых шара. Тогда можно использовать аргументы Боба, чтобы понять, что если вместо этого вы откроете три X-дверцы, то найдете три черных шара. Это аналогично такому рассуждению: если Алиса открывает X, а Боб и Чарли открывают Y, тогда вы знаете наверняка, что количество черных шаров будет нечетным. Следовательно, если вы знаете, что Боб и Чарли, открыв дверцы Y, нашли по белому шару, то Алиса найдет черный, когда откроет дверцу X. Аналогично, если Алиса и Чарли, открыв дверцы Y, нашли по белому шару, то Боб найдет черный, когда откроет дверцу X. Наконец, если Алиса и Боб, открыв дверцы Y, нашли по белому шару, то Чарли должен найти черный, когда откроет дверцу X. Итак, мы видим, что[1]

$$Y_A Y_B Y_C = 000 \rightarrow X_A X_B X_C = 111. \tag{4.114}$$

Не так ли?

Боб: Ну, возможно, это достаточно логично, но насколько это полезно? Мы не знаем, что обнаружим внутри ящика, пока не откроем его. Вы предположили, что $Y_A Y_B Y_C = 000$, но мы никогда не знаем это заранее.

ГГЦМ: Безусловно, но подождите. Да, вы правы, что мы не можем знать заранее, что мы найдем, если откроем дверцу Y на каждом ящике. Но для трех ящиков имеется только восемь возможностей, и мы можем легко их все перечислить. А для каждой из этих восьми возможностей для $Y_A Y_B Y_C$ мы можем использовать те же рассуждения, что и раньше, чтобы сделать вывод о значении $X_A X_B X_C$. Мы получаем таблицу

[1] Здесь 0 обозначает белый шар, а 1 — черный; Y_A означает то, что находит Алиса, когда открывает дверцу Y на своем ящике, *и так далее*.

типа этой:

$$\begin{aligned}
Y_A Y_B Y_C = 000 &\rightarrow X_A X_B X_C = 111 \\
Y_A Y_B Y_C = 001 &\rightarrow X_A X_B X_C = 001 \\
Y_A Y_B Y_C = 010 &\rightarrow X_A X_B X_C = 010 \\
Y_A Y_B Y_C = 100 &\rightarrow X_A X_B X_C = 100 \\
Y_A Y_B Y_C = 011 &\rightarrow X_A X_B X_C = 100 \\
Y_A Y_B Y_C = 101 &\rightarrow X_A X_B X_C = 010 \\
Y_A Y_B Y_C = 110 &\rightarrow X_A X_B X_C = 001 \\
Y_A Y_B Y_C = 111 &\rightarrow X_A X_B X_C = 111
\end{aligned} \tag{4.115}$$

Боб: Хорошо, ну и что?

ГГЦМ: Есть нечто замечательное в этой таблице, Боб! Взгляни на значения $X_A X_B X_C$... Каждое из них имеет нечетное количество единиц. Это и есть наше предсказание. Когда вы все будете открывать дверцу X на ваших ящиках, вы всегда будете находить нечетное количество черных шаров! Может быть один, может быть три, но всегда *нечетное количество*.

Конечно же, Алиса, Боб и Чарли восхищены проницательностью ГГЦМ. Они предложили продолжить эксперимент, который был одобрен и щедро профинансирован. Наконец, наступает долгожданный день, когда они в первый раз выполняют эксперимент. И когда Алиса, Боб и Чарли, каждый, открывает дверцу X на своем ящике, можете ли вы угадать, что они обнаруживают? Три белых шара. Вах!!!

Подозревая ошибку, Алиса, Боб и Чарли очень тщательно повторяют эксперимент, снова и снова, и снова... Но в каждом испытании, всякий раз, они находят четное количество черных шаров, когда они открывают дверцу X на всех трех ящиках. Иногда — ни одного, иногда — два, но никогда — один, и никогда — три. То, что они обнаруживали, всегда было прямо противоположно тому, что предсказывали ГГЦМ, исходя из принципа эйнштейновской локальности!

Снова отчаяние приводит жаждущих просвещения Алису, Боба и Чарли в библиотеку. После некоторого изучения учебника по квантовой механике и основательного допроса сотрудника лаборатории Алисы они понимают, что их три ящика были приготовлены в квантовом ГГЦМ-состоянии:

$$|\psi\rangle_{ABC} = \frac{1}{\sqrt{2}}\left(|000\rangle_{ABC} + |111\rangle_{ABC}\right), \tag{4.116}$$

являющемся одновременно собственным состоянием трех наблюдаемых

$$\mathbf{Z}_A \otimes \mathbf{Z}_B \otimes \mathbf{1}_C, \quad \mathbf{1}_A \otimes \mathbf{Z}_B \otimes \mathbf{Z}_C, \quad \mathbf{X}_A \otimes \mathbf{X}_B \otimes \mathbf{X}_C \tag{4.117}$$

с единичным собственным значением. А так как $\mathbf{ZX} = i\mathbf{Y}$, они понимают, что это состояние обладает свойствами:[1]

$$\begin{aligned} \mathbf{Y}_A \otimes \mathbf{Y}_B \otimes \mathbf{X}_C &= -1, \\ \mathbf{X}_A \otimes \mathbf{Y}_B \otimes \mathbf{Y}_C &= -1, \\ \mathbf{Y}_A \otimes \mathbf{X}_B \otimes \mathbf{Y}_C &= -1, \\ \mathbf{X}_A \otimes \mathbf{X}_B \otimes \mathbf{X}_C &= 1. \end{aligned} \tag{4.118}$$

Открывая ящик с помощью дверцы X или дверцы Y, Алиса, Боб и Чарли выполняют измерение наблюдаемых $\mathbf{X}$ или $\mathbf{Y}$, результат которого $+1$ означает белый шар, а результат -1 — черный шар. Таким образом, если приготовлено трехкубитовое состояние (4.116), то уравнение (4.118) говорит, что нечетное количество черных шаров будет обнаруживаться, если дверца Y открывается на двух ящиках, а дверца X — на третьем, в то время как четное количество черных шаров будет обнаруживаться, если на всех трех ящиках открывается дверца X. Это поведение, недвусмысленно предсказываемое квантовой механикой, именно то, что казалось таким обескураживающим Алисе, Бобу и Чарли и их консервативным собратьям, приверженцам эйнштейновской локальности.

После дополнительного глубокого изучения учебника по квантовой механике Алиса, Боб и Чарли постепенно приходят к пониманию изъяна в их рассуждениях. Они знакомятся с принципом дополнительности Бора, принципом непримиримой несовместимости некоммутирующих наблюдаемых. Они поняли, что для того чтобы прийти к своему предсказанию, они должны были *постулировать* результат измерения $\mathbf{YYY}$, а затем делать заключение о результатах измерения $\mathbf{XXX}$. Пренебрегая требующими серьезного отношения предостережениями Нильса Бора, они пали жертвой самого пагубного заблуждения.

Как они и надеялись, эксперимент с тремя ящиками принес дальнейшее признание Алисе и Бобу, а также должность Чарли. Конечно, эксперимент с тремя монетами уже убедительно опроверг эйнштейновскую локальность; несмотря на это, эксперимент с тремя ящиками имел другой характер. В эксперименте с монетами Алиса и Боб могли открыть только

[1] Здесь используются обозначения $\mathbf{X}_A = \boldsymbol{\sigma}_1^{(A)}$, $\mathbf{Y}_B = \boldsymbol{\sigma}_2^{(B)}$ и так далее. Равенства в уравнении (4.118) следует понимать как символические. Они обозначают, что (4.116) является собственным состоянием соответствующих наблюдаемых с собственными значениями ∓ 1. — *Прим. ред.*

две из трех монет, обнаруживая любую из четырех возможных конфигураций: ОО, ОР, РО, РР. Лишь выполнив множество испытаний, они смогли накопить убедительные статистические доказательства нарушения неравенства Белла. В противоположность этому в эксперименте с тремя ящиками Алиса, Боб и Чарли нашли результат, не согласующийся с эйнштейновской локальностью в каждом отдельном испытании, в котором они открывали дверцу X на всех трех ящиках!

4.7. Упражнения

4.1. Теорема Харди. Боб (в Бостоне) и Клер (в Чикаго) делят множество идентично приготовленных копий двухкубитового состояния

$$|\psi\rangle = \sqrt{1-2x}\,|00\rangle + \sqrt{x}\,|01\rangle + \sqrt{x}\,|10\rangle, \tag{4.119}$$

где x — вещественное число, лежащее между 0 и $1/2$. Они проводят множество испытаний, в которых каждый измеряет свой кубит в базисе $\{|0\rangle, |1\rangle\}$, и узнают, что если результатом Боба является $|1\rangle$, то результат Клер всегда $|0\rangle$, а если результатом Клер является $|1\rangle$, тогда у Боба всегда $|0\rangle$.

Боб и Клер проводят дальнейшие эксперименты, в которых Боб выполняет измерение в базисе $\{|0\rangle, |1\rangle\}$, а Клер — в ортонормированном базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$. Они обнаруживают, что если результат Боба $|0\rangle$, то результатом Клер всегда является $|\varphi\rangle$ и никогда — $|\varphi^\perp\rangle$. Аналогично, если Клер измеряет в базисе $\{|0\rangle, |1\rangle\}$, а Боб — в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, тогда если результат Клер $|0\rangle$, то результатом Боба всегда является $|\varphi\rangle$ и никогда — $|\varphi^\perp\rangle$.

а) Выразить базис $\{|\varphi\rangle, |\varphi^\perp\rangle\}$ через $\{|0\rangle, |1\rangle\}$.

Теперь Боб и Клер интересуются, что произойдет, если они оба будут измерять в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$. Их друг Альберт, ярый сторонник локального реализма, предсказывает, что невозможно обоим получить результат $|\varphi^\perp\rangle$ (предсказание известное как *теорема Харди*). Альберт аргументирует это следующим образом.

> Когда Боб и Клер выполняют измерение в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, имеет смысл рассмотреть, что могло бы случиться, если бы вместо этого кто-то один из них выполнил измерение в базисе $\{|0\rangle, |1\rangle\}$.

Итак, предположим, что Боб и Клер оба измеряют в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$ и что они оба получают результат $|\varphi^\perp\rangle$. Теперь, если бы Боб вместо этого измерял в базисе $\{|0\rangle, |1\rangle\}$, то мы могли бы быть уверены в том, что его результат — $|1\rangle$, так как эксперимент показывает, что если бы Боб получил $|0\rangle$, то Клер не могла бы получить $|\varphi^\perp\rangle$. Аналогично, если бы Клер измеряла в базисе $\{|0\rangle, |1\rangle\}$, то она наверняка получила бы результат $|1\rangle$. Мы приходим к выводу, что если бы Боб и Клер оба измеряли в базисе $\{|0\rangle, |1\rangle\}$, то они оба получили бы результат $|1\rangle$. Но это противоречит эксперименту, который показывает, что если Боб и Клер оба выполняют измерение в базисе $\{|0\rangle, |1\rangle\}$, то невозможно им обоим получить результат $|1\rangle$. Следовательно, мы вынуждены сделать вывод, что если Боб и Клер измеряют в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, то они не могут одновременно получить результат $|\varphi^\perp\rangle$.

Несмотря на впечатляющую аргументацию Альберта, Боб и Клер решают исследовать, какое предсказание может быть получено из квантовой механики.

a) Если Боб и Клер оба измеряют в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, каково квантово-механическое предсказание для вероятности $P(x)$ того, что они оба получат результат $|\varphi^\perp\rangle$?

b) Найдите «максимальное нарушение» теоремы Харди: покажите, что максимальным значением $P(x)$ является $P[(3-\sqrt{5})/2] = (5\sqrt{5}-11)/2 \approx 0{,}0902$.

c) Боб и Клер проводят эксперимент, подтверждающий предсказание квантовой механики. Что ошибочно в рассуждениях Альберта?

4.2. Закрытие лазейки детектирования. Напомним, что *неравенство КГШХ*

$$|\langle ab\rangle + \langle a'b\rangle + \langle ab'\rangle - \langle a'b'\rangle| \leqslant 2 \tag{4.120}$$

справедливо, если случайные переменные a, b, a', b' принимают значения ± 1 и подчиняются совместному распределению вероятностей. Максимальное нарушение этого неравенства квантово-механическими предсказаниями имеет место, когда левая часть равна $2\sqrt{2}$, что достигается, когда Алиса и Боб делят максимально запутанное состояние $|\phi^+\rangle$, a, a' — результаты измерения кубита Алисы вдоль осей $\hat{x}$

и $\hat{z}$, а b, b' — результаты измерения кубита Боба вдоль осей $(\hat{x}+\hat{z})/\sqrt{2}$ и $(\hat{x}-\hat{z})/\sqrt{2}$.

Алиса и Боб выполнили замечательный эксперимент, измерив поляризацию запутанной фотонной пары и подтвердили предсказываемое квантовой механикой нарушение неравенства КГШХ. Альберт настроен скептически. Он обращает внимание на то, что используемые в их эксперименте детекторы не очень эффективны. По большей части, если Алиса регистрирует фотон, то Боб — нет, а если Боб регистрирует фотон, то Алиса – нет. Следовательно, они отбрасывают данные для большинства фотонных пар и оставляют результаты только при совпадении детектирования двух фотонов. В своем анализе данных Алиса и Боб предполагают, что их результаты основаны на репрезентативной выборке измеряемых наблюдаемых, подчиняющихся некоторому распределению вероятностей. Однако Альберт доказывает, что их выводы могут оказаться недостоверными, если состояние детектируемого фотона скоррелировано с результатом измерения поляризации.

Алиса и Боб интересуются, насколько им необходимо поднять эффективность детекторов, чтобы выполнить эксперимент, который убедит Альберта.

Алиса может ориентировать свой детектор вдоль любой оси, и если она направила его вдоль оси $\hat{a}$, то в идеале ее детектор будет щелкать, когда спин ее кубита направлен вверх вдоль оси $\hat{a}$, но ввиду неэффективности детектора иногда он не срабатывает, даже если кубит ориентирован вверх. Пусть теперь для каждого номера i фотонной пары: $x_i \in \{0,1\}$ - переменная, обозначающая сработал ли детектор Алисы, ориентированный вдоль оси $\hat{a}$, а именно, если щелчок был, то $x_i = 1$, а если нет, то $x_i = 0$. Поскольку детектор неидеальный, то x_i может быть равно нулю, даже если кубит ориентирован вверх вдоль $\hat{a}$. Аналогично $x'_i \in \{0,1\}$ обозначает, сработал ли детектор Алисы, ориентированный вдоль оси $\hat{a}'$, $y_i \in \{0,1\}$ обозначает, сработал ли детектор Боба, ориентированный вдоль $\hat{b}$, а $y'_i \in \{0,1\}$ обозначает, сработал ли детектор Боба, ориентированный вдоль $\hat{b}'$. В предположении локального реализма каждой паре можно сопоставить значения x, x', y, y', определяемые локальными скрытыми переменными.

Алиса и Боб свободны в выборе ориентации своих детекторов в каждом измерении; следовательно, их выборка значений x, x', y, y' репре-

зентативна и они выводят из своих измерений следующие значения:

$$\begin{aligned} P_{++}(ab) &= \frac{1}{N}\sum_{i=1}^{N} x_i y_i, \\ P_{-+}(a'b) &= \frac{1}{N}\sum_{i=1}^{N} x'_i y_i, \\ P_{+-}(ab') &= \frac{1}{N}\sum_{i=1}^{N} x_i y'_i, \\ P_{--}(a'b') &= \frac{1}{N}\sum_{i=1}^{N} x'_i y'_i, \end{aligned} \tag{4.121}$$

где N — полное количество испытанных пар. Здесь, например, $P_{++}(ab)$ — вероятность того, что оба детектора сработают, когда Алиса и Боб ориентируют их вдоль осей $\hat{a}$ и $\hat{b}$, соответственно (с учетом влияния несовершенства детекторов).

a) Покажите, что если $x, x', y, y' \in \{0, 1\}$, то

$$xy + xy' + x'y - x'y' \leqslant x + y. \tag{4.122}$$

b) Покажите, что

$$P_{++}(ab) + P_{-+}(a'b) + P_{+-}(ab') - P_{--}(a'b') \leqslant P_{+\cdot}(a) + P_{\cdot+}(b); \tag{4.123}$$

здесь $P_{+\cdot}(a)$ обозначает вероятность того, что детектор Алисы щелкнет, если он ориентирован вдоль оси $\hat{a}$, а $P_{\cdot+}(b)$ обозначает вероятность того, что детектор Боба щелкнет, если он ориентирован вдоль оси $\hat{b}$.

c) Теперь сравним это с предсказаниями квантовой механики, где детектор Алисы имеет эффективность η_A, а детектор Боба — η_B. Это означает, что детектор Алисы щелкает с вероятностью $P = \eta_A P_{\text{perf}}$, где P_{perf} — вероятность щелчка идеального детектора, и аналогично для детектора Боба. Выбирая a, b, a', b' максимально нарушающими неравенство КГШХ, покажите, что предсказания квантовой механики нарушают неравенство (4.123), если только

$$\frac{\eta_A \eta_B}{\eta_A + \eta_B} > \frac{1}{1+\sqrt{2}}. \tag{4.124}$$

Таким образом, если $\eta_A = \eta_B$, то Алисе и Бобу необходимы детекторы с эффективностью выше 82,84%, чтобы преодолеть возражения Альберта.

4.3. Телепортация с помощью непрерывных переменных. Один полный ортонормированный базис в гильбертовом пространстве двух частиц на вещественной прямой представляет собой базис (сепарабельных) собственных состояний оператора положения $\{|q_1\rangle \otimes |q_2\rangle\}$. Другой – запутанный базис $\{|Q, P\rangle\}$, где

$$|Q, P\rangle = \frac{1}{\sqrt{2\pi}} \int dq\, e^{iPq} |q\rangle \otimes |q + Q\rangle\,; \tag{4.125}$$

они являются одновременными собственными состояниями оператора относительного положения $\mathbf{Q} = \mathbf{q}_2 - \mathbf{q}_1$ и оператора полного импульса $\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2$.

a) Проверьте, что

$$\langle Q', P' | Q, P\rangle = \delta(Q' - Q)\delta(P' - P). \tag{4.126}$$

b) Поскольку состояния $\{|Q, P\rangle\}$ образуют базис, мы можем разложить собственные состояния положений как

$$|q_1\rangle \otimes |q_2\rangle = \int dQ\, dP\, |Q, P\rangle\langle Q, P|\big(|q_1\rangle \otimes |q_2\rangle\big). \tag{4.127}$$

Вычислите коэффициенты разложения $\langle Q, P|\big(|q_1\rangle \otimes |q_2\rangle\big)$.

c) Алиса и Боб приготовили запутанное состояние $|Q, P\rangle_{AB}$ двух частиц A и B; Алиса оставила себе частицу A, а Боб — частицу B. Алиса получила неизвестный волновой пакет $|\psi\rangle_C = \int dq\, |q\rangle_C\, {}_C\langle q|\psi\rangle_C$, который она намерена телепортировать Бобу. Составьте протокол, который они могут выполнить, чтобы осуществить телепортацию. Что должна измерить Алиса? Какую информацию она должна послать Бобу? Что должен сделать Боб, получив эту информацию, чтобы частица B была приготовлена в состоянии $|\psi\rangle_B$?

4.4. Телепортация со смешанными состояниями. Операциональный способ определения запутанного состояния заключается в том, что оно может быть использовано для телепортации неизвестного квантового состояния с лучшей точностью воспроизведения, чем этого можно было бы добиться с помощью одних только локальных операций и классической связи. В этом упражнении вы покажете, что существуют смешанные состояния, в этом смысле запутанные, но тем не менее не нарушающие никакого неравенства Белла. Следовательно, для смешанных

состояний (в противоположность чистым состояниям) понятия «запутанный» и «нарушающий неравенство Белла» не эквивалентны.

Рассмотрите «шумящую» запутанную пару с матрицей плотности

$$\rho(\lambda) = (1-\lambda)|\psi^-\rangle\langle\psi^-| + \frac{\lambda}{4}\mathbf{1}. \tag{4.128}$$

a) Найдите точность воспроизведения F, которой можно достичь, если состояние $\rho(\lambda)$ используется для телепортации одного кубита от Алисы к Бобу. [**Указание**. Вспомните, что вы показали в одном из предыдущих упражнений, что «случайное гадание» имеет точность воспроизведения $F = 1/2$.]

b) При каких значениях λ найденная в (a) точность воспоизведения лучше той, которой можно добиться, если Алиса измеряет свой кубит и посылает Бобу классическое сообщение? [**Указание.** Раньше вы показали, что можно достичь значения $F = 2/3$, если Алиса измеряет свой кубит. Фактически это наилучшее возможное значение F, достижимое в классической связи.]

c) Вычислите

$$\text{Prob}(\uparrow_{\hat{n}}\uparrow_{\hat{m}}) \equiv \text{tr}\left(\mathbf{E}_A(\hat{n})\mathbf{E}_B(\hat{m})\rho(\lambda)\right), \tag{4.129}$$

где $\mathbf{E}_A(\hat{n})$ — проекция кубита Алисы на состояние $|\uparrow_{\hat{n}}\rangle$, а $\mathbf{E}_B(\hat{m})$ — проекция кубита Боба на состояние $|\uparrow_{\hat{m}}\rangle$.

d) Рассмотрите случай $\lambda = 1/2$. Покажите, что в этом случае состояние $\rho(\lambda)$ не нарушает неравенства Белла. [**Указание.** Достаточно построить модель локальных скрытых переменных, которая при $\lambda = 1/2$ корректно воспроизводит найденные в (c) спиновые корреляции.] Предположите, что скрытая переменная $\hat{\alpha}$ однородно распределена на единичной сфере и что существуют функции f_A и f_B такие, что

$$\text{Prob}_A(\uparrow_{\hat{n}}) = f_A(\hat{\alpha}\cdot\hat{n}), \qquad \text{Prob}_B(\uparrow_{\hat{m}}) = f_B(\hat{\alpha}\cdot\hat{m}). \tag{4.130}$$

Задача состоит в том, чтобы найти f_A и f_B (где $0 \leqslant f_{A,B} \leqslant 1$), обладающие свойствами

$$\int_{\hat{\alpha}} f_A(\hat{\alpha}\cdot\hat{n}) = \frac{1}{2}, \qquad \int_{\hat{\alpha}} f_B(\hat{\alpha}\cdot\hat{m}) = \frac{1}{2},$$

$$\int_{\hat{\alpha}} f_A(\hat{\alpha}\cdot\hat{n}) f_B(\hat{\alpha}\cdot\hat{m}) = \text{Prob}(\uparrow_{\hat{n}}\uparrow_{\hat{m}}). \tag{4.131}$$

4.5. Распределение квантового ключа. Алиса и Боб хотят выполнить протокол распределения квантового ключа. Алиса имеет все необходимое, чтобы приготовить любое из двух состояний: $|u\rangle$ или $|v\rangle$. В подходящем базисе эти два состояния могут быть представлены как

$$|u\rangle = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad |v\rangle = \begin{pmatrix} \sin\alpha \\ \cos\alpha \end{pmatrix}, \tag{4.132}$$

где $0 < \alpha < \pi/4$. Алиса выбирает наугад, что послать Бобу, $|u\rangle$ или $|v\rangle$, а Боб должен выполнить измерение, чтобы определить, что она послала. Так как эти два состояния не ортогональны, Боб не может различить их с абсолютной точностью.

a) Боб понимает, что он не может рассчитывать на то, что всякий раз он сможет идентифицировать кубит Алисы, поэтому он довольствуется процедурой, которая лишь иногда обеспечивает успех. Он выполняет ПОЗМ с тремя возможными исходами: $\neg|u\rangle$, $\neg|v\rangle$, или НЕ ЗНАЮ. Если он получает результат $\neg|u\rangle$, он уверен, что было послано $|v\rangle$, а если он получает результат $\neg|v\rangle$, он уверен, что было послано $|u\rangle$. Если получен результат НЕ ЗНАЮ, тогда его измерение неубедительно (не позволяет сделать определенного вывода). Эта ПОЗМ определяется операторами

$$\begin{gathered} \mathbf{F}_{\neg u} = A(\mathbf{1} - |u\rangle\langle u|), \quad \mathbf{F}_{\neg v} = A(\mathbf{1} - |v\rangle\langle v|), \\ \mathbf{F}_{\mathrm{DK}} = (1 - 2A)\mathbf{1} + A(|u\rangle\langle u| + |v\rangle\langle v|) \end{gathered} \tag{4.133}$$

(DK — Don't Know — НЕ ЗНАЮ), где A — положительное вещественное число. Какое значение A должен выбрать Боб, чтобы минимизировать вероятность результата НЕ ЗНАЮ, и чему равна эта минимальная вероятность НЕ ЗНАЮ (при условии, что Алиса выбирает $|u\rangle$ или $|v\rangle$ с равной вероятностью)? **[Указание.** Если A слишком велико, то $\mathbf{F}_{\mathrm{DK}}$ будет иметь отрицательные собственные значения, а уравнения (4.133) не будут представлять ПОЗМ.]

b) Разработайте протокол распределения квантового ключа, используя исходные данные Алисы и ПОЗМ Боба.

c) Конечно, Ева тоже хочет знать, что Алиса посылает Бобу. Надеясь на то, что Алиса и Боб не заметят, она перехватывает каждый посылаемый Алисой кубит, выполняя ортогональное измерение, проецирующее его на базис $\left\{\binom{1}{0}, \binom{0}{1}\right\}$. Если она получает результат $\binom{1}{0}$, то она пересылает Бобу $|u\rangle$, а если — $\binom{0}{1}$, то пересылает ему $|v\rangle$. Следовательно, всякий раз, когда ПОЗМ Боба имеет

убедительный результат, Ева знает, каков он. Но вмешательство вызывает обнаруживаемые ошибки; иногда Боб получает «убедительный» результат, который на самом деле отличается от того, что посылала Алиса. Какова вероятность такой ошибки?

4.6. Минимальное возмущение. В упражнении 2.1 вы исследовали игру, в которой Алиса решает наудачу (равновероятно), какое чистое состояние одного кубита приготовить из двух возможных:

$$|\psi\rangle = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix} \quad \text{или} \quad |\tilde{\psi}\rangle = \begin{pmatrix}\sin\alpha\\ \cos\alpha\end{pmatrix}, \tag{4.134}$$

и посылает это состояние Бобу. Выполняя ортогональное измерение в базисе $\{|0\rangle, |1\rangle\}$, Боб может идентифицировать состояние с минимальной вероятностью ошибки

$$(P_{\text{error}})_{\text{optimal}} = \sin^2\alpha = \frac{1}{2}(1 - \sin\theta), \tag{4.135}$$

где мы определили θ соотношением

$$\langle\psi|\tilde{\psi}\rangle \equiv \cos\theta = \sin(2\alpha). \tag{4.136}$$

Но допустим теперь, что Ева хочет *перехватить* это состояние, пока оно движется от Алисы к Бобу. Как и Боб, она желает извлечь оптимальную информацию, отличающую $|\psi\rangle$ от $|\tilde{\psi}\rangle$, и при этом минимизировать вносимое ее вмешательством возмущение, так чтобы Алисе и Бобу было невозможно заметить, что здесь что-то не так.

Ева понимает, что оптимальную ПОЗМ можно осуществить с помощью операторов измерений

$$\mathbf{M}_0 = |\phi_0\rangle\langle 0|, \qquad \mathbf{M}_1 = |\phi_1\rangle\langle 1| \tag{4.137}$$

с произвольными векторами $|\phi_0\rangle$ и $|\phi_1\rangle$. Если Ева выполняет это измерение, то Боб получает состояние

$$\rho' = \cos^2\alpha|\phi_0\rangle\langle\phi_0| + \sin^2\alpha|\phi_1\rangle\langle\phi_1|, \tag{4.138}$$

если Алиса послала $|\psi\rangle$, и состояние

$$\tilde{\rho}' = \sin^2\alpha|\phi_0\rangle\langle\phi_0| + \cos^2\alpha|\phi_1\rangle\langle\phi_1|, \tag{4.139}$$

если Алиса послала $|\tilde{\psi}\rangle$.

Ева хочет, чтобы средняя точность воспроизведения получаемого Бобом состояния была как можно больше. Величина, которую она хочет минимизировать, называемая в дальнейшем «возмущением» D, измеряет, насколько близка к единице эта средняя точность воспроизведения

$$D = 1 - \frac{1}{2}(F + \tilde{F}), \tag{4.140}$$

где

$$F = \langle\psi|\rho'|\psi\rangle, \qquad \tilde{F} = \langle\tilde{\psi}|\tilde{\rho}'|\tilde{\psi}\rangle. \tag{4.141}$$

Целью этого упражнения является проверить, насколько эффективно Ева может сократить возмущение с помощью подходящего выбора своих измерительных операторов.

a) Покажите, что $F + \tilde{F}$ может быть представлено в виде

$$F + \tilde{F} = \langle\phi_0|A|\phi_0\rangle + \langle\phi_1|B|\phi_1\rangle, \tag{4.142}$$

где

$$\begin{aligned} A &= \begin{pmatrix} 1 - 2\cos^2\alpha\sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & 2\cos^2\alpha\sin^2\alpha \end{pmatrix}, \\ B &= \begin{pmatrix} 2\cos^2\alpha\sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & 1 - 2\cos^2\alpha\sin^2\alpha \end{pmatrix}. \end{aligned} \tag{4.143}$$

b) Покажите, что если $|\phi_0\rangle$ и $|\phi_1\rangle$ выбраны оптимально, то минимальное возмущение, которое может быть достигнуто, равно

$$D_{\min}(\cos^2\theta) = \frac{1}{2}\left(1 - \sqrt{1 - \cos^2\theta + \cos^4\theta}\right). \tag{4.144}$$

[**Указание.** Мы можем выбрать $|\phi_0\rangle$ и $|\phi_1\rangle$, чтобы независимо максимизировать два слагаемых в уравнении (4.142). Максимальным значением является максимальное собственное значение оператора A, которое может быть выражено как $\lambda_{\max} = \frac{1}{2}(1 + \sqrt{1 - 4\det A})$, поскольку сумма собственных значений равна единице.] Конечно, Ева могла бы сделать возмущение еще меньшим, если ее устроит меньшая, чем оптимальная, вероятность правильного определения сообщения Алисы.

c) Изобразите график функции $D_{\min}(\cos^2\theta)$. Истолкуйте ее значения при $\cos\theta = 1$ и $\cos\theta = 0$. При каком значении θ $D_{\min}$ максимальна? Найдите $D_{\min}$ и $(p_{\text{error}})_{\text{optimal}}$ для этого значения θ.

4.7. Приближенное клонирование. Теорема о невозможности клонирования показывает, что невозможно построить унитарную машину, которая будет делать идеальные копии неизвестного квантового состояния. Но допустим, что нас устроит *неидеальная* копия — какой точности воспроизведения мы можем добиться?

Рассмотрим машину, действующую на трехкубитовое состояние в соответствии с

$$\begin{aligned} |000\rangle_{ABC} &\to \sqrt{\frac{2}{3}}|00\rangle_{AB}|0\rangle_C + \sqrt{\frac{1}{3}}|\psi^+\rangle_{AB}|1\rangle_C, \\ |100\rangle_{ABC} &\to \sqrt{\frac{2}{3}}|11\rangle_{AB}|1\rangle_C + \sqrt{\frac{1}{3}}|\psi^-\rangle_{AB}|0\rangle_C. \end{aligned} \tag{4.145}$$

a) Являются ли такой прибор в принципе физически реализуемым?

Если машина действует на начальное состояние $|\psi\rangle_A|00\rangle_{BC}$, то она производит чистое запутанное состояние трех кубитов $|\Psi\rangle_{ABC}$. Но если мы наблюдаем один только кубит A, то его конечным состоянием является оператор плотности $\rho'_A = \mathrm{tr}_{BC}\left(|\Psi\rangle_{ABC\,ABC}\langle\Psi|\right)$. Аналогично конечным состоянием отдельно наблюдаемого кубита B является ρ'_B. Нетрудно видеть, что $\rho'_A = \rho'_B$ — идентичные, но не идеальные копии исходного чистого состояния $|\psi\rangle_A$.

b) Отображение начального состояния $|\psi\rangle_{A\,A}\langle\psi|$ кубита A на конечное состояние ρ'_A определяет супероператор \$. Найдите его представление операторной суммы.

c) Найдите ρ'_A для $|\psi\rangle_A = a|0\rangle_A + b|1\rangle_A$ и вычислите его точность воспроизведения $F \equiv {}_A\langle\psi|\rho'_A|\psi\rangle_A$.

4.8. Прости нас, дядюшка Альберт. Рассмотрим n-кубитовое «кот-состояние»

$$|\psi\rangle_n = \frac{1}{2}\left(|000\ldots0\rangle + |111\ldots1\rangle\right). \tag{4.146}$$

Это состояние можно охарактеризовать как одновременное собственное состояние (с единичным собственным значением) n операторов

$$\begin{aligned} &\sigma_3 \otimes \sigma_3 \otimes 1 \otimes 1 \otimes \cdots \otimes 1 \otimes 1 \otimes 1, \\ &1 \otimes \sigma_3 \otimes \sigma_3 \otimes 1 \otimes \cdots \otimes 1 \otimes 1 \otimes 1, \\ &\qquad\cdots \\ &1 \otimes 1 \otimes 1 \otimes 1 \otimes \cdots \otimes 1 \otimes \sigma_3 \otimes \sigma_3, \\ &\sigma_1 \otimes \sigma_1 \otimes \sigma_1 \otimes \cdots \otimes \sigma_1 \otimes \sigma_1 \otimes \sigma_1. \end{aligned} \tag{4.147}$$

a) Покажите, что $|\psi\rangle_n$ является собственным состоянием оператора

$$(\sigma_1 + i\sigma_2)^{\otimes n} + (\sigma_1 - i\sigma_2)^{\otimes n}, \tag{4.148}$$

и вычислите его собственное значение.

b) Если мы верим в локальные скрытые переменные, тогда мы верим, что для каждого из n-кубитов σ_1 и σ_2 имеют определенные значения, коль скоро скрытые переменные определены. Если это так, то что можно сказать относительно *модулей* $(\sigma_1 + i\sigma_2)^{\otimes n}$ или $(\sigma_1 - i\sigma_2)^{\otimes n}$, предполагая определенные значения скрытых переменных?

c) Из (b) выведите верхнюю границу для

$$\frac{1}{2}\left|(\sigma_1 + i\sigma_2)^{\otimes n} + (\sigma_1 - i\sigma_2)^{\otimes n}\right|, \tag{4.149}$$

следующую из гипотезы о локальных скрытых переменных.

d) Сравните это с (a). Что сказал бы Эйнштейн?

4.9. Манипулирование запутыванием. a) Двадцать пять игроков команды Янки из Нью-Йорка и двадцать пять игроков комады Святых Отцов из Сан-Диего хотят разделить пятьдесят кубитов «кот-состояния». Янки готовят 26-кубитовое «кот-состояние» и дают один из кубитов Алисе; то же делают и Святые Отцы. Теперь Алиса должна соединить эти и приготовить 50-кубитовое состояние. Как ей это сделать? [**Указание.** Подумайте о стабилизаторе.]

b) Присоединившись к Янки, Алиса приняла на хранение один из кубитов их 25-кубитового «кот-состояния». Но ее подкупили! Алисе поручено извлечь имеющийся у ней кубит из «кот-состояния», сохранив неповрежденным 24-кубитовое состояние остальных игроков. Как ей это сделать? [**Указание.** Подумайте о стабилизаторе.]

4.10. Критерий Переса — Городецки в d измерениях. Напомним, что *состояние Вернера* пары кубитов может быть представлено как

$$\rho(\lambda) = \lambda|\phi^+\rangle\langle\phi^+| + \frac{1}{4}(1-\lambda)\mathbf{1} \tag{4.150}$$

и что *частичное транспонирование* ρ_{AB}^{PT} парного оператора плотности ρ_{AB} определяется как

$$\rho_{AB}^{PT} \equiv (\mathbf{1}_A \otimes \mathbf{T}_B)\rho_{AB}, \tag{4.151}$$

где $\mathbf{T}$ – операция транспонирования, действующая в базисе $\{|i\rangle\}$ как

$$\mathbf{T}(|i\rangle\langle j|) = |j\rangle\langle i|. \tag{4.152}$$

На лекции мы видели, что частичное транспонирование состояния Вернера $\rho(\lambda)$ отрицательно при $\lambda > 1/3$; следовательно, согласно *критерию Переса–Городецки* состояние Вернера несепарабельно при $\lambda > 1/3$.

a) Естественный способ обобщения состояния Вернера на пару d-мерных систем – рассмотреть

$$\rho_\Phi(\lambda) = \lambda|\Phi\rangle\langle\Phi| + \frac{1}{d^2}(1-\lambda)\mathbf{1}, \tag{4.153}$$

где $|\Phi\rangle$ – максимально запутанное состояние

$$|\Phi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|i\rangle\otimes|i\rangle. \tag{4.154}$$

Покажите, что

$$(|\Phi\rangle\langle\Phi|)^{PT} = \frac{1}{d}(\mathbf{1} - 2\mathbf{E}_{\text{antisym}}), \tag{4.155}$$

где $\mathbf{E}_{\text{antisym}}$ – проектор на пространство, антисимметричное относительно перестановки двух систем: A и B.

b) При каких значениях λ частичное транспонирование $\rho_\Phi(\lambda)$ отрицательно?

c) Если состояние Вернера двух кубитов выбрано в виде

$$\rho(\lambda) = \lambda|\psi^-\rangle\langle\psi^-| + \frac{1}{4}(1-\lambda)\mathbf{1}, \tag{4.156}$$

тогда другой естественный способ обобщить состояние Вернера на пару d-мерных систем – рассмотреть

$$\rho_{\text{anti}}(\lambda) = \frac{2\lambda}{d(d-1)}\mathbf{E}_{\text{antisym}} + \frac{1}{d^2}(1-\lambda)\mathbf{1}. \tag{4.157}$$

При каких значениях λ частичное транспонирование $\rho_{\text{anti}}(\lambda)$ отрицательно?

Глава 5

Теория квантовой информации

Теория квантовой информации настолько обширный предмет, что вполне могла бы занимать нас весь семестр. Но вследствие недостатка времени (мне не терпится перейти к квантовым вычислениям) я не смогу осветить этот предмет так глубоко, как мне бы этого хотелось. Мы удовольствуемся отрывочным введением в некоторые основные идеи и результаты. Возможно, лекции будут носить более описательный характер, нежели в первом семестре, с более частыми рассуждениями на пальцах и с большим количеством деталей, оставленных для домашних упражнений. Вероятно, эту главу следовало бы назвать: «Теория квантовой информации для нетерпеливых» [1].

Теория квантовой информации имеет дело с четырьмя главными темами.

(1) Передача классической информации по квантовым каналам связи (будет обсуждаться).

(2) Компромисс между получением информации о квантовом состоянии и его возмущением (этот вопрос мы кратко обсуждали в главе 4 в связи с квантовой криптографией, но здесь разберемся с ним основательно).

(3) Количественная характеристика квантового запутывания (которой мы кратко коснемся).

(4) Передача квантовой информации по квантовым каналам связи. (Мы обсудим случай канала без помех, отложив обсуждение канала с помехами до тех пор, пока не познакомимся с квантовыми корректирующими кодами.)

Эти темы объединяет часто повторяющийся лейтмотив: интерпретация и применения энтропии фон Неймана.

---

[1] Читателю, интересующемуся более строгим математическим изложением основных результатов квантовой теории информации, можно порекомендовать книгу А.С. Холево, *Введение в квантовую теорию информации*, МЦНМО, М.: 2002; более подробное изложение теории классической и квантовой информации на физическом уровне строгости можно найти в книге M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2001; перевод на русский язык М. Нильсен, И. Чанг, *Квантовые вычисления и квантовая информация*, М.: Мир, 2006. – *Прим. ред.*

5.1. Шеннон для «чайников»

Прежде чем мы сможем понять энтропию фон Неймана и ее значение для квантовой информации, мы должны обсудить энтропию Шеннона и ее значение для информации классической.

В своей основополагающей статье 1948 г. Клод Шеннон установил два основных результата теории классической информации. Им были решены две центральных проблемы.

(1) Насколько можно *сжать* сообщение, то есть насколько избыточна информация? («Теорема кодирования без помех»).

(2) С какой *скоростью* мы можем надежно передавать сообщения по каналу с помехами; то есть насколько избыточным должно быть содержание сообщения, чтобы быть защищенным от ошибок? («Теорема кодирования для канала с помехами (шумом)».)

Оба вопроса касаются *избыточности* — насколько, в среднем, *неожиданна* следующая буква сообщения. Согласно одной из ключевых идей Шеннона, удобную количественную меру избыточности предоставляет *энтропия*.

Я назвал этот раздел «Шеннон для чайников», поскольку я попытаюсь быстро объяснить основные идеи Шеннона с минимальным количеством ε-ов и δ. Таким образом, я смогу втиснуть теорию классической информации примерно в двенадцать страниц.

5.1.1. Энтропия Шеннона и сжатие данных

Сообщением называется строка из букв, выбранных из содержащего k букв алфавита

$$\{a_1, a_2, \ldots, a_k\}. \tag{5.1}$$

Предположим, что буквы a_x в сообщении статистически независимы и каждая из них появляется с заданной *a priori* вероятностью $p(a_x)$, где $\sum_{x=1}^{k} p(a_x) = 1$. Простейшим примером служит двоичный алфавит, в котором 0 появляется с вероятностью $1 - p$, а 1 — с вероятностью p (где $0 \leqslant p \leqslant 1$).

Рассмотрим длинные, содержащие n букв ($n \gg 1$), сообщения. Нас интересует, можно ли сжать сообщение до более короткой строки, несущей, по существу, ту же информацию?

Согласно закону больших чисел при очень больших n типичные строки содержат (в двоичном случае) примерно $n(1-p)$ нулей и примерно np единиц. Количество различных строк такого типа (типичных строк) по порядку величины равно биномиальному коэффициенту $\binom{n}{np}$ и из формулы Стирлинга $\log n! = n\log n - n + O(\log n)$ мы получаем

$$\begin{aligned}\log\binom{n}{np} &= \log\left(\frac{n!}{(np)![n(1-p)]!}\right) \simeq n\log n - n - \\ &- [np\log np - np + n(1-p)\log n(1-p) - n(1-p)] = \\ &= nH(p), \end{aligned} \tag{5.2}$$

где

$$H(p) = -p\log p - (1-p)\log(1-p) \tag{5.3}$$

— функция, называемая *энтропией*. Следовательно, количество типичных строк имеет порядок $2^{nH(p)}$. (Логарифмы здесь понимаются по основанию два, если не оговаривается иное.)

Чтобы передать, по существу, всю информацию, переносимую строкой из n битов, достаточно выбрать блоковый код, присваивающий положительное целое число каждой типичной строке. Этот блоковый код имеет около $2^{nH(p)}$ слов (появляющихся с *a priori* одинаковой вероятностью), так что любое из них мы можем идентифицировать, используя двоичную строку длиной $nH(p)$. Поскольку $0 \leqslant H(p) \leqslant 1$ при $0 \leqslant p \leqslant 1$ и $H(p) = 1$ только при $p = 1/2$, блоковый код сокращает сообщение при любом $p \neq 1/2$ (когда 0 и 1 не равновероятны). Это результат Шеннона. Главная идея заключается в том, что нам не нужно кодовое слово для каждой последовательности букв, а только для *типичных* последовательностей. Вероятность того, что действительное сообщение атипично, асимптотически (то есть в пределе $n \to \infty$) мала.

Это рассуждение очевидным образом обобщается на случай k букв, когда буква x появляется с вероятностью $p(x)$.[1] В строке, содержащей n букв, x обычно возникает приблизительно $np(x)$ раз, а количество типичных строк имеет порядок

$$\frac{n!}{\prod_x (np(x))!} \simeq 2^{nH(X)}, \tag{5.4}$$

[1] Ансамбль, в котором каждая из n букв извлекается из распределения X, будет обозначаться как X^n.

где мы вновь воспользовались асимптотической формулой Стирлинга, а

$$H(X) = -\sum_x p(x) \log p(x) \tag{5.5}$$

– *энтропия Шеннона* (или просто энтропия) ансамбля $X = \{x, p(x)\}$. Выбирая блоковый код, присваивающий целые числа типичным последовательностям, можно сжать до $nH(X)$ битов информацию, содержащуюся в строке из n букв. В этом смысле выбранная из ансамбля буква x несет в среднем $H(X)$ битов информации.

Это рассуждение полезно переформулировать на несколько ином языке. Отдельное n-буквенное сообщение

$$x_1 x_2 \ldots x_n \tag{5.6}$$

возникает с вероятностью, *a priori* равной

$$P(x_1 x_2 \ldots x_n) = p(x_1)p(x_2)\cdots p(x_n), \tag{5.7}$$

$$\log P(x_1 x_2 \ldots x_n) = \sum_{i=1}^{n} \log p(x_i). \tag{5.8}$$

Применяя к этой сумме центральную предельную теорему, мы приходим к выводу, что для «большинства последовательностей»

$$-\frac{1}{n} \log P(x_1 x_2 \ldots x_n) \sim \langle -\log p(x) \rangle \equiv H(X), \tag{5.9}$$

где угловые скобки обозначают среднее значение по распределению вероятностей, управляющему случайной переменной x.

Конечно, на языке ε-ов и δ можно дать точную формулировку этого утверждения. Для любых $\varepsilon, \delta > 0$ и для достаточно больших n каждая «типичная последовательность» имеет вероятность P, удовлетворяющую неравенству

$$H(X) - \delta < -\frac{1}{n} \log P(x_1 x_2 \ldots x_n) < H(X) + \delta, \tag{5.10}$$

а суммарная вероятность всех типичных последовательностей превышает $1 - \varepsilon$.[1] Или, другими словами, каждая из последовательностей букв, возникающих с превосходящей $1 - \varepsilon$ суммарной вероятностью («типичные

[1] Фактически это один из вариантов *закона больших чисел*. Его строгую математическую формулировку можно найти в любом учебнике по теории вероятностей или в книге А. С. Холево, *Введение в квантовую теорию информации*, МЦНМО, М.: 2002. – *Прим. ред.*

последовательности»), появляется с вероятностью P такой, что

$$2^{-n(H+\delta)} \leqslant P \leqslant 2^{-n(H-\delta)}. \tag{5.11}$$

Из уравнения (5.11) можно вывести верхнюю и нижнюю грани для *количества* $N(\varepsilon, \delta)$ типичных последовательностей (так как сумма вероятностей всех типичных последовательностей должна лежать между $1-\varepsilon$ и единицей):

$$(1-\varepsilon)2^{n(H-\delta)} \leqslant N(\varepsilon, \delta) \leqslant 2^{n(H+\delta)}. \tag{5.12}$$

С помощью блокового кода длиной $n(H+\delta)$ битов мы можем закодировать все типичные последовательности. Тогда, независимо от того, как закодированы атипичные последовательности, вероятность ошибки (декодирования) будет меньше, чем ε.

И наоборот, если мы попытаемся сжать сообщение до меньшего, чем $H-\delta'$, количества битов на одну букву, то не сможем добиться малой частоты ошибок при $n \to \infty$, так как будем не в состоянии однозначно присвоить кодовые слова всем типичным последовательностям. Вероятность успешного декодирования сообщения P_{success} будет ограничена сверху

$$P_{\text{success}} \leqslant 2^{n(H-\delta')}2^{-n(H-\delta)} + \varepsilon' = 2^{-n(\delta'-\delta)} + \varepsilon'. \tag{5.13}$$

Мы можем корректно декодировать только $2^{n(H-\delta')}$ типичных сообщений, каждое из которых возникает с вероятностью, меньшей чем $2^{-n(H-\delta)}$ (ε' добавлена, чтобы учесть вероятность того, что нам удается корректно декодировать атипичные сообщения). А так как δ может быть выбрана сколь угодно малой, то при $n \to \infty$ малой становится и эта вероятность успеха.

Таким образом, оптимальный код асимптотически сжимает каждую букву до $H(X)$ битов. Это и есть теорема Шеннона о кодировании в отсутствии шума.

5.1.2. Взаимная информация

Энтропия Шеннона $H(X)$ количественно определяет, сколько в среднем информации передается буквой, извлеченной из ансамбля X. То есть сообщает, сколько (асимптотически при $n \to \infty$, где n — количество извлеченных букв) необходимо битов, чтобы закодировать эту информацию.

Взаимная информация $I(X;Y)$ количественно определяет степень корреляции двух сообщений. Как много мы узнаем о сообщении, извлеченном из X^n, прочитав сообщение, извлеченное из Y^n?

Допустим, например, что мы хотим послать сообщение от отправителя к получателю. Однако в канале связи имеется шум, так что полученное

сообщение (y) может отличаться от посланного (x). Канал с шумом можно характеризовать условной вероятностью $p(y|x)$ — вероятностью того, что будет получено y, если послано x. Предположим, что буква x посылается с *a priori* известной вероятностью $p(x)$. Мы хотим количественно определить, что мы узнаем об x, получив y; какой объем информации мы приобретаем?

Как уже говорилось, энтропия $H(X)$ дает отнесенную к одной букве количественную меру моего априорного незнания сообщения до его получения; то есть вам необходимо передать мне (без искажений) nH битов, чтобы (асимптотически) точно определить конкретное сообщение из n букв. Но после ознакомления с сообщением y, я могу использовать теорему Бейеса, чтобы скорректировать распределение вероятностей для x:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}. \tag{5.14}$$

[Мне известны $p(y|x)$, если я знаком со свойствами канала, и $p(x)$, если я знаю априорные вероятности появления букв; таким образом, я могу вычислить $p(y) = \sum_x p(y|x)p(x)$.] Благодаря приобретенному новому знанию я стал более осведомлен относительно x, чем ранее. С полученными мной y-ми, используя оптимальный код, вы можете полностью определить конкретную строку из n букв, посылая мне

$$H(X|Y) = \langle -\log p(x|y) \rangle \tag{5.15}$$

битов на каждую букву[1]. $H(X|Y)$ называется «условной энтропией». Из $p(x|y) = p(x,y)/p(y)$ мы видим, что

$$\begin{aligned} H(X|Y) &= \langle -\log p(x,y) + \log p(y) \rangle = \\ &= H(X,Y) - H(Y) \end{aligned} \tag{5.16}$$

[1]Вряд ли следует понимать это утверждение в буквальном смысле. Для того, чтобы отправитель мог восстановить посылаемую им строку из n букв X, он должен получать по параллельному каналу *без шума* информацию о каждой букве выходящего сообщения Y и по нему же отправлять исправления, если произошла ошибка передачи. [См. C. Shannon, A mathematical theory of communication, Bell System Techn. J., **27**, № 3, 379–423; № 4, 623–656 (1948). Русский перевод: К. Шеннон, Математическая теория связи, в книге К. Шеннон, *Работы по теории информации и кибернетике*, ИЛ, Москва (1063), стр. 243–332.] Скорее условную энтропию $H(X|Y)$ следует интерпретировать как количество информации, *теряемой* при передаче сообщения X через канал с шумом.

Кроме этого, обратим внимание на то, что в (5.15), а также в уравнениях (5.16), (5.17) и (5.19), угловые скобки обозначают усреднение по совместному распределению вероятностей $p(x,y)$. Следовательно все эти величины определяют информационные характеристики не конкретных n-буквенных строк, а всего совместного ансамбля входящих и выходящих сообщений. — *Прим. ред.*

и аналогично

$$H(Y|X) \equiv \langle -\log p(y|x) \rangle = \\ = \left\langle -\log \frac{p(x,y)}{p(x)} \right\rangle = H(X,Y) - H(X). \tag{5.17}$$

Таким образом, $H(X|Y)$ можно интерпретировать как количество *дополнительных* битов на одну букву, необходимых для полного определения x и y при известном y. Очевидно, что эта величина не может быть отрицательной.

Информация об X, приобретаемая при знакомстве с Y, измеряется тем, насколько *сокращается* отнесенное к одной букве количество битов, необходимое для идентификации X при известном Y. Таким образом:

$$I(X;Y) \equiv H(X) - H(X|Y) = \\ = H(X) + H(Y) - H(X,Y) = \\ = H(Y) - H(Y|X). \tag{5.18}$$

$I(X;Y)$ называется взаимной информацией. Она, очевидно, симметрична относительно перестановки X и Y; количество информации об X, получаемое при знакомстве с Y, равно количеству информации об Y, получаемому при знакомстве с X. Знакомство с Y не может *уменьшить* мое знание об X, следовательно, $I(X;Y)$ очевидно неотрицательна. (Неравенства $H(X) \geqslant H(X|Y) \geqslant 0$ легко доказываются с учетом свойства выпуклости логарифмической функции[1].)

Конечно, если X и Y полностью некоррелированы, то мы имеем $p(x,y) = p(x)p(y)$ и

$$I(X;Y) \equiv \left\langle \log \frac{p(x,y)}{p(x)p(y)} \right\rangle = 0; \tag{5.19}$$

естественно, что, знакомясь с Y, мы ничего не можем узнать об X, если между ними нет корреляции!

5.1.3. Теорема о кодировании для канала с шумом

Если мы хотим установить связь через канал с шумом, то мы, очевидно, можем повысить надежность передачи посредством избыточности

[1] См., например, T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, J. Wiley & Sons, New York, 1991.

информации. Например, я могу многократно посылать каждый бит, а получатель – прислушиваться к голосу большинства, чтобы его декодировать.

Но всегда ли для данного канала можно найти код, гарантирующий сколь угодно высокую надежность (при $n \to \infty$)? А что можно сказать о *быстродействии* таких кодов; сколько битов потребуется для каждой буквы сообщения?

Фактически Шеннон показал, что любой канал может быть использован для сколь угодно надежной связи с конечной (ненулевой) скоростью, пока существует хоть *какая-нибудь* корреляция между его входом и выходом. Более того, он нашел полезное выражение для оптимальной скорости коммуникации, которая может быть достигнута. Эти результаты составляют содержание «теоремы о кодировании для канала связи с шумом».

Предположим для определенности, что мы пользуемся двоичным алфавитом, каждая буква которого (0 и 1) появляется с априорной вероятностью $1/2$. Предположим также, что канал является «двоичным симметричным каналом» – он действует на каждый бит независимо, с вероятностью p инвертируя его значение и оставляя невредимым с вероятностью $1 - p$. То есть условные вероятности равны

$$\begin{aligned} p(0|0) &= 1 - p, \quad p(0|1) = p, \\ p(1|0) &= p, \quad p(1|1) = 1 - p. \end{aligned} \tag{5.20}$$

Мы хотим построить семейство кодов растущего блочного размера n такого, чтобы вероятность ошибки декодирования стремилась к нулю при $n \to \infty$. Если количество закодированных в блоке битов равно k, то код заключается в выборе 2^k «кодовых слов» из 2^n возможных n-битовых строк. Определим быстродействие кода R (число битов информации, приходящихся на один передаваемый бит) как

$$R = \frac{k}{n}. \tag{5.21}$$

Нам нужно разработать такой код, чтобы кодовые строки находились как можно «дальше друг от друга». Другими словами, для данного быстродействия R мы хотим максимизировать количество битов, которые должны инвертироваться, чтобы одно кодовое слово заменилось другим (это количество называется «расстоянием Хэмминга» между двумя кодовыми словами).

Для любой входящей строки длиной n битов, ошибки, как правило, будут вызывать инвертирование примерно np битов – следовательно, вход

обычно рассеивается в одну из примерно $2^{nH(p)}$ типичных выходящих строк (заполняющих «сферу радиуса Хэмминга» np, окружающую входящую строку). Для надежного декодирования, входящие кодовые слова следует выбирать таким образом, чтобы было маловероятным перекрытие сфер ошибок двух разных кодовых слов. В противном случае два разных входа иногда будут давать один и тот же выход, что с неизбежностью приведет к ошибкам декодирования. Если мы хотим избавиться от таких двусмысленностей декодирования, полное число строк, содержащихся во всех 2^{nR} сферах ошибок, не должно превышать полного количества битов 2^n в выходящем сообщении; мы требуем выполнения

$$2^{nH(p)}2^{nR} \leqslant 2^n \tag{5.22}$$

или

$$R \leqslant 1 - H(p) \equiv C(p). \tag{5.23}$$

Если надежность передачи достаточно высока, мы не можем ожидать, что быстродействие кода превзойдет $C(p)$. Но достижимо ли на самом деле быстродействие $R = C(p)$ (асимптотически)?

Фактически возможна передача с R, сколь угодно близким к $C(p)$ и сколь угодно малой вероятностью ошибки. По-видимому, самой остроумной из идей Шеннона была демонстрация того, что $C(p)$ может быть достигнуто учетом среднего по «случайным кодам». [Очевидно, что случайный выбор кода — не самый разумный способ, но, возможно это покажется удивительным, оказывается, что случайное кодирование достигает такого же высокого быстродействия (асимптотически при больших n), как и любая другая схема кодирования.] Поскольку C представляет собой оптимальное быстродействие при надежной передаче данных по каналу с шумом, она называется *емкостью канала связи* или *пропускной способностью канала связи*.

Предположим, что 2^{nR} кодовых слов представляют собой случайную выборку из ансамбля X^n. Сообщение (одно из кодовых слов) послано. Чтобы его декодировать, изобразим вокруг полученного сообщения «сферу Хэмминга», содержащую

$$2^{n[H(p)+\delta]} \tag{5.24}$$

строк. Сообщение декодируется содержащимся в этой сфере кодовым словом в предположении, что оно существует и единственно. Если такое кодовое слово не существует или оно не единственно, то будем считать, что произошла ошибка декодирования.

Насколько вероятна ошибка декодирования? Мы выбрали сферу декодирования достаточно большой, так что отсутствие достоверного кодового

слова внутри сферы атипично, следовательно, мы должны беспокоиться лишь о том, что ее займут более одного достоверного кодового слова. Поскольку всего имеется 2^n возможных строк, то окружающая выходящую строку сфера Хэмминга содержит долю

$$\frac{2^{n[H(p)+\delta]}}{2^n} = 2^{-n[C(p)-\delta]} \tag{5.25}$$

от общего количества строк. Таким образом, вероятность того, что одно из 2^{nR} случайно выбранных кодовых слов «по несчастью» займет эту сферу, равна

$$2^{-n[C(p)-R-\delta]}. \tag{5.26}$$

А так как δ мы можем выбрать сколь угодно малым, то R можно взять настолько близким к $C(p)$ [но все же меньшим, чем $C(p)$], насколько это необходимо, чтобы вероятность такой ошибки оставалась экспоненциально малой при $n \to \infty$.

Пока мы показали, что мала *средняя* вероятность ошибки, которую мы усредняем по выбору случайного кода, а для каждого конкретного кода — еще и по всем кодовым словам. Таким образом, должен существовать один частный код со средней (усредненной по кодовым словам) вероятностью ошибки, меньшей чем ε. Но нам хотелось бы иметь более сильный результат — вероятность ошибки мала для *каждого* кодового слова.

Чтобы установить этот более сильный результат, обозначим через P_i вероятность ошибки декодирования i-го посланного кодового слова. Мы продемонстрировали существование кода такого, что

$$\frac{1}{2^{nR}} \sum_{i=1}^{2^{nR}} P_i < \varepsilon. \tag{5.27}$$

Пусть $N_{2\varepsilon}$ обозначает количество кодовых слов с $P_i > 2\varepsilon$. Тогда мы приходим к выводу, что

$$\frac{1}{2^{nR}}(N_{2\varepsilon})2\varepsilon < \varepsilon \quad \text{или} \quad N_{2\varepsilon} < 2^{nR-1}, \tag{5.28}$$

то есть можно отбросить максимум половину кодовых слов, чтобы добиться $P_i < 2\varepsilon$ для *каждого* кодового слова. Быстродействие сконструированного нами нового кода равно

$$\text{Rate} = R - \frac{1}{n}, \tag{5.29}$$

что стремится к R при $n \to \infty$.

Таким образом, $C(p) = 1 - H(p)$ представляет собой максимальное быстродействие, которое может быть достигнуто асимптотически со сколь угодно малой вероятностью ошибки.

Рассмотрим теперь, как обобщить эти доказательства на более общие алфавиты и каналы. Пусть имеется канал связи, характеризуемый набором $p(y|x)$, и определенное распределение вероятностей $X = \{x, p(x)\}$ для входящих букв. Мы посылаем строки из n букв и предполагаем, что канал действует на каждую букву независимо. (О действующем таким образом канале говорят как о «канале без памяти».) Конечно, как только заданы $p(y|x)$ и $X = \{x, p(x)\}$, так сразу определены $p(x|y)$ и $Y = \{y, p(y)\}$.

Чтобы установить достижимое быстродействие, вновь рассмотрим усреднение по случайным кодам, где кодовые слова выбираются с *априорной* вероятностью, определяемой ансамблем X^n. Таким образом, с высокой вероятностью они будут выбраны из типичного набора строк букв, содержащего около $2^{nH(X)}$ таких типичных строк.

Для типичного, принадлежащего Y^n, получаемого сообщения существует около $2^{nH(X|Y)}$ сообщений, которые могли бы быть посланы. Мы можем декодировать полученное сообщение, сопоставляя ему «сферу», содержащую $2^{n[H(X|Y)+\delta]}$ возможных входов. Если внутри этой сферы имеется единственное кодовое слово, то им декодируется полученное сообщение.

Как и раньше, маловероятно, что внутри сферы не окажется ни одного кодового слова, однако мы должны исключить возможность того, что их там больше одного. Каждая сфера декодирования содержит долю

$$\frac{2^{n[H(X|Y)+\delta]}}{2^{nH(X)}} \cdot 2^{-n[H(X)-H(X|Y)-\delta]} = \\ = 2^{-n[I(X;Y)-\delta]} \tag{5.30}$$

от общего числа типичных входов. Если имеется 2^{nR} кодовых слов, то вероятность того, что одно из них случайно окажется внутри сферы декодирования, равна

$$2^{nR}2^{-n[I(X;Y)-\delta]} = 2^{-n[I(X;Y)-R-\delta]}. \tag{5.31}$$

Поскольку δ может быть выбрана сколь угодно малой, то R можно взять настолько близким к I (но все же меньшим, чем I), насколько это необходимо, чтобы вероятность ошибки декодирования оставалась экспоненциально малой при $n \to \infty$.

Это доказательство показывает, что, когда мы усредняем по случайным кодам и по кодовым словам, вероятность ошибки остается малой при любом быстродействии $R < I$. Тогда те же самые рассуждения, что и выше, показывают существование особого кода с вероятностью ошибки $< \varepsilon$ для каждого кодового слова. Это приемлемый результат, поскольку он согласуется с нашей интерпретацией I, как информации, которую мы приобретаем о входящем X, получая сигнал Y. То есть I представляет собой отнесенную к одной букве информацию, которую мы можем послать по данному каналу связи.

Взаимная информация $I(X;Y)$ зависит не только от условных вероятностей $p(y|x)$, характеризующих канал связи, но также и от априорных вероятностей $p(x)$ появления букв. Приведенное выше доказательство случайного кодирования применимо при любом выборе вероятностей $p(x)$, следовательно, мы показали, что безошибочная передача возможна при любом быстродействии R, меньшем чем

$$C = \max_{\{p(x)\}} I(X;Y). \tag{5.32}$$

C называется *емкостью канала* или *пропускной способностью канала* и зависит только от условных вероятностей, определяющих данный канал.

Мы показали, что достижимо любое быстродействие $R < C$, но может ли R превзойти C (при условии, что по-прежнему вероятность ошибки стремится к нулю при $n \to \infty$)? Доказательство того, что C является верхней границей быстродействия, в общем случае может показаться более тонким, чем для двоичного симметричного канала — вероятности ошибок для разных букв различны, и мы свободны в использовании этого при создании кода. Будем, однако, рассуждать следующим образом:

Допустим, что мы выбрали 2^{nR} строк из n букв в качестве кодовых слов. Рассмотрим ансамбль (обозначаемый как $\tilde{X}^n$), в котором каждое кодовое слово возникает с одинаковой вероятностью ($= 2^{-nR}$). Тогда очевидно, что

$$H(\tilde{X}^n) = nR. \tag{5.33}$$

Посылая кодовые слова через канал связи, мы получаем ансамбль $\tilde{Y}^n$ выходящих состояний.

Поскольку мы предполагаем, что канал действует на каждую букву независимо, условная вероятность для строки из n букв факторизуется:

$$p(y_1 y_2 \ldots y_n | x_1 x_2 \ldots x_n) = p(x_1|y_1)p(x_2|y_2)\cdots p(x_n|y_n), \tag{5.34}$$

а отсюда следует, что условная энтропия удовлетворяет условию

$$H(\tilde{Y}^n|\tilde{X}^n) = \langle -\log p(y^n|x^n)\rangle = \sum_i \langle -\log p(y_i|x_i)\rangle = \\ = \sum_i H(\tilde{Y}_i|\tilde{X}_i), \tag{5.35}$$

где $\tilde{X}_i$ и $\tilde{Y}_i$ – частные (маргинальные) распределения вероятностей для i-ой буквы, определяемые нашим распределением по кодовым словам. Напомним, что нам также известно, что $H(X,Y) \leqslant H(X) + H(Y)$ или

$$H(\tilde{Y}^n) \leqslant \sum_i H(\tilde{Y}_i). \tag{5.36}$$

Отсюда следует, что

$$I(\tilde{Y}^n;\tilde{X}^n) = H(\tilde{Y}^n) - H(\tilde{Y}^n|\tilde{X}^n) \leqslant \sum_i \left[H(\tilde{Y}_i) - H(\tilde{Y}_i|\tilde{X}_i)\right] = \\ = \sum_i I(\tilde{Y}_i;\tilde{X}_i) \leqslant nC; \tag{5.37}$$

взаимная информация посланного и полученного сообщений ограничена сверху суммой отнесенных к каждой букве взаимных информаций, а взаимная информация для каждой буквы ограничена сверху емкостью канала связи [поскольку C определяется как максимум $I(X;Y)$].

Вспоминая о симметрии взаимной информации, мы имеем

$$I(\tilde{X}^n;\tilde{Y}^n) = H(\tilde{X}^n) - H(\tilde{X}^n|\tilde{Y}^n) = \\ = nR - H(\tilde{X}^n|\tilde{Y}^n) \leqslant nC. \tag{5.38}$$

Теперь если мы в состоянии надежно декодировать при $n \to \infty$, то это означает, что входящее кодовое слово полностью определяется получаемым сигналом или что условная энтропия входа (в расчете на одну букву) должна стать малой:

$$\frac{1}{n}H(\tilde{X}^n|\tilde{Y}^n) \to 0. \tag{5.39}$$

Если безошибочность передачи возможна, то в пределе $n \to \infty$ уравнение (5.38) принимает вид

$$R \leqslant C. \tag{5.40}$$

Быстродействие не может превзойти емкость канала связи. [Вспомним, что условная энтропия, в отличие от взаимной информации, *не симметрична*. Действительно, $H(\tilde{Y}^n|\tilde{X}^n)/n$ *не становится* малым, поскольку канал вносит неопределенность в то, какое сообщение будет получено. Но если мы можем декодировать точно, то, коль скоро сигнал получен, исчезает неопределенность в том, какое кодовое слово было послано.]

Мы показали, что емкость C представляет собой максимальное достижимое быстродействие связи через канал с шумом, при котором вероятность ошибки декодирования стремится к нулю при стремящемся к бесконечности количестве букв в сообщении. В этом состоит теорема Шеннона о кодировании для канала связи с шумом.

Конечно, использованный нами метод (усреднение по случайным кодам) доказательства того, что равенство $R = C$ асимптотически достижимо, не очень конструктивен. Так как случайный код не имеет структуры или схемы, то кодирование и декодирование будут довольно громоздкими (нам нужна экспоненциально большая книга кодов). Тем не менее эта теорема важна и полезна, поскольку она говорит о том, что в принципе достижимо и, более того, что недостижимо, даже в принципе. К тому же, поскольку $I(X;Y)$ является вогнутой функцией от $X = \{x, p(x)\}$ (при фиксированном $\{p(y|x)\}$), то она имеет единственный локальный максимум, а C для интересуемого канала связи часто может быть вычислена (по крайней мере численно).

5.2. Энтропия фон Неймана

В теории классической информации мы часто рассматриваем источник, который готовит сообщения из n букв ($n \gg 1$), причем каждая буква независимо извлекается из ансамбля $X = \{x, p(x)\}$. Мы видели, что информационная энтропия Шеннона $H(X)$ (асимптотически при $n \to \infty$) представляет собой количество приходящихся на одну букву несжимаемых битов информации.

Нас также могут интересовать корреляции между сообщениями. Корреляции между двумя ансамблями букв X и Y характеризуются условными вероятностями $p(y|x)$. Мы видели, что взаимная информация

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{5.41}$$

представляет собой приходящееся на одну букву количество битов информации об X, которую мы можем получить, читая Y (и наоборот). Если условные вероятности $p(y|x)$ характеризуют канал с шумом, то $I(X;Y)$ –

приходящееся на одну букву количество информации, которое может быть передано через канал связи (при *a priori* заданном распределении вероятностей для X-ов).

Мы хотели бы распространить эти понятия на *квантовую* информацию. Представим источник, который готовит сообщения из n букв, но теперь каждая буква выбирается из ансамбля квантовых состояний. Алфавит сигналов представляет собой множество квантовых состояний $\boldsymbol{\rho}_x$, каждое из которых появляется с определенной *априорной* вероятностью p_x.

Как мы уже подробно обсуждали, если наблюдателю неизвестно, какая буква приготовлена, то вероятность любого результата любого измерения буквы, выбранной из этого ансамбля, можно полностью охарактеризовать матрицей плотности

$$\boldsymbol{\rho} = \sum_x p_x \boldsymbol{\rho}_x; \tag{5.42}$$

для ПОЗМ $\{\mathbf{F}_a\}$ мы имеем

$$\mathrm{Prob}(a) = \mathrm{tr}\left(\mathbf{F}_a \boldsymbol{\rho}\right). \tag{5.43}$$

Для этой (или любой другой) матрицы плотности можно определить энтропию фон Неймана

$$S(\boldsymbol{\rho}) = -\,\mathrm{tr}\left(\boldsymbol{\rho}\log\boldsymbol{\rho}\right). \tag{5.44}$$

Конечно, если мы выберем ортонормированный базис $\{|a\rangle\}$, диагонализующий $\boldsymbol{\rho}$:

$$\boldsymbol{\rho} = \sum_a \lambda_a |a\rangle\langle a|, \tag{5.45}$$

то

$$S(\boldsymbol{\rho}) = H(A), \tag{5.46}$$

где $H(A)$ — энтропия Шеннона ансамбля $A = \{a, \lambda_a\}$.

В том случае, когда алфавит сигналов состоит из взаимно ортогональных чистых состояний, квантовый источник сводится к классическому; все сигнальные состояния идеально различимы и $S(\boldsymbol{\rho}) = H(A)$. Более интересен квантовый источник, сигнальные состояния которого $\boldsymbol{\rho}$ взаимно не коммутируют. Мы докажем, что энтропия фон Неймана является количественной мерой несжимаемой информации, содержащейся в квантовом источнике (в том случае, когда сигнальные состояния являются чистыми), почти как энтропия Шеннона является количественной мерой информации, содержащейся в классическом источнике.

На самом деле мы обнаружим, что энтропия фон Неймана играет двойственную роль. Она является количественной мерой не только *квантовой* информации, содержащейся в одной букве ансамбля (минимальное количество приходящихся на одну букву кубитов, необходимое для надежного кодирования информации), но и содержащейся в ней *классической* информации (максимальное количество приходящейся на одну букву информации — в битах, а не в кубитах — которое можно получить с помощью наилучшего измерения). Мы увидим, что энтропия фон Неймана входит в квантовую информацию еще одним, третьим способом, количественно определяя запутывание бинарного чистого состояния. Таким образом, теория квантовой информации в значительной мере занимается интерпретацией и применениями энтропии фон Неймана, подобно тому как классическая теория информации главным образом занимается интерпретацией и применениями энтропии Шеннона.

Фактически необходимый для развития квантовой теории информации математический аппарат очень похож на математику Шеннона (типичные последовательности, случайное кодирование,...); похож настолько, что временами скрывается то, что в концептуальном плане они на самом деле весьма различны. Центральной проблемой квантовой теории информации является то, что неортогональные чистые квантовые состояния нельзя идеально различить — особенность, не имеющая классического аналога.

5.2.1. Математические свойства $S(\rho)$

Имеется несколько часто используемых свойств $S(\rho)$ [многие из которых являются близкими аналогами свойств $H(X)$]. Ниже я привожу список некоторых из этих свойств. Большей частью их доказательства не сложны (заметным исключением является доказательство сильной субаддитивности) и включены в упражнения в конце главы[1].

(1) Чистота. Чистое состояние $\rho = |\varphi\rangle\langle\varphi|$ имеет $S(\rho) = 0$.

(2) Инвариантность. Энтропия не изменяется при унитарных преобразованиях базиса

$$S(\mathbf{U}\rho\mathbf{U}^{-1}) = S(\rho). \tag{5.47}$$

[1]Некоторые доказательства можно также найти в обзоре A. Wehrl, *General Properties of Entropy*, Rev. Mod. Phys. **50**, 221 (1978); или в главе 9 книги A. Peres, *Quantum Theory: Concept and Methods*, Kluwer Academic Publishers, New York et al 2002. [Подробное обсуждение математических свойств энтропии фон Неймана можно найти в книге М. Нильсен, И. Чанг, *Квантовые вычисления и квантовая информация*, М.: Мир, 2006. — *Прим. ред.*]

Это очевидно, поскольку $S(\rho)$ зависит только от собственных значений ρ.

(3) Максимум. Если ρ имеет D ненулевых собственных значений, то

$$S(\rho) \leqslant \log D, \tag{5.48}$$

где равенство достигается, когда все ненулевые собственные значения равны между собой. (Энтропия максимальна, когда квантовые состояния *равновероятны*.)

(4) Вогнутость. Для $\lambda_1, \lambda_2, \ldots \lambda_n \geqslant 0$ и $\lambda_1 + \lambda_2 + \ldots + \lambda_n = 1$

$$S(\lambda_1\rho_1 + \lambda_2\rho_2 + \ldots + \lambda_n\rho_n) \geqslant \lambda_1 S(\rho_1) + \lambda_2 S(\rho_2) + \ldots \lambda_n S(\rho_n). \tag{5.49}$$

То есть энтропия фон Неймана тем больше, чем нам *меньше известно* о том, как было приготовлено состояние. Это свойство является следствием выпуклости логарифмической функции.

(5) Энтропия измерения. Предположим, что в состоянии ρ измеряется наблюдаемая

$$\mathbf{A} = \sum_y |a_y\rangle a_y \langle a_y|, \tag{5.50}$$

так что результат a_y появляется с вероятностью

$$p(a_y) = \langle a_y|\rho|a_y\rangle. \tag{5.51}$$

Тогда энтропия Шеннона ансамбля всех исходов измерения $Y = \{a_y, p(a_y)\}$ удовлетворяет

$$H(Y) \geqslant S(\rho), \tag{5.52}$$

где равенство достигается для коммутирующих $\mathbf{A}$ и ρ. Математически это утверждение означает, что в любом базисе $S(\rho)$ возрастает при замене нулями всех недиагональных матричных элементов ρ. Физически это означает, что случайность результата измерения минимизуется, если выбирается измерение наблюдаемой, коммутирующей с матрицей плотности. Но если мы измеряем «плохую» наблюдаемую, то результат будет менее предсказуем.

(6) Энтропия приготовления. Если чистое состояние случайным образом извлекается из ансамбля $\{|\varphi_x\rangle, p_x\}$, так что матрица плотности равна

$$\rho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x|, \tag{5.53}$$

то

$$H(Y) \geqslant S(\rho), \tag{5.54}$$

где равенство достигается, если сигнальные состояния $|\varphi_x\rangle$ взаимно ортогональны. Это утверждение указывает на то, что перемешивание неортогональных чистых состояний ведет к *потере различимости*. [Мы не можем полностью восстановить информацию о том, какое состояние было приготовлено, поскольку, как мы обсудим позже, достижимый при выполнении измерения прирост информации не может превзойти $S(\rho)$.]

(7) Субаддитивность. Рассмотрим бипарную систему AB в состоянии ρ_{AB}. Тогда

$$S(\rho_{AB}) \leqslant S(\rho_A) + S(\rho_B), \tag{5.55}$$

где $\rho_A = \mathrm{tr}_B\, \rho_{AB}$, $\rho_B = \mathrm{tr}_A\, \rho_{AB}$, а равенство достигается при $\rho_{AB} = \rho_A \otimes \rho_B$. Таким образом, энтропия *аддитивна* для некоррелированных систем, в противном случае энтропия всей системы меньше суммы энтропий ее частей. Это свойство является аналогом свойства энтропии Шеннона

$$H(X,Y) \leqslant H(X) + H(Y), \tag{5.56}$$

(или $I(X;Y) \geqslant 0$); оно имеет место, поскольку некоторая информация в XY (или AB) закодирована в корреляциях между X и Y (A и B).

(8) Сильная субаддитивность. Для любого состояния ρ_{ABC} трехкомпонентной системы

$$S(\rho_{ABC}) + S(\rho_B) \leqslant S(\rho_{AB}) + S(\rho_{BC}). \tag{5.57}$$

Это свойство называется «сильной» субаддитивностью, поскольку оно сводится к (обычной) субаддитивности в случае одномерной B. Доказательство соответствующего свойства энтропии Шеннона довольно просто, однако для энтропии фон Неймана оно оказывается на удивление трудным[1]. Свойство сильной субаддитивности легче запомнить, если интерпретировать его следующим образом: AB и BC можно рассматривать как две *перекрывающиеся* подсистемы. Энтропия их объединения (ABC) плюс энтропия их пересечения (B) не превышает

[1] Набросок доказательства сильной субаддитивности энтропии фон Неймана приведен в статье A. Wehrl, *General Properties of Entropy*, Rev. Mod. Phys. **50**, 221 (1978); [На русском языке см. М. Нильсен, И. Чанг, *Квантовые вычисления и квантовая информация*, М.: Мир, 2006. – *Прим. ред.*]

сумму энтропий подсистем (AB и BC). Мы увидим, что сильная субаддитивность имеет глубокие и важные следствия.

(9) Неравенство треугольника (Неравенство Араки – Либа). Для бинарной системы

$$S(\rho_{AB}) \geqslant |S(\rho_A) - S(\rho_B)|. \tag{5.58}$$

Неравенство треугольника резко контрастирует с аналогичным свойством энтропии Шеннона

$$H(X,Y) \geqslant H(X), H(Y) \tag{5.59}$$

или

$$H(X|Y), H(Y|X) \geqslant 0. \tag{5.60}$$

Энтропия Шеннона классической бинарной системы превосходит энтропию Шеннона любой ее части – во всей системе содержится больше информации о ней, нежели в ее части! Но это не так для энтропии фон Неймана. В предельном случае бинарного чистого квантового состояния мы имеем $S(\rho_A) = S(\rho_B)$ (не равные нулю, если состояние запутано), в то время как $S(\rho_{AB}) = 0$. Бинарное состояние определенным образом приготовлено, но если мы измеряем наблюдаемые различных подсистем, то результаты измерений с неизбежностью становятся случайными и непредсказуемыми. Мы не можем различить, как было приготовлено состояние, наблюдая две подсистемы отдельно, поскольку информация закодирована скорее в нелокальных квантовых корреляциях. Сопоставление положительности условной энтропии Шеннона (в классическом случае) с неравенством треугольника (в квантовом случае) замечательно характеризует ключевое различие между квантовой и классической информацией.

5.2.2. Энтропия и термодинамика

Конечно, понятие энтропии впервые было введено в науку в термодинамике. Здесь я ненадолго отвлекусь на некоторые термодинамические приложения математических свойств $S(\rho)$.

Существует два различных (но связанных между собой) возможных подхода к основаниям квантовой статистической физики. В первом мы рассматриваем эволюцию изолированной (замкнутой) квантовой системы, но производим некоторое *сглаживание* (*coarse graining*), чтобы определить термодинамические переменные. Во втором подходе, который, возможно, физически более мотивирован, мы рассматриваем *открытую* систе-

му, квантовую систему в контакте с окружением, и следим за ее эволюцией, не контролируя окружение.

Для открытой системы определяющим математическим свойством энтропии фон Неймана является ее *субаддитивность*. Если система (A) и окружение (E) первоначально не коррелированы друг с другом

$$\boldsymbol{\rho}_{AE} = \boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_E, \tag{5.61}$$

то энтропия аддитивна

$$S(\boldsymbol{\rho}_{AE}) = S(\boldsymbol{\rho}_A) + S(\boldsymbol{\rho}_E). \tag{5.62}$$

Предположим теперь, что открытая система эволюционирует в течение некоторого времени. Эволюция описывается унитарным оператором $\mathbf{U}_{AE}$, действующим на комбинированную систему $A + E$:

$$\boldsymbol{\rho}_{AE} \rightarrow \boldsymbol{\rho}'_{AE} = \mathbf{U}_{AE}\boldsymbol{\rho}_{AE}\mathbf{U}_{AE}^{-1}, \tag{5.63}$$

а поскольку унитарная эволюция сохраняет S, то

$$S(\boldsymbol{\rho}'_{AE}) = S(\boldsymbol{\rho}_{AE}). \tag{5.64}$$

Наконец, применим свойство субаддитивности к состоянию $S(\boldsymbol{\rho}'_{AE})$, получая в результате

$$S(\boldsymbol{\rho}_A) + S(\boldsymbol{\rho}_E) = S(\boldsymbol{\rho}'_{AE}) \leqslant S(\boldsymbol{\rho}'_A) + S(\boldsymbol{\rho}'_B), \tag{5.65}$$

где равенство имеет место в случае, когда A и E остаются некоррелированными. Если мы определим «полную» энтропию Вселенной как сумму энтропии системы и энтропии окружения, то придем к выводу, что *энтропия Вселенной не может убывать*. Это одна из формулировок второго закона термодинамики. Заметим, однако, чтобы вывести этот «закон», мы предположили, что в начальном состоянии система и окружение были некоррелированы.

Обычно взаимодействие системы и окружения *будет* генерировать корреляции, так что (в предположении *отсутствия* начальных корреляций) энтропия действительно *будет нарастать*. Вспомните из нашего обсуждения основного уравнения в § 3.5, что обычно окружение быстро «забывает», так что, если наше время разрешения достаточно велико, то в каждый момент времени систему и окружение (фактически) можно рассматривать как

«первоначально» некоррелированные (марковское приближение). В этом предположении «полная» энтропия будет монотонно возрастать, асимптотически приближаясь к своему теоретическому максимуму, максимальному достижимому значению, согласующемуся со всеми законами сохранения (энергии, заряда, барионного числа и т. д.).

Действительно, обычное предположение, лежащее в основании квантовой статистической физики, состоит в том, что система и окружение находятся в «наиболее вероятной конфигурации», максимизирующей $S(\rho_A) + S(\rho_E)$. В этой конфигурации все «доступные» состояния равновероятны.

С микроскопической точки зрения первоначально закодированная в системе информация (наша способность отличать одно начальное состояние от другого, первоначально ортогонального, состояния) теряется; она оказывается закодированной в квантовом запутывании системы и окружения. В принципе эта информация могла бы быть восстановлена, но на практике локальным наблюдателям это совершенно недоступно. Следовательно, мы наблюдаем термодинамическую необратимость.

Конечно, мы можем применить эти рассуждения к большой замкнутой системе (всей Вселенной?). Мы можем разделить систему на малую ее часть и остаток (окружение малой части). Тогда сумма энтропий этих частей будет неубывающей. Это частный тип сглаживания. Эта часть замкнутой системы ведет себя подобно открытой системе, поэтому для больших систем микроканонический и канонический ансамбли статистической механики дают одинаковые предсказания.

5.3. Сжатие квантовых данных

Что является квантовым аналогом теоремы о кодировании без шума?

Рассмотрим длинное сообщение, состоящее из n букв, где каждая буква случайным образом выбирается из ансамбля чистых состояний

$$\{|\varphi_x\rangle, p_x\}, \tag{5.66}$$

а сами $|\varphi_x\rangle$ не обязательно ортогональны. (Например, $|\varphi_x\rangle$ может представлять собой состояние поляризации одного фотона.) Таким образом, каждая буква описывается матрицей плотности

$$\rho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x|, \tag{5.67}$$

а все сообщение целиком — матрицей плотности

$$\rho^n = \rho \otimes \rho \otimes \cdots \otimes \rho. \tag{5.68}$$

Зададим вопрос: насколько *избыточна* эта квантовая информация? Мы хотели бы придумать *квантовый код*, который *позволит* сжать сообщение в более узкое гильбертово пространство без потери точности его воспроизведения. Например, допустим, что у нас есть устройство квантовой памяти (жесткий диск квантового компьютера?), и нам известны *статистические* свойства записанных данных (то есть мы знаем ρ). Сжимая данные, мы хотим сэкономить объем памяти.

Оптимальное сжатие, которое может быть достигнуто, было найдено Беном Шумахером. Можете ли вы угадать ответ? Наилучшим возможным сжатием, совместимым со сколь угодно высокой точностью воспроизведения при $n \to \infty$ является сжатие в гильбертово пространство $\mathcal{H}$ с

$$\log(\dim \mathcal{H}) \simeq nS(\rho). \tag{5.69}$$

В этом отношении энтропия фон Неймана представляет собой количество *кубитов* квантовой информации, переносимых одной буквой сообщения. Например, если сообщение состоит из n фотонных состояний поляризации, то мы можем сжать его до $m = nS(\rho)$ фотонов – сжатие всегда возможно, за исключением случая $\rho = \frac{1}{2}\mathbf{1}$. (Мы не можем сжать случайные кубиты точно так же, как не можем сжать случайные биты.)

Доказательство теоремы Шумахера не составляет труда, если известны и понятны результаты Шеннона. Большой заслугой Шумахера была правильная постановка вопроса, что позволило впервые дать точную (квантово-) информационную теоретическую интерпретацию энтропии фон Неймана[1].

5.3.1. Сжатие квантовых данных: пример

Прежде чем обсуждать в общем виде протокол Шумахера сжатия квантовых данных, полезно рассмотреть простой пример. Предположим, что нашими буквами являются отдельные кубиты, извлекаемые из ансамбля

$$\begin{aligned} |\uparrow_z\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad p = \frac{1}{2}, \\ |\uparrow_x\rangle &= \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad p = \frac{1}{2}, \end{aligned} \tag{5.70}$$

[1] Как мы вскоре увидим, интерпретация $S(\rho)$ на языке *классической* информации, закодированной в квантовых состояниях, действительно была известна раньше.

так что матрица плотности каждой буквы имеет вид

$$\boldsymbol{\rho} = \frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z| + \frac{1}{2}|\uparrow_x\rangle\langle\uparrow_x| = \\ = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}. \tag{5.71}$$

Как очевидно из симметрии, собственными состояниями $\boldsymbol{\rho}$ являются кубиты, ориентированные вверх и вниз вдоль оси $\hat{n} = \frac{1}{\sqrt{2}}(\hat{x} + \hat{z})$:

$$|0'\rangle \equiv |\uparrow_{\hat{n}}\rangle = \begin{pmatrix} \cos\frac{\pi}{8} \\ \sin\frac{\pi}{8} \end{pmatrix}, \\ |1'\rangle \equiv |\downarrow_{\hat{n}}\rangle = \begin{pmatrix} \sin\frac{\pi}{8} \\ -\cos\frac{\pi}{8} \end{pmatrix}; \tag{5.72}$$

соответствующие им собственные значения:

$$\lambda(0') = \frac{1}{2} + \frac{1}{2\sqrt{2}} = \cos^2\frac{\pi}{8}, \\ \lambda(1') = \frac{1}{2} - \frac{1}{2\sqrt{2}} = \sin^2\frac{\pi}{8}; \tag{5.73}$$

[очевидно, что $\lambda(0') + \lambda(1') = 1$, а $\lambda(0')\lambda(1') = 1/8 = \det\boldsymbol{\rho}$]. Собственное состояние $|0'\rangle$ одинаково (и довольно сильно) перекрывается с обоими сигнальными состояниями

$$|\langle 0'|\uparrow_z\rangle|^2 = |\langle 0'|\uparrow_x\rangle|^2 = \cos^2\frac{\pi}{8} = 0{,}8535; \tag{5.74}$$

перекрытия состояния $|1'\rangle$ тоже одинаковы (но относительно слабы)

$$|\langle 1'|\uparrow_z\rangle|^2 = |\langle 1'|\uparrow_x\rangle|^2 = \sin^2\frac{\pi}{8} = 0{,}1465. \tag{5.75}$$

Таким образом, если мы не знаем, какое состояние было послано, $|\uparrow_z\rangle$ или $|\uparrow_x\rangle$, то лучшей нашей догадкой является $|\psi\rangle = |0'\rangle$. Это предположение имеет максимальную *точность воспроизведения*

$$F = \frac{1}{2}|\langle\uparrow_z|\psi\rangle|^2 + \frac{1}{2}|\langle\uparrow_x|\psi\rangle|^2 \tag{5.76}$$

среди всех возможных состояний кубита $|\psi\rangle$ ($F = 0{,}8535$).

Теперь представим, что Алисе нужно послать Бобу три буквы. Но она может позволить себе послать только два кубита (квантовые каналы требуют очень больших расходов!). Тем не менее она хочет, чтобы Боб реконструировал ее состояние с максимально возможной точностью воспроизведения.

Она могла бы послать Бобу две из имеющихся у нее трех букв и предложить Бобу угадать $|0'\rangle$ для третьей. Тогда Боб получает две буквы с $F = 1$ и имеет $F = 0{,}8535$ для третьей; следовательно, полная $F = 0{,}8535$. Но существует ли более разумная процедура, достигающая более высокой точности воспроизведения?

Лучшая процедура действительно *существует*. Диагонализовав ρ, мы разложили гильбертово пространство одного кубита на «вероятное» (натянутое на $|0'\rangle$) и «маловероятное» (натянутое на $|1'\rangle$) одномерные подпространства. Подобным образом мы можем разложить гильбертово пространство трех кубитов на вероятное и маловероятное подпространства. Если произвольное сигнальное состояние имеет вид $|\psi\rangle = |\psi_1\rangle|\psi_2\rangle|\psi_3\rangle$ (с каждым из трех кубитов, находящихся в состоянии $|\uparrow_z\rangle$ или $|\uparrow_x\rangle$), то

$$
\begin{aligned}
|\langle 0'0'0'|\psi\rangle|^2 &= \cos^6\frac{\pi}{8} = 0{,}6219,\\
|\langle 0'0'1'|\psi\rangle|^2 = |\langle 0'1'0'|\psi\rangle|^2 = |\langle 1'0'0'|\psi\rangle|^2 &= \cos^4\frac{\pi}{8}\sin^2\frac{\pi}{8} = 0{,}1067,\\
|\langle 0'1'1'|\psi\rangle|^2 = |\langle 1'0'1'|\psi\rangle|^2 = |\langle 1'1'0'|\psi\rangle|^2 &= \cos^2\frac{\pi}{8}\sin^4\frac{\pi}{8} = 0{,}0183,\\
|\langle 1'1'1'|\psi\rangle|^2 &= \sin^6\frac{\pi}{8} = 0{,}0031.
\end{aligned}
\tag{5.77}
$$

Таким образом, мы можем разложить пространство на вероятное подпространство Λ, натянутое на $\{|0'0'0'\rangle, |0'0'1'\rangle, |0'1'0'\rangle, |1'0'0'\rangle\}$, и его ортогональное дополнение $\Lambda^\perp$. Если мы выполняем («грубое») измерение, проецирующее сигнальное состояние на Λ или $\Lambda^\perp$, то вероятность проецирования на вероятное подпространство равна

$$P_{\text{likely}} = 0{,}6219 + 3 \times 0{,}1067 = 0{,}9419, \tag{5.78}$$

тогда как вероятность проецирования на маловероятное подпространство —

$$P_{\text{unlikely}} = 3 \times 0{,}0183 + 0{,}0031 = 0{,}0581. \tag{5.79}$$

Чтобы выполнить это грубое измерение, Алиса могла бы, например, сначала применить унитарное преобразование U, превращающее четыре высоко вероятных базисных состояния в

$$|\cdot\rangle|\cdot\rangle|0\rangle, \tag{5.80}$$

а четыре маловероятных базисных состояния в

$$|\cdot\rangle|\cdot\rangle|1\rangle\,; \tag{5.81}$$

затем Алиса измеряет третий кубит, чтобы закончить грубое измерение. Если результатом является $|0\rangle$, то ее входящее состояние было спроецировано (в действительности) на Λ. Она посылает Бобу два оставшихся (неизмеренные) кубита. Когда Боб получает это (сжатое) двухкубитовое состояние $|\psi_{\rm comp}\rangle$, он развертывает его, присоединяя $|0\rangle$ и применяя $\mathbf{U}^{-1}$,

$$|\psi'\rangle = \mathbf{U}^{-1}(|\psi_{\rm comp}\rangle|0\rangle). \tag{5.82}$$

Если измерение Алисы третьего кубита дает $|1\rangle$, то она спроецировала свое входящее состояние на маловероятное подпространство $\Lambda^{\perp}$. Лучшее, что она может сделать в этом случае, это послать состояние, которое Боб развертывает в самое вероятное состояние $|0'0'0'\rangle$; то есть она посылает такое состояние $|\psi_{\rm comp}\rangle$, что

$$|\psi'\rangle = \mathbf{U}^{-1}(|\psi_{\rm comp}\rangle|0\rangle) = |0'0'0'\rangle. \tag{5.83}$$

Таким образом, если Алиса кодирует трехкубитовое сигнальное состояние $|\psi\rangle$, посылает два кубита Бобу, а Боб декодирует как только что описано, тогда он получает состояние

$$|\psi\rangle\langle\psi| \;\rightarrow\; \boldsymbol{\rho} = \mathbf{E}|\psi\rangle\langle\psi|\mathbf{E} + |0'0'0'\rangle\langle\psi|(\mathbf{1}-\mathbf{E})|\psi\rangle\langle 0'0'0'|, \tag{5.84}$$

где $\mathbf{E}$ – проекционный оператор на Λ. Достигаемая в этой процедуре точность воспроизведения равна

$$F = \langle\psi|\boldsymbol{\rho}|\psi\rangle = \big(\langle\psi|\mathbf{E}|\psi\rangle\big)^2 + \big(\langle\psi|(\mathbf{1}-\mathbf{E})|\psi\rangle\big)\big(\langle\psi|0'0'0'\rangle\big)^2 = \\ = (0{,}9419\dot{})^2 + 0{,}0581 \times 0{,}6219 = 0{,}9234. \tag{5.85}$$

Это действительно лучше наивной процедуры посылки двух из трех кубитов, каждого с идеальной точностью воспроизведения.

Когда мы посылаем более длинные сообщения с большим количеством букв, точность воспроизведения сжатия улучшается. Энтропия фон Неймана однокубитового ансамбля равна

$$S(\boldsymbol{\rho}) = H\left(\cos^2\frac{\pi}{8}\right) = 0{,}60088\ldots \tag{5.86}$$

Следовательно, согласно теореме Шумахера мы можем сократить длинное сообщение на фактор (скажем) 0,6009 и тем не менее достичь очень высокой точности воспроизведения.

5.3.2. Кодирование Шумахера в общем

Ключом к теореме Шеннона о кодировании в отсутствии шума является то, что мы без большой потери точности воспроизведения можем кодировать типичные последовательности и игнорировать остальные. Чтобы количественно описать сжимаемость квантовой информации, перейдем от понятия типичной *последовательности* к понятию типичного *подпространства*. Ключом к теореме Шумахера о квантовом кодировании в отсутствии шума является то, что мы без большой потери точности воспроизведения можем кодировать типичные подпространства и игнорировать их ортогональные дополнения.

Рассмотрим сообщение, состоящее из n букв, где каждая буква является чистым квантовым состоянием, извлекаемым из ансамбля $\{|\varphi_x\rangle, p_x\}$, так что матрица плотности одной буквы равна

$$\rho = \sum_x p_x |\varphi_x\rangle\langle\varphi_x|. \tag{5.87}$$

Более того, буквы извлекаются независимо, так что матрица плотности всего сообщения

$$\rho^n = \rho \otimes \rho \otimes \cdots \otimes \rho. \tag{5.88}$$

Мы хотим доказать, что для больших n почти все носители этой матрицы плотности занимают подпространство полного гильбертова пространства сообщений, причем размерность этого подпространства асимптотически стремится к $2^{nS(\rho)}$.

Этот вывод непосредственно следует из соответствующего классического утверждения, если мы рассматриваем ортонормированный базис, в котором ρ диагональна. Работая в этом базисе, по существу, мы можем рассматривать наш квантовый источник информации как эффективный классический источник, производящий сообщения, которые представляют собой строки из собственных состояний ρ. Вероятность каждого такого сообщения определяется произведением соответствующих собственных значений ρ. Для заданных n и δ определим типичное подпространство Λ как пространство, натянутое на собственные векторы ρ^n, с собственными значениями, удовлетворяющими

$$2^{-n(S+\delta)} \leqslant \lambda \leqslant 2^{-n(S-\delta)}. \tag{5.89}$$

Непосредственно пользуясь результатом Шеннона, мы приходим к выводу, что для любых δ, $\varepsilon > 0$ и при достаточно большом n сумма подчиняющихся

этому условию собственных значений $\boldsymbol{\rho}^n$ удовлетворяет неравенству

$$\mathrm{tr}(\boldsymbol{\rho}^n \mathbf{E}) > 1 - \varepsilon \tag{5.90}$$

(где $\mathbf{E}$ обозначает проекционный оператор на типичное подпространство), а количество $\dim \Lambda$ таких собственных значений удовлетворяет неравенству

$$(1-\varepsilon)2^{n(S-\delta)} \leqslant \dim \Lambda \leqslant 2^{n(S+\delta)}. \tag{5.91}$$

Наша стратегия кодирования состоит в том, чтобы отправлять состояния, действительно принадлежащие типичному подпространству. Например, мы можем выполнить грубое измерение, проецирующее входящее сообщение на Λ или на $\Lambda^\perp$; с вероятностью $P_\Lambda = \mathrm{tr}(\boldsymbol{\rho}^n \mathbf{E}) > 1 - \varepsilon$ результат будет принадлежать Λ. В таком случае спроецированное состояние кодируется и посылается. Асимптотически вероятность другого результата пренебрежимо мала, поэтому не так уж важно, что мы будем делать в этом случае.

Кодирование спроецированного состояния просто упаковывает его, чтобы оно могло переноситься минимальным количеством кубитов. Например, мы применяем унитарное преобразование базиса $\mathbf{U}$, которое превращает каждое состояние $|\psi_{\mathrm{typ}}\rangle$ из Λ в состояние вида

$$\mathbf{U}|\psi_{\mathrm{typ}}\rangle = |\psi_{\mathrm{comp}}\rangle|0_{\mathrm{rest}}\rangle, \tag{5.92}$$

где $|\psi_{\mathrm{comp}}\rangle$ — состояние $n(S+\delta)$ кубитов, а $|0_{\mathrm{rest}}\rangle$ обозначает состояние $|0\rangle \otimes \ldots \otimes |0\rangle$ остальных кубитов. Алиса посылает $|\psi_{\mathrm{comp}}\rangle$ Бобу, который декодирует его, присоединяя $|0_{\mathrm{rest}}\rangle$ и применяя $\mathbf{U}^{-1}$.

Предположим, что

$$|\varphi_i\rangle = |\varphi_{x_1(i)}\rangle|\varphi_{x_2(i)}\rangle \cdots |\varphi_{x_n(i)}\rangle \tag{5.93}$$

обозначает произвольное сообщение, представляющее собой одно из n-буквенных чистых состояний, которое может быть послано. После того как выполнены только что описанные кодирование, передача и декодирование, Боб получает реконструированное состояние

$$|\varphi_i\rangle\langle\varphi_i| \to \boldsymbol{\rho}_i' = \mathbf{E}|\varphi_i\rangle\langle\varphi_i|\mathbf{E} + \boldsymbol{\rho}_{i,Junk}\langle\varphi_i|(\mathbf{1}-\mathbf{E})|\varphi_i\rangle, \tag{5.94}$$

где $\boldsymbol{\rho}_{i,Junk}$ — состояние, выбираемое нами для отправления, если грубое измерение дает результат $\Lambda^\perp$. Что можно сказать относительно точности воспроизведения этой процедуры?

Точность воспроизведения изменяется от сообщения к сообщению (в противоположность обсуждавшемуся выше примеру), поэтому мы рассматриваем точность воспроизведения, усредненную по ансамблю возможных сообщений:

$$F = \sum_i p_i \langle\varphi_i|\boldsymbol{\rho}'_i|\varphi_i\rangle = \sum_i p_i \langle\varphi_i|\mathbf{E}|\varphi_i\rangle\langle\varphi_i|\mathbf{E}|\varphi_i\rangle + \\ + \sum_i p_i \langle\varphi_i|\boldsymbol{\rho}'_{i,Junk}|\varphi_i\rangle\langle\varphi_i|\mathbf{1}-\mathbf{E}|\varphi_i\rangle \geqslant \sum_i p_i \|\mathbf{E}|\varphi_i\rangle\|^4, \quad (5.95)$$

где последнее неравенство справедливо, поскольку вклад «мусора» неотрицателен. Так как для любого вещественного числа

$$(x-1)^2 \geqslant 0 \quad \text{или} \quad x^2 \geqslant 2x-1, \quad (5.96)$$

мы имеем (полагая $x = \|\mathbf{E}|\varphi_i\rangle\|^2$)

$$\|\mathbf{E}|\varphi_i\rangle\|^4 \geqslant 2\|\mathbf{E}|\varphi_i\rangle\|^2 - 1 \quad (5.97)$$

и, следовательно,

$$F \geqslant \sum_i p_i \big(2\langle\varphi_i|\mathbf{E}|\varphi_i\rangle - 1\big) = 2\,\mathrm{tr}\,(\boldsymbol{\rho}^n\mathbf{E}) - 1 > 2(1-\varepsilon) - 1 = 1 - 2\varepsilon. \quad (5.98)$$

Итак, мы показали, что можно сжать сообщение до объема, несколько меньшего, чем $n(S+\delta)$ кубитов, обеспечивая в то же время сколь угодно высокую при больших n усредненную точность воспроизведения.

Следовательно, мы установили, что с несущественной потерей точности воспроизведения сообщение может быть сжато до $S+\delta$ кубитов в расчете на одну букву. Возможно ли дальнейшее сжатие?

Допустим, что Боб декодирует полученное сообщение $\boldsymbol{\rho}_{\mathrm{comp,i}}$ путем присоединения кубитов и применения преобразования $\mathbf{U}^{-1}$, получая при этом

$$\boldsymbol{\rho}' = \mathbf{U}^{-1}(\boldsymbol{\rho}_{\mathrm{comp,i}} \otimes |0\rangle\langle 0|)\mathbf{U} \quad (5.99)$$

(«унитарное декодирование»). Допустим, что $\boldsymbol{\rho}_{\mathrm{comp,i}}$ было сжато до $n(S-\delta)$ кубитов. Тогда, *независимо от того, как было закодировано входящее сообщение*, все декодированные сообщения будут принадлежать подпространству Λ' размерности $2^{n(S-\delta)}$ гильбертова пространства Боба. (Мы не предполагаем здесь, что Λ' не имеет ничего общего с типичным подпространством.)

Если входящим сообщением является $|\varphi_i\rangle$, то реконструированное Бобом сообщение — $\boldsymbol{\rho}_i'$, которое может быть диагонализовано

$$\boldsymbol{\rho}_i' = \sum_{a_i} |a_i\rangle \lambda_{a_i} \langle a_i|, \tag{5.100}$$

где векторы $|a_i\rangle$ — взаимно ортогональные состояния из Λ'. Точность воспроизведения реконструированного сообщения равна

$$F_i = \langle \varphi_i | \boldsymbol{\rho}_i' | \varphi_i \rangle = \sum_{a_i} \lambda_{a_i} \langle \varphi_i | a_i \rangle \langle a_i | \varphi_i \rangle \leqslant$$

$$\leqslant \sum_{a_i} \langle \varphi_i | a_i \rangle \langle a_i | \varphi_i \rangle \leqslant \langle \varphi_i | \mathbf{E}' | \varphi_i \rangle, \tag{5.101}$$

где $\mathbf{E}'$ — ортогональный проектор на подпространство Λ'. Следовательно, усредненная точность воспроизведения удовлетворяет

$$F = \sum_i p_i F_i \leqslant \sum_i p_i \langle \varphi_i | \mathbf{E}' | \varphi_i \rangle = \mathrm{tr}(\boldsymbol{\rho}^n \mathbf{E}'). \tag{5.102}$$

Но поскольку $\mathbf{E}'$ проецирует на пространство размерности $2^{n(S-\delta)}$, то $\mathrm{tr}(\boldsymbol{\rho}^n \mathbf{E}')$ не может быть больше суммы $2^{n(S-\delta)}$ наибольших собственных значений $\boldsymbol{\rho}^n$. Из свойств типичных подпространств следует, что эта сумма принимает сколь угодно малое значение; при достаточно большом n

$$F \leqslant \mathrm{tr}(\boldsymbol{\rho}^n \mathbf{E}') < \varepsilon. \tag{5.103}$$

Таким образом, мы показали, что если мы попытаемся сжать сообщение до $S - \delta$ кубитов на одну букву, тогда при достаточно большом n точность воспроизведения неизбежно станет плохой. Итак, мы приходим к выводу, что $S(\boldsymbol{\rho})$ кубитов на одну букву является оптимальным сжатием квантовой информации, которое может быть достигнуто, если мы хотим получить хорошую точность воспроизведения при n стремящемся к бесконечности. В этом состоит теорема Шумахера о кодировании в отсутствии шума.

Приведенное выше доказательство применимо к любой мыслимой схеме кодирования, но только к ограниченному классу схем декодирования (унитарное декодирование). Конечно, может быть рассмотрена более общая схема декодирования, описываемая *супероператором*. Тогда потребуется более высокая техника для доказательства того, что невозможно лучшее сжатие, чем S кубитов на одну букву. Но вывод остается неизменным. Суть в том, что $S - \delta$ кубитов недостаточно для того, чтобы различить все типичные состояния.

Подводя итог, отметим тесную аналогию между теоремами Шеннона и Шумахера о кодировании в отсутствии шума. В классическом случае почти все длинные сообщения являются типичными последовательностями, так что мы можем кодировать только их и тем не менее иметь малую вероятность ошибки. В квантовом случае почти все длинные сообщения имеют почти единичное перекрытие с типичным подпространством, так что мы можем кодировать только его и тем не менее достигать высокой точности воспроизведения.

Фактически Алиса могла бы послать Бобу эффективно классическую информацию — строку $x_1 x_2 \ldots x_n$, закодированную во взаимно ортогональных квантовых состояниях — тогда Боб, следуя этим классическим инструкциям, мог бы реконструировать состояние Алисы. Таким способом они могли бы добиться высоко надежного сжатия до $H(X)$ битов (или кубитов) в расчете на одну букву. Но если буквы извлекаются из ансамбля *неортогональных* чистых состояний, то эта степень сжатия не оптимальна; часть классической информации о приготовлении состояния становится избыточной, поскольку неортогональные состояния не могут быть идеально различимыми. Таким образом, кодирование Шумахера может продвинуться дальше, достигая оптимального сжатия $S(\rho)$ кубитов на одну букву сообщения. Информация упакована более эффективно, но дорогой ценой — Боб получил то, что имела ввиду Алиса, но он не может узнать — что. В противоположность классическому случаю, Боб не может выполнить никакого измерения, чтобы корректно дешифровать сообщение Алисы. Попытка прочитать сообщение неизбежно внесет в него возмущение.

5.3.3. Кодирование смешанного состояния: информация Холево

Теорема Шумахера характеризует сжимаемость ансамбля чистых состояний. Но что если буквы извлекаются из ансамбля смешанных состояний? В этом случае сжимаемость надежно не установлена и является предметом текущих исследований[1].

Нетрудно видеть, что для смешанных состояний $S(\rho)$ уже не будет ответом. Чтобы привести тривиальный пример, предположим, что некоторое смешанное состояние ρ_0 с энтропией $S(\rho_0) \neq 0$ выбирается с вероятностью $p_0 = 1$. Тогда сообщение всегда равно $\rho_0 \otimes \rho_0 \otimes \cdots \otimes \rho_0$ и не несет никакой информации; Боб может идеально реконструировать сообщение, *ничего не получая* от Алисы. Следовательно, это сообщение можно сжать до нуля кубитов на одну букву, что меньше, чем $S(\rho_0) > 0$.

[1] См. M. Horodecki, *Limits for Compression of Quantum Information Carried by Ensembles of Mixed States*, Phys. Rev., **A57**, 3364–3369 (1997); quant-ph/9712035.

Чтобы построить менее тривиальный пример, вспомним, что для ансамбля взаимно ортогональных чистых состояний энтропия Шеннона равна энтропии фон Неймана:

$$H(X) = S(\rho). \tag{5.104}$$

так что классическая и квантовая сжимаемости совпадают. Это справедливо, поскольку ортогональные состояния идеально различимы. Фактически, если Алиса хочет послать Бобу сообщение

$$|\varphi_{x_1}\rangle|\varphi_{x_2}\rangle\cdots|\varphi_{x_n}\rangle, \tag{5.105}$$

то она может послать классическое сообщение $x_1 \ldots x_n$, а Боб может реконструировать состояние с идеальной точностью воспроизведения.

Теперь предположим, что буквы извлекаются из ансамбля взаимно ортогональных *смешанных* состояний $\{\rho_x, p_x\}$:

$$\operatorname{tr} \rho_x \rho_y = 0, \qquad x \neq y; \tag{5.106}$$

то есть ρ_x и ρ_y имеют носители во взаимно ортогональных подпространствах гильбертова пространства. Эти смешанные состояния также идеально различимы, то есть опять сообщения, по существу, классические и, следовательно, могут быть сжаты до $H(X)$ кубитов на одну букву. Например, мы можем расширить гильбертово пространство наших букв $\mathcal{H}_A$ до более широкого пространства $\mathcal{H}_A \otimes \mathcal{H}_B$ и выбрать очищение каждого ρ_x, то есть чистое состояние $|\varphi_x\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$, такое что

$$\operatorname{tr}_B \left(|\varphi_x\rangle_{AB\ AB}\langle\varphi_x|\right) = (\rho_x)_A. \tag{5.107}$$

Эти чистые состояния взаимно ортогональны, а ансамбль $\{|\varphi_x\rangle_{AB}, p_x\}$ имеет энтропию фон Неймана $H(X)$; следовательно, мы можем выполнить сжатие сообщения

$$|\varphi_{x_1}\rangle_{AB}\cdots|\varphi_{x_n}\rangle_{AB} \tag{5.108}$$

по Шумахеру до $H(X)$ кубитов на одну букву (асимптотически). После развертывания этого состояния Боб может взять частичный след, «выбрасывая» подсистему B, и таким образом реконструировать сообщение Алисы.

Чтобы сделать разумное предположение о том, какое выражение характеризует сжимаемость сообщения, построенного из алфавита смешанных состояний, мы могли бы поискать выражение, которое сводится к $S(\rho)$ для ансамбля чистых состояний, и — к $H(X)$ для ансамбля взаимно ортогональных смешанных состояний. Выбирая базис, в котором

$$\rho = \sum_x p_x \rho_x \tag{5.109}$$

является блочно-диагональным, мы видим, что

$$\begin{aligned} S(\rho) &= -\operatorname{tr}\rho\log\rho = -\sum_x \operatorname{tr}\left(p_x\rho_x\right)\log\left(p_x\rho_x\right) = \\ &= -\sum_x p_x\log p_x - \sum_x p_x \operatorname{tr}\rho_x\log\rho_x = \\ &= H(X) + \sum_x p_x S(\rho_x) \end{aligned} \tag{5.110}$$

(вспоминая, что $\operatorname{tr}\rho_x = 1$ для каждого x). Следовательно, мы можем записать энтропию Шеннона в виде

$$H(X) = S(\rho) - \sum_x p_x S(\rho_x) \equiv \chi(\mathcal{E}). \tag{5.111}$$

Величина $\chi(\mathcal{E})$ называется *информацией Холево* ансамбля $\mathcal{E} = \{\rho_x, p_x\}$. Очевидно, она зависит не только от матрицы плотности ρ, но и от конкретного способа реализации ρ как ансамбля смешанных состояний. Мы напли, что как для ансамбля чистых состояний, так и для ансамбля *взаимно ортогональных* смешанных состояний информация Холево $\chi(\mathcal{E})$ представляет собой оптимальное количество кубитов на одну букву, которого можно достичь, если мы сжимаем сообщение, сохраняя высокую точность воспроизведения при больших n.

Информация Холево может рассматриваться как обобщение энтропии фон Неймана, переходящее в $S(\rho)$ для ансамбля чистых состояний. Она также является близким аналогом взаимной информации

$$I(Y;X) = H(Y) - H(Y|X) \tag{5.112}$$

в классической теории информации, сообщающей нам, насколько в среднем уменьшается энтропия Шеннона ансамбля Y, когда мы узнаем значение X; аналогично

$$\chi(\mathcal{E}) = S(\rho) - \sum_x p_x S(\rho_x) \tag{5.113}$$

говорит нам, насколько в среднем уменьшается энтропия фон Неймана ансамбля, когда мы узнаем, как он был приготовлен. Подобно классической взаимной информации, информация Холево всегда неотрицательна, как это следует из свойства вогнутости $S(\rho)$

$$S\left(\sum p_x\rho_x\right) \geqslant \sum p_x S(\rho_x). \tag{5.114}$$

Теперь мы хотим исследовать связь между информацией Холево и сжимаемостью сообщений, построенных из алфавита *неортогональных* смешанных состояний. Фактически можно показать, что в общем случае невозможно сжатие с высокой точностью воспроизведения до объема, меньшего чем χ на одну букву сообщения.

Чтобы установить этот результат, воспользуемся свойством «монотонности» χ, доказанным Линдбладом и Ульманом: супероператор не может увеличивать информацию Холево. То есть если $\$$ — произвольный супероператор, действующий на ансамбль смешанных состояний как

$$\$\colon \mathcal{E} = \{\boldsymbol{\rho}_x, p_x\} \rightarrow \mathcal{E}' = \{\$(\boldsymbol{\rho}_x), p_x\}, \tag{5.115}$$

то

$$\chi(\mathcal{E}') \leqslant \chi(\mathcal{E}). \tag{5.116}$$

Монотонность Линдблада – Ульмана тесно связана с сильной субаддитивностью энтропии фон Неймана, что вы покажете в домашнем упражнении.

Монотонность χ обеспечивает еще одно свидетельство того, что χ характеризует количество информации, закодированной в квантовой системе. Декогерентизация, описываемая супероператором, может лишь сохранить или сократить эту величину информации, но не увеличить ее. Заметим, что в противоположность этому энтропия фон Неймана не монотонна. Супероператор может преобразовать начальное чистое состояние в смешанное, увеличивая $S(\boldsymbol{\rho})$. Однако другой супероператор преобразует любое смешанное состояние в «основное» $|0\rangle\langle 0|$ и, следовательно, уменьшает энтропию начального смешанного состояния до нуля. Было бы ошибкой интерпретировать это уменьшение S как «приобретение информации», так как наша способность отличить разные возможные приготовления полностью утрачена. Соответственно распад в основное состояние сокращает до нуля информацию Холево, отражая то, что мы потеряли возможность реконструировать начальное состояние.

Рассмотрим теперь сообщения из n букв, независимо извлекаемых из ансамбля $\mathcal{E} = \{\boldsymbol{\rho}_x, p_x\}$; ансамбль всех таких входящих сообщений обозначается как $\mathcal{E}^{(n)}$. Пусть разработан код, который сжимает сообщения так, что они все занимают гильбертово пространство $\tilde{\mathcal{H}}^{(n)}$; ансамбль сжатых сообщений обозначается как $\tilde{\mathcal{E}}^{(n)}$. Тогда развертывание выполняется супероператором $\$$

$$\$\colon \tilde{\mathcal{E}}^{(n)} \rightarrow \mathcal{E}'^{(n)}, \tag{5.117}$$

чтобы получить ансамбль $\mathcal{E}'^{,(n)}$ выходящих сообщений.

Теперь предположим, что эта схема кодирования имеет высокую точность воспроизведения. Чтобы свести к минимуму техническую сторону, не будем вдаваться в детали того, как следует охарактеризовать количественно точность воспроизведения $\mathcal{E}'^{(n)}$ относительно $\mathcal{E}^{(n)}$. Мы просто примем, что если $\mathcal{E}'^{(n)}$ имеет высокую точность воспроизведения, то для любого δ и при достаточно больших n

$$\frac{1}{n}\chi\bigl(\mathcal{E}^{(n)}\bigr) - \delta \leqslant \frac{1}{n}\chi\bigl(\mathcal{E}'^{(n)}\bigr) \leqslant \frac{1}{n}\chi\bigl(\mathcal{E}^{(n)}\bigr) + \delta; \tag{5.118}$$

отнесенная к одной букве информации Холево выходящего и входящего сообщений приближаются друг к другу. Поскольку входящие сообщения являются произведениями состояний, то из аддитивности $S(\boldsymbol{\rho})$ следует, что

$$\chi\bigl(\mathcal{E}^{(n)}\bigr) = n\chi(\mathcal{E}), \tag{5.119}$$

а из монотонности Линдблада – Ульмана мы знаем, что

$$\chi\bigl(\mathcal{E}'^{(n)}\bigr) \leqslant \chi\bigl(\tilde{\mathcal{E}}^{(n)}\bigr). \tag{5.120}$$

Комбинируя уравнения (5.118)–(5.120), находим, что

$$\frac{1}{n}\chi\bigl(\tilde{\mathcal{E}}^{(n)}\bigr) \geqslant \chi(\mathcal{E}) - \delta. \tag{5.121}$$

Наконец, $\chi\bigl(\tilde{\mathcal{E}}^{(n)}\bigr)$ ограничена сверху величиной $S\bigl(\tilde{\boldsymbol{\rho}}^{(n)}\bigr)$, которая, в свою очередь, ограничена сверху числом $\log\dim\tilde{\mathcal{H}}^{(n)}$. Так как δ можно выбрать сколь угодно малой, мы приходим к выводу, что асимптотически при $n \to \infty$

$$\frac{1}{n}\log\dim\tilde{\mathcal{H}}^{(n)} \geqslant \chi(\mathcal{E}); \tag{5.122}$$

хорошо воспроизводимое сжатие, до менее чем $\chi(\mathcal{E})$ кубитов на одну букву, невозможно.

Нередко возникает соблазн предположить, что сжатие до $\chi(\mathcal{E})$ кубитов на одну букву сообщения асимптотически достижимо. С середины января 1998 г. это предположение все еще ждет своего доказательства или опровержения.

5.4. Доступная информация

Тесная аналогия между информацией Холево $\chi(\mathcal{E})$ и классической взаимной информацией $I(X;Y)$, а также монотонность χ наводят на мысль,

что χ связана с количеством *классической* информации, которая может храниться в квантовой системе и извлекаться из нее. В этом разделе мы дадим точную формулировку этой связи.

Предыдущий раздел был посвящен количественному определению объема *квантовой* информации — измеряемой в кубитах — в сообщениях, построенных из алфавита квантовых состояний. Теперь же мы обратимся к совсем другому вопросу. Мы хотим количественно определить объем *классической* информации — измеряемой в битах — которую можно извлечь из таких сообщений, в частности, в случае, когда алфавит включает взаимно неортогональные буквы.

Но почему мы должны быть столь неразумны, чтобы хранить классическую информацию в неортогональных квантовых состояниях, которые нельзя идеально различить? Несомненно, следует избегать такого способа хранения информации, так как это ведет к деградации классического сигнала. Но, возможно, мы не можем обойтись без него. Допустим, например, я инженер связи и интересуюсь существенными физическими ограничениями классической емкости широкополосного оптического волокна. Чтобы получить наилучшую пропускную способность классической информации на единицу мощности, нам, очевидно, следует выбрать кодирование информации в отдельных фотонах, а чтобы добиться высокого темпа, мы должны увеличивать количество передаваемых за одну секунду фотонов. Но если мы сжимаем фотонные волновые пакеты достаточно тесно друг с другом, они начнут перекрываться и мы не сможем их идеально различать. Как максимизировать передаваемую в этом случае классическую информацию? Другой важный пример: допустим, что я — физик-экспериментатор и хочу использовать тонкую квантовую систему, чтобы сконструировать очень чувствительный прибор, измеряющий действие классической силы на систему. Мы можем моделировать силу как свободный параметр x в гамильтониане системы $\mathbf{H}(x)$. В зависимости от значения x состояние системы будет эволюционировать к различным возможным конечным (неортогональным) состояниям $\boldsymbol{\rho}_x$. Как много информации относительно x может получить наш прибор?

Несмотря на то, что с точки зрения физики эта проблема сильно отличается от сжимаемости квантовой информации, математически они связаны между собой. Мы обнаружим, что центральную роль в нашем обсуждении играют энтропия фон Неймана и ее обобщение, информация Холево.

Предположим, например, что Алиса готовит чистое квантовое состояние, извлекая его из ансамбля $\mathcal{E} = \{|\varphi_x\rangle, p_x\}$. Бобу известен ансамбль, но не конкретное выбранное Алисой состояние. Он хочет получить максимально возможную информацию относительно x.

Боб собирает информацию, выполняя обобщенное измерение, ПОЗМ $\{\mathbf{F}_y\}$. Если Алиса приготовила x, то Боб получит результат измерения y с условной вероятностью

$$p(y|x) = \langle\varphi_x|\mathbf{F}_y|\varphi_x\rangle. \tag{5.123}$$

Эти условные вероятности вместе с ансамблем X определяют количество в среднем получаемой Бобом информации, взаимную информацию $I(X;Y)$ приготовления и результата измерения.

Боб свободен в выборе своего измерения. «Наилучшее» возможное измерение, максимизирующее получение информации, называется *оптимальным измерением*, определяемым ансамблем. Максимальное получение информации равно

$$\mathrm{Acc}(\mathcal{E}) = \max_{\{\mathbf{F}_y\}} I(X;Y), \tag{5.124}$$

где max определяется по всем возможным ПОЗМ-ам. Эта величина называется *доступной информацией* ансамбля $\mathcal{E}$.

Конечно, если состояния $|\varphi_x\rangle$ взаимно ортогональны, то они идеально различимы. Условная вероятность результата ортогонального измерения

$$\mathbf{E}_y = |\varphi_y\rangle\langle\varphi_y| \tag{5.125}$$

равна

$$p(y|x) = \delta_{y,x}, \tag{5.126}$$

так что $H(X|Y) = 0$, а $I(X;Y) = H(X)$. Это измерение, очевидно, оптимально — приготовление полностью определено — так что для ансамбля взаимно ортогональных (чистых или смешанных) состояний

$$\mathrm{Acc}(\mathcal{E}) = H(X). \tag{5.127}$$

Однако проблема становится гораздо интереснее, когда сигнальными состояниями являются неортогональные чистые состояния. Для этого случая не известно ни одного полезного общего выражения для $\mathrm{Acc}(\mathcal{E})$, но существует верхняя граница

$$\mathrm{Acc}(\mathcal{E}) \leqslant S(\boldsymbol{\rho}). \tag{5.128}$$

Мы видели, что эта граница достигается в случае ортогональных сигнальных состояний, когда $S(\boldsymbol{\rho}) = H(X)$. В общем случае из классической теории информации известно, что $I(X;Y) \leqslant H(X)$; но для неортогональных

состояний $S(\rho) < H(X)$, так что неравенство (5.128) определяет наилучшую границу. Тем не менее эта граница не является точной, во многих случаях $\mathrm{Acc}(\mathcal{E})$ строго меньше $S(\rho)$.

Более определенное соотношение между $\mathrm{Acc}(\mathcal{E})$ и $S(\rho)$ мы получим, если рассмотрим доступную информацию в расчете на одну букву в сообщении, содержащем n букв. Теперь Боб имеет большую свободу — он может решить выполнить коллективное измерение всех n букв и таким образом получить больше информации, чем если бы он ограничивался измерением только по одной букве за один раз. Более того, Алиса может решить приготовить скорее ансамбль частных, максимально различимых, сообщений (код), нежели произвольные сообщения, каждая буква которых извлечена из ансамбля $\mathcal{E}$.

Тогда мы увидим, что Алиса и Боб могут найти такой код, чтобы маргинальным (частным) ансамблем каждой буквы был $\mathcal{E}$, а отнесенная к одной букве доступная информация асимптотически стремилась к $S(\rho)$ при $n \to \infty$. В этом смысле $S(\rho)$ характеризует доступную информацию в ансамбле *чистых* квантовых состояний.

Более того, заменой энтропии фон Неймана на информацию Холево эти результаты обобщаются на ансамбли смешанных квантовых состояний. Доступная информация ансамбля смешанных состояний $\{\rho_x, p_x\}$ удовлетворяет неравенству

$$\mathrm{Acc}(\mathcal{E}) \leqslant \chi(\mathcal{E}), \tag{5.129}$$

результат, известный как *граница Холево*. Эта граница в общем случае не является точной (хотя она достигается для ансамблей взаимно ортогональных смешанных состояний). Однако если Алиса и Боб выбирают n-буквенный код, в котором частным ансамблем для каждой буквы является $\mathcal{E}$, а Боб выполняет коллективное оптимальное обобщенное измерение (ПОЗМ) всех n букв, тогда максимальной достижимой информацией на одну букву является $\chi(\mathcal{E})$, если потребовать, чтобы все кодовые слова представляли собой *произведения* состояний. В этом смысле $\chi(\mathcal{E})$ характеризует доступную информацию в ансамбле *смешанных* квантовых состояний.

Алфавит из смешанных квантовых состояний может возникнуть, если Алиса попытается послать Бобу чистые квантовые состояния через квантовый канал с шумом. Вследствие декогерентизации в канале связи, Боб получает смешанные состояния, которые он должен декодировать. В этом случае $\chi(\mathcal{E})$ характеризует максимальное количество классической информации, которое может быть передано Бобу через квантовый канал с шумом.

Например, Алиса может послать Бобу n фотонов в определенных состояниях поляризации. Если предположить, что шум действует на каждый

фотон независимо и что Алиса посылает фотоны в незапутанных состояниях, тогда $\chi(\mathcal{E})$ – максимальное количество информации, которое может быть передано Бобу с каждым фотоном. Поскольку

$$\chi(\mathcal{E}) \leqslant S(\boldsymbol{\rho}) \leqslant 1, \tag{5.130}$$

отсюда, в частности, следует, что отдельный (незапутанный) фотон может переносить, самое большее, один бит классической информации.

5.4.1. Граница Холево

Граница Холево для доступной информации не относится к разряду очевидных теорем, но подобно многим интересным результатам квантовой теории информации, она становится очевидной, коль скоро установлена сильная субаддитивность энтропии фон Неймана. Здесь мы предположим наличие свойства сильной субаддитивности и покажем, что отсюда следует граница Холево.

Напомним исходные данные: Алиса готовит квантовое состояние, извлекаемое из ансамбля $\mathcal{E} = \{\boldsymbol{\rho}_x, p_x\}$, а затем Боб выполняет ПОЗМ $\{\mathbf{F}_y\}$. Совместным распределением вероятностей, управляющим приготовлением Алисы x и результатом Боба y является

$$p(x, y) = p_x \operatorname{tr}\left(\mathbf{F}_y \boldsymbol{\rho}_x\right). \tag{5.131}$$

Мы хотим показать, что

$$I(X;Y) \leqslant \chi(\mathcal{E}). \tag{5.132}$$

Поскольку сильная субаддитивность является свойством трех подсистем, нам нужно определить три системы, к которым оно будет применяться. Наша стратегия состоит в приготовлении входящей системы X, в которой хранится классическая запись того, какое приготовление было выбрано, и выходящей системы Y, классические корреляции которой с X управляются совместным распределением $p(x, y)$. Тогда, применяя свойство сильной субаддитивности к X, Y и нашей квантовой системе Q, мы сможем связать $I(X;Y)$ с $\chi(\mathcal{E})$.

Допустим, что начальным состоянием системы XQY является

$$\boldsymbol{\rho}_{XQY} = \sum_x p_x |x\rangle\langle x| \otimes \boldsymbol{\rho}_x \otimes |0\rangle\langle 0|, \tag{5.133}$$

где векторы $|x\rangle$ – взаимно ортогональные чистые состояния входящей системы X, а $|0\rangle$ – частное чистое состояние выходящей системы Y. Вычис-

ляя частичные следы, мы видим, что

$$\boldsymbol{\rho}_X = \sum_x p_x |x\rangle\langle x| \;\rightarrow\; S(\boldsymbol{\rho}_X) = H(X),$$
$$\boldsymbol{\rho}_Q = \sum_x p_x \boldsymbol{\rho}_x \equiv \boldsymbol{\rho} \;\rightarrow\; S(\boldsymbol{\rho}_{QY}) = S(\boldsymbol{\rho}_Q) = S(\boldsymbol{\rho}), \tag{5.134}$$

а так как векторы $|x\rangle$ взаимно ортогональны, мы также имеем

$$S(\boldsymbol{\rho}_{XQY}) = S(\boldsymbol{\rho}_{XQ}) = -\sum_x \operatorname{tr}\left(p_x \boldsymbol{\rho}_x \log p_x \boldsymbol{\rho}_x\right) =$$
$$= H(X) + \sum_x p_x S(\boldsymbol{\rho}_x). \tag{5.135}$$

Теперь выполним унитарное преобразование, которое «отпечатывает» результат измерения Боба на выходящей системе Y. Предположим пока, что Боб выполняет ортогональное измерение $\{\mathbf{E}_y\}$, где

$$\mathbf{E}_y \mathbf{E}_{y'} = \delta_{y,y'} \mathbf{E}_y \tag{5.136}$$

(вскоре мы кратко рассмотрим более общие ПОЗМ-ы). Наше унитарное преобразование $\mathbf{U}_{QY}$ действует на QY согласно

$$\mathbf{U}_{QY} : |\varphi\rangle_Q \otimes |0\rangle_Y = \sum_y \mathbf{E}_y |\varphi\rangle_Q \otimes |y\rangle_Y \tag{5.137}$$

(где векторы $|y\rangle_Y$ взаимно ортогональны) и, следовательно, преобразует $\boldsymbol{\rho}_{XQY}$ как

$$\mathbf{U}_{QY} : \boldsymbol{\rho}_{XQY} \;\rightarrow\; \boldsymbol{\rho}'_{XQY} = \sum_{x,y,y'} p_x |x\rangle\langle x| \otimes \mathbf{E}_y \boldsymbol{\rho}_x \mathbf{E}_{y'} \otimes |y\rangle\langle y'|. \tag{5.138}$$

Поскольку энтропия фон Неймана инвариантна относительно унитарного изменения базиса, то

$$S(\boldsymbol{\rho}'_{XQY}) = S(\boldsymbol{\rho}_{XQY}) = H(X) + \sum_x p_x S(\boldsymbol{\rho}_x),$$
$$S(\boldsymbol{\rho}'_{QY}) = S(\boldsymbol{\rho}_{QY}) = S(\boldsymbol{\rho}), \tag{5.139}$$

а вычисляя частичный след уравнения (5.138) и используя (5.136), мы находим

$$\boldsymbol{\rho}'_{XY} = \sum_{x,y} p_x \operatorname{tr}\left(\mathbf{E}_y \boldsymbol{\rho}_x\right) |x\rangle\langle x| \otimes |y\rangle\langle y| = \\ = \sum_{x,y} p(x,y) |x,y\rangle\langle x,y| \;\rightarrow\; S(\boldsymbol{\rho}'_{XY}) = H(X,Y). \tag{5.140}$$

Отсюда, очевидно, следует, что

$$\boldsymbol{\rho}'_Y = \sum_y p(y) |y\rangle\langle y| \;\rightarrow\; S(\boldsymbol{\rho}'_Y) = H(Y). \tag{5.141}$$

Применим теперь свойство сильной субаддитивности в форме

$$S(\boldsymbol{\rho}'_{XQY}) + S(\boldsymbol{\rho}'_Y) \leqslant S(\boldsymbol{\rho}'_{XY}) + S(\boldsymbol{\rho}'_{QY}), \tag{5.142}$$

которая принимает вид

$$H(X) + \sum_x p_x S(\boldsymbol{\rho}_x) + H(Y) \leqslant H(X,Y) + S(\boldsymbol{\rho}) \tag{5.143}$$

или

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \leqslant S(\boldsymbol{\rho}) - \sum_x p_x S(\boldsymbol{\rho}_x) = \chi(\mathcal{E}). \tag{5.144}$$

Это и есть граница Холево.

Одним из способов рассмотрения более общих ПОЗМ является расширение системы путем присоединения к ней еще одной подсистемы Z. Тогда мы конструируем унитарное преобразование $\mathbf{U}_{QYZ}$, действующее как

$$\mathbf{U}_{QYZ} \,:\, |\varphi\rangle_Q \otimes |0\rangle_Y \otimes |0\rangle_Z = \sum_y \sqrt{\mathbf{F}_y} |\varphi\rangle_Q \otimes |y\rangle_Y \otimes |y\rangle_Z, \tag{5.145}$$

так что

$$\boldsymbol{\rho}'_{XQYZ} = \sum_{x,y,y'} p_x |x\rangle\langle x| \otimes \sqrt{\mathbf{F}_y}\boldsymbol{\rho}_x\sqrt{\mathbf{F}_{y'}} \otimes |y\rangle\langle y'| \otimes |y\rangle\langle y'|. \tag{5.146}$$

Тогда вычисление частичного следа по Z дает

$$\boldsymbol{\rho}'_{XQY} = \sum_{x,y} p_x |x\rangle\langle x| \otimes \sqrt{\mathbf{F}_y}\boldsymbol{\rho}_x\sqrt{\mathbf{F}_y} \otimes |y\rangle\langle y| \tag{5.147}$$

и

$$\rho'_{XY} = \sum_{x,y} p_x \operatorname{tr}\left(\mathbf{F}_y \rho_x\right) |x\rangle\langle x| \otimes |y\rangle\langle y| =$$
$$= \sum_{x,y} p(x,y) |x,y\rangle\langle x,y| \to S(\rho'_{XY}) = H(X,Y). \qquad (5.148)$$

Оставшаяся часть доказательства проводится так же, как и выше.

5.4.2. Улучшение различимости: метод Переса — Вутерса

Чтобы лучше познакомиться с концепцией доступной информации, рассмотрим однокубитовый пример. Алиса готовит одно из трех возможных чистых состояний:

$$|\varphi_1\rangle = |\uparrow_{\hat{n}_1}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
$$|\varphi_2\rangle = |\uparrow_{\hat{n}_2}\rangle = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \qquad (5.149)$$
$$|\varphi_3\rangle = |\uparrow_{\hat{n}_3}\rangle = \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix};$$

объект со спином-$1/2$ ориентирован в одном из трех направлений, симметрично распределенных в xz-плоскости. Каждое состояние *a priori* имеет вероятность $1/3$. Очевидно, что «сигнальные состояния» Алисы не ортогональны:

$$\langle\varphi_1|\varphi_2\rangle = \langle\varphi_1|\varphi_3\rangle = \langle\varphi_2|\varphi_3\rangle = -\frac{1}{2}. \qquad (5.150)$$

Задачей Боба является: выполняя подходящее измерение, как можно больше узнать о том, что приготовлено Алисой. Матрицей плотности ансамбля Алисы является

$$\rho = \frac{1}{3}\left(|\varphi_1\rangle\langle\varphi_1| + |\varphi_2\rangle\langle\varphi_2| + |\varphi_3\rangle\langle\varphi_3|\right) = \frac{1}{2}\mathbf{1}, \qquad (5.151)$$

энтропия которой $S(\rho) = 1$. Следовательно, граница Холево говорит нам, что взаимная информация приготовления Алисы и результата измерения Боба не может превышать одного бита.

Хотя фактически доступная информация существенно меньше допускаемого границей Холево одного бита. В этом случае ансамбль Алисы достаточно симметричен, поэтому нетрудно угадать оптимальное измерение. Боб может выбрать ПОЗМ с тремя результатами, где

$$\mathbf{F}_a = \frac{2}{3}\bigl(\mathbf{1} - |\varphi_a\rangle\langle\varphi_a|\bigr), \qquad a = 1, 2, 3; \tag{5.152}$$

мы видим, что

$$p(a|b) = \langle\varphi_b|\mathbf{F}_a|\varphi_b\rangle = \begin{cases} 0, & a = b, \\ \frac{1}{2}, & a \neq b. \end{cases} \tag{5.153}$$

Следовательно, результат измерения a *исключает* возможность того, что Алиса приготовила a, но оставляет *a posteriori* равными вероятности ($p = 1/2$) двух других состояний. Полученная Бобом информация равна

$$I = H(X) - H(X|Y) = \log_2 3 - 1 = 0{,}58496. \tag{5.154}$$

Чтобы показать, что это измерение действительно оптимально, мы можем сослаться на вариант теоремы Дэвиса, которая гарантирует, что оптимальная ПОЗМ может быть выбрана с тремя $\mathbf{F}_a$, образующими, как и три состояния входящего ансамбля, семейство симметрии третьего порядка. Этот результат серьезно ограничивает возможные ПОЗМ, так что можно проверить с помощью явных вычислений, что (5.152) оптимальна. Таким образом, мы нашли, что доступная информация в ансамбле $\mathcal{E} = \{|\varphi_a\rangle, p_a = 1/3\}$ равна

$$\mathrm{Acc}(\mathcal{E}) = \log_2 \frac{3}{2} = 0{,}58496\ldots. \tag{5.155}$$

Граница Холево не достигается.

Допустим теперь, что у Алисы достаточно много денег и она может позволить себе послать Бобу два кубита, каждый из которых снова извлечен из ансамбля $\mathcal{E}$. Ее естественным решением будет приготовить для этого одно из *девяти* состояний

$$|\varphi_a\rangle|\varphi_b\rangle, \qquad a, b = 1, 2, 3, \tag{5.156}$$

с вероятностью $p_{ab} = 1/9$ каждое. Тогда наилучшая стратегия Боба, дающая, как и раньше, взаимную информацию 0,58496 битов на один кубит, состоит в выполнении ПОЗМ (5.152) на каждом из двух кубитов.

Но Алиса и Боб намерены поступить лучше. После обсуждения проблемы с А. Пересом и В. Вутерсом они выбирают другую стратегию. Алиса приготовит одно из *трех* двухкубитовых состояний

$$|\Phi_a\rangle = |\varphi_a\rangle|\varphi_a\rangle, \qquad a = 1, 2, 3, \tag{5.157}$$

каждое из которых появляется с априорной вероятностью $p_a = 1/3$. Рассматриваемый как один кубит, выбор Алисы управляется ансамблем $\mathcal{E}$, но теперь между ее двумя кубитами имеется (классическая) корреляция – оба они приготовлены одним способом.

Три вектора $|\Phi_a\rangle$ линейно независимы и, следовательно, образуют линейную оболочку трехмерного подпространства четырехмерного гильбертова пространства двух кубитов. В домашнем упражнении вы покажете, что матрица плотности

$$\boldsymbol{\rho} = \frac{1}{3}\sum_{a=1}^{3} |\Phi_a\rangle\langle\Phi_a| \tag{5.158}$$

имеет ненулевые собственные значения $1/2$, $1/4$ и $1/4$, так что

$$S(\boldsymbol{\rho}) = -\frac{1}{2}\log\frac{1}{2} - 2\left(\frac{1}{4}\log\frac{1}{4}\right) = \frac{3}{2}. \tag{5.159}$$

Граница Холево требует, чтобы доступная информация на *один кубит* была меньше, чем $3/4$ бита. По крайней мере это согласуется с тем, что можно превзойти ее значение 0,58496 на один кубит, достигнутое в методе девяти состояний.

На первый взгляд, может показаться, что Алиса не в состоянии передать такое количество классической информации Бобу, если она решает послать одно из всего лишь трех, вместо девяти, возможных состояний. Однако после некоторых размышлений этот вывод становится не очевидным. Действительно, Алиса имеет меньший выбор сигналов, но эти сигналы *более различимы*; вместо (5.150) мы имеем

$$\langle\Phi_a|\Phi_b\rangle = \frac{1}{4}, \qquad a \neq b. \tag{5.160}$$

Бобу следует использовать эту улучшенную различимость при выборе своего измерения. В частности, он найдет более выгодным выполнить *коллективное* измерение двух кубитов вместо измерения их по одному.

Теперь уже не очевидно, каким будет оптимальное измерение Боба. Но он может привлечь общую процедуру, которая, хотя и не обязательно оптимальна, но по крайней мере обычно достаточно хороша. Назовем

построенную с помощью этой процедуры ПОЗМ «достаточно хорошим измерением» (или ДХИ).

Рассмотрим некоторый набор векторов $|\tilde{\Phi}_a\rangle$, которые не предполагаются ортогональными или нормированными. Мы хотим придумать ПОЗМ, которая может достаточно хорошо различать эти векторы. Прежде всего, построим

$$\mathbf{G} = \sum_a |\tilde{\Phi}_a\rangle\langle\tilde{\Phi}_a|. \tag{5.161}$$

Это положительный оператор в подпространстве, натянутом на векторы $|\tilde{\Phi}_a\rangle$. Следовательно, в этом подпространстве он имеет обратный оператор $\mathbf{G}^{-1}$, а обратный оператор имеет положительный квадратный корень $\mathbf{G}^{-1/2}$. Теперь мы определяем

$$\mathbf{F}_a = \mathbf{G}^{-1/2}|\tilde{\Phi}_a\rangle\langle\tilde{\Phi}_a|\mathbf{G}^{-1/2} \tag{5.162}$$

и видим, что в линейной оболочке векторов $|\tilde{\Phi}_a\rangle$

$$\sum_a \mathbf{F}_a = \mathbf{G}^{-1/2}\left(\sum_a |\tilde{\Phi}_a\rangle\langle\tilde{\Phi}_a|\right)\mathbf{G}^{-1/2} = \\ = \mathbf{G}^{-1/2}\mathbf{G}\mathbf{G}^{-1/2} = \mathbf{1}. \tag{5.163}$$

Если необходимо, мы можем присоединить к этим $\mathbf{F}_a$ еще один положительный оператор, проектор $\mathbf{F}_0$ на ортогональное дополнение рассматриваемого подпространства и таким образом построить ПОЗМ. Она представляет собой ДХИ, связанное с данным набором векторов $|\tilde{\Phi}_a\rangle$.

В частном случае, когда векторы $|\tilde{\Phi}_a\rangle$ ортогональны

$$|\tilde{\Phi}_a\rangle = \sqrt{\lambda_a}|\Phi_a\rangle \tag{5.164}$$

(где $|\Phi_a\rangle$ ортонормированы), мы имеем

$$\mathbf{F}_a = \sum_{a,b,c} \left(|\Phi_b\rangle\lambda_b^{-1/2}\langle\Phi_b|\right)\left(\lambda_a|\Phi_a\rangle\langle\Phi_a|\right)\left(|\Phi_c\rangle\lambda_c^{-1/2}\langle\Phi_c|\right) = \\ = |\Phi_a\rangle\langle\Phi_a|, \tag{5.165}$$

то есть идеально различающее векторы $|\Phi_a\rangle$ ортогональное и, следовательно, оптимальное измерение. Если векторы $|\Phi_a\rangle$ линейно независимы, но

не ортогональны, то ДХИ снова является ортогональным измерением (поскольку n одномерных операторов в n-мерном пространстве могут образовать ПОЗМ, если только они взаимно ортогональны), но в этом случае измерение может оказаться не оптимальным.

В домашней работе вы построите ДХИ для векторов $|\Phi_a\rangle$ из (5.157) и покажете, что

$$p(a|a) = \langle\Phi_a|\mathbf{F}_a|\Phi_a\rangle = \frac{1}{3}\left(1+\frac{1}{\sqrt{2}}\right)^2 = 0{,}971405,$$

$$p(b|a) = \langle\Phi_a|\mathbf{F}_b|\Phi_a\rangle = \frac{1}{6}\left(1-\frac{1}{\sqrt{2}}\right)^2 = 0{,}0142977 \qquad (5.166)$$

(при $b \neq a$). Отсюда следует, что условная энтропия входа

$$H(X|Y) = 0{,}215893, \qquad (5.167)$$

а поскольку $H(X) = \log_2 3 = 1{,}58496$, то приобретаемая информация

$$I = H(X) - H(X|Y) = 1{,}36907, \qquad (5.168)$$

взаимная информация равна 0,684535 битов на один кубит. Таким образом, улучшенная различимость сигналов Алисы действительно оправдала себя — мы превзошли 0,58496 битов, которые можно было извлечь из одного кубита. Мы все еще не достигли границы Холево ($I < 1{,}5$ в этом случае), хотя и подошли к ней несколько ближе, чем раньше.

Этот пример, впервые описанный Пересом и Вутерсом, преподносит несколько полезных уроков. Во-первых, Алиса в состоянии послать Бобу большее количество информации, «сократив» свой набор кодовых слов. Ей лучше сделать выбор из меньшего количества более различимых сигналов, чем из большего количества менее различимых сигналов. Алфавит из трех букв кодирует больше, чем алфавит из девяти букв.

Во-вторых, Боб способен считать больше информации, если он выполняет коллективное измерение, вместо измерения каждого кубита по отдельности. Его оптимальное ортогональное измерение проецирует сигнал Алисы на базис *запутанных* состояний.

Описанное здесь ДХИ «оптимально» в том смысле, что оно дает наибольшее приобретение информации по сравнению с любым *известным* измерением. Скорее всего это действительно максимальное значение I, которое может быть достигнуто при любом измерении, но я не доказал этого.

5.4.3. Достижимость границы Холево: чистые состояния

Усвоив эти уроки, мы можем показать, что для заданного ансамбля чистых состояний можно построить n-буквенные кодовые слова, которые асимптотически достигают доступной информации $S(\rho)$ на одну букву.

Мы должны выбрать код, ансамбль кодовых слов, которые может приготовить Алиса, и «декодирующую наблюдаемую» — ПОЗМ, которую будет использовать Боб, пытаясь различить кодовые слова. Наша задача состоит в том, чтобы показать, что Алиса может выбрать $2^{n(S-\delta)}$ таких кодовых слов, что с пренебрежимо малой вероятностью ошибки при $n \to \infty$ Боб может определить, какое из них было послано. Мы не будем вникать во все детали доказательства, а удовольствуемся пониманием того, почему этот результат весьма правдоподобен.

Конечно, главной идеей является привлечение случайного кодирования. Алиса выбирает произведение сигнальных состояний

$$|\varphi_{x_1}\rangle|\varphi_{x_2}\rangle\cdots|\varphi_{x_n}\rangle, \tag{5.169}$$

случайным образом извлекая каждую букву из ансамбля $\mathcal{E} = \{|\varphi_x\rangle, p_x\}$. Как мы видели, для типичного кода каждое типичное кодовое слово имеет большое перекрытие с типичным подпространством $\Lambda^{(n)}$, размерность которого $\dim\Lambda^{(n)} > 2^{n[S(\rho)-\delta]}$. Более того, для типичного кода управляющий каждой буквой частный ансамбль близок к $\mathcal{E}$.

Поскольку при больших n типичное подпространство очень велико, Алиса может выбрать много кодовых слов, тем не менее оставаясь уверенной в том, что характерное перекрытие двух типичных кодовых слов очень мало. С эвристической точки зрения типичные кодовые слова случайным образом распределены в типичном подпространстве, а в среднем два случайных единичных вектора в пространстве размерности D имеют перекрытие $1/D$. Следовательно, если $|u\rangle$ и $|w\rangle$ — два кодовых слова, то

$$\langle|\langle u|w\rangle|^2\rangle_\Lambda < 2^{-n(S-\delta)}. \tag{5.170}$$

Здесь $\langle\cdot\rangle_\Lambda$ обозначает среднее по случайным типичным кодовым словам.

Вы можете убедиться в том, что типичные кодовые слова действительно однородно распределены в типичном подпространстве, как видно из дальнейшего: усредненное по ансамблю перекрытие случайных кодовых слов $|\varphi_{x_1}\rangle\ldots|\varphi_{x_n}\rangle$ и $|\varphi_{y_1}\rangle\ldots|\varphi_{y_n}\rangle$ равно

$$\begin{aligned} &= \sum p_{x_1}\cdots p_{x_n}p_{y_1}\cdots p_{y_n}\left(|\langle\varphi_{x_1}|\varphi_{y_1}\rangle|^2\cdots|\langle\varphi_{x_n}|\varphi_{y_n}\rangle|^2\right) = \\ &= \mathrm{tr}\left(\rho\otimes\cdots\otimes\rho\right)^2. \end{aligned} \tag{5.171}$$

Теперь предположим, что мы ограничили след типичным подпространством $\Lambda^{(n)}$; это пространство имеет размерность $\dim\Lambda^{(n)} < 2^{n(S+\delta)}$, а собственные значения сужения оператора $\boldsymbol{\rho}^{(n)} = \boldsymbol{\rho} \otimes \cdots \otimes \boldsymbol{\rho}$ в подпространство $\Lambda^{(n)}$ удовлетворяют неравенству $\lambda < 2^{-n(S-\delta)}$. Следовательно:

$$\langle |\langle u|w\rangle|^2 \rangle_\Lambda = \mathrm{tr}_\Lambda \left[\boldsymbol{\rho}^{(n)}\right]^2 < 2^{n(S+\delta)} \left[2^{-n(S-\delta)}\right]^2 = 2^{-n(S-3\delta)}, \qquad (5.172)$$

где tr_Λ обозначает след по типичному подпространству.

Теперь предположим, что выбрано $2^{n(S-\delta)}$ случайных кодовых слов $\{|u_i\rangle\}$. Тогда если $|u_j\rangle$ – произвольное фиксированное кодовое слово, то

$$\sum_{i \neq j} \langle |\langle u_i|u_j\rangle|^2 \rangle < 2^{n(S-\delta)} 2^{-n(S-\delta')} = 2^{-n(\delta-\delta')} + \varepsilon; \qquad (5.173)$$

здесь суммирование ведется по всем кодовым словам, а усреднение больше не ограничивается типичными кодовыми словами – ε в правой части возникает от атипичного случая. Теперь при любом фиксированном δ и при достаточно большом n мы можем выбрать δ' и ε настолько малыми, насколько нам это нужно; таким образом, при усреднении по кодам и кодовым словам внутри кода последние становятся хорошо различимыми при $n \to \infty$.

Теперь призовем на помощь несколько стандартных «шеннонизмов»: так как уравнение (5.173) справедливо в среднем по кодам, то оно справедливо и для некоторого частного кода. [Более того, поскольку почти все коды обладают тем свойством, что частный (маргинальный) ансамбль каждой буквы близок к $\mathcal{E}$, то существует код с таким свойством, удовлетворяющий (5.173).] Теперь уравнение (5.173) справедливо, если мы усредняем по частному кодовому слову $|u_j\rangle$. Но, отбрасывая не больше половины кодовых слов, можно быть уверенными в том, что любое и каждое кодовое слово хорошо отличимо от всех остальных.

Итак, Алиса может выбрать $2^{n(S-\delta)}$ хорошо различимых кодовых слов, которые становятся взаимно ортогональными в пределе $n \to \infty$. При конечном n Боб может выполнить ДХИ, которое стремится к оптимальному ортогональному измерению при $n \to \infty$. Следовательно, отнесенная к одной букве доступная информация

$$\frac{1}{n}\mathrm{Acc}(\tilde{\mathcal{E}}^{(n)}) = S(\boldsymbol{\rho}) - \delta \qquad (5.174)$$

достижима, где $\tilde{\mathcal{E}}^{(n)}$ обозначает ансамбль n-буквенных кодовых слов Алисы.

Конечно, при любом конечном n ПОЗМ Боба будет представлять собой сложное коллективное измерение, выполняемое на всех n буквах. Чтобы дать настоящее доказательство достижимости, необходимо тщательно проанализировать ПОЗМ и пределы вероятности ее ошибки. Это было выполнено Хаусладеном и др[1]. Приведенное здесь доказательство на пальцах, по крайней мере, показывает, почему их вывод не удивителен.

Из границы Холево и субаддитивности энтропии следует также, что отнесенная к одной букве доступная информация асимптотически не может превзойти $S(\boldsymbol{\rho})$. Граница Холево утверждает, что

$$\mathrm{Acc}(\tilde{\mathcal{E}}^{(n)}) \leqslant S(\tilde{\boldsymbol{\rho}}^{(n)}), \tag{5.175}$$

где $\tilde{\boldsymbol{\rho}}^{(n)}$ обозначает матрицу плотности кодовых слов, а субаддитивность энтропии предполагает, что

$$S(\tilde{\boldsymbol{\rho}}^{(n)}) \leqslant \sum_{i=1}^{n} S(\tilde{\boldsymbol{\rho}}_i), \tag{5.176}$$

где $\tilde{\boldsymbol{\rho}}_i$ — приведенная матрица плотности i-ой буквы. А так как каждая $\tilde{\boldsymbol{\rho}}_i$ асимптотически стремится к $\boldsymbol{\rho}$, то

$$\lim_{n\to\infty} \frac{1}{n}\mathrm{Acc}(\tilde{\mathcal{E}}^{(n)}) \leqslant \lim_{n\to\infty} \frac{1}{n} S(\tilde{\boldsymbol{\rho}}^{(n)}) \leqslant S(\boldsymbol{\rho}). \tag{5.177}$$

Чтобы получить это ограничение, мы не делали никаких предположений относительно кода, за исключением того, что частный ансамбль каждой буквы асимптотически стремится к $\mathcal{E}$. В частности, это ограничение справедливо, даже если кодовые слова являются не сепарабельными, а запутанными состояниями. Таким образом, мы показали, что $S(\boldsymbol{\rho})$ является оптимальной доступной информацией на одну букву.

Можно определить разновидность емкости канала связи, связанной с конкретным алфавитом чистых квантовых состояний, «емкость для заданного алфавита». Предположим, что Алиса обеспечена источником квантовых состояний. Она может создать любое из состояний $|\varphi_x\rangle$, но выбор априорных вероятностей этих состояний зависит от нее. Емкость для заданного алфавита C_{f_a} представляет собой максимальную доступную информацию на одну букву, которой она может достичь при наилучшем возможном

[1] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland and W. K. Wooters. Classical information capacity of a quantum channel. *Phys. Rev.* **A 54** (1996) 1869–1876. [См. также А. С. Холево. *Введение в квантовую теорию информации.* — МЦНМО, М.: 2002. *Прим. ред.*]

распределении $\{p_x\}$. Мы нашли, что

$$C_{f_a} = \max_{\{p_x\}} S(\boldsymbol{\rho}). \tag{5.178}$$

C_{f_a} представляет собой оптимальное количество классических битов, которое можно (асимптотически) закодировать в одной букве данного конкретного алфавита квантовых состояний из источника.

5.4.4. Достижимость границы Холево: смешанные состояния

Теперь мы хотели бы распространить предыдущие рассуждения на более общий случай. Будем рассматривать n-буквенные сообщения, в которых частным ансамблем для каждой буквы является ансамбль *смешанных* состояний

$$\mathcal{E} = \{\boldsymbol{\rho}_x, p_x\}. \tag{5.179}$$

Мы хотим показать, что в расчете на одну букву (асимптотически при $n \to \infty$) можно переслать $\chi(\mathcal{E})$ битов классической информации. Вновь нашей задачей является: (1) конкретизировать код, которым могут пользоваться Алиса и Боб, ансамбль которого (по крайней мере асимптотически) буква за буквой создает ансамбль $\mathcal{E}$; (2) конкретизировать декодирующую наблюдаемую Боба, ПОЗМ, которую он будет использовать, пытаясь различить кодовые слова; (3) показать, что при $n \to \infty$ вероятность ошибки Боба стремится к нулю. Как и при обсуждении случая чистых состояний, я не буду здесь представлять полное доказательство (см. Холево[1], а также Шумахер и Вестморленд[2]). Вместо этого я предложу вашему вниманию аргументы (даже с большим, чем ранее, количеством объяснений на пальцах, если такое возможно), показывающие, что их вывод разумен.

Как обычно, мы будем демонстрировать достижимость с помощью метода случайного кодирования. Алиса выбирает кодовые слова из смешанных состояний, каждая буква которых извлекается из ансамбля $\mathcal{E}$. То есть кодовое слово

$$\boldsymbol{\rho}_{x_1} \otimes \boldsymbol{\rho}_{x_2} \otimes \cdots \otimes \boldsymbol{\rho}_{x_n} \tag{5.180}$$

[1] A. S. Holevo. The Capacity of the Quantum Channel with General Signal States. *IEEE Trans. Inf. Theory*, **44** (1998) 269–273; quant-ph/9611023.

[2] B. Schumacher and M. D. Westmoreland. Sending Classical Information Via Noisy Quantum Channels. *Phys. Rev.* **A 56** (1997) 131–138. [На русском языке доказательство теоремы Холево–Шумахера–Вестморленда (ХШВ) можно найти в книге М. Нильсен, И. Чанг, *Квантовые вычисления и квантовая информация*, М.: Мир, 2006. – *Прим. ред.*]

выбирается с вероятностью $p_{x_1}p_{x_2}\cdots p_{x_n}$. Идея в том, что *каждое* типичное кодовое слово может рассматриваться как ансамбль чистых состояний, почти все носители которого находятся в определенном типичном подпространстве. Если перекрытия типичных подпространств различных кодовых слов малы, тогда Боб будет в состоянии выполнить ПОЗМ, которая с малой вероятностью ошибки идентифицирует характеристику типичного подпространства сообщения Алисы.

Какова размерность типичного подпространства типичного кодового слова? Если мы *усредняем* по кодовым словам, то средняя энтропия кодового слова равна

$$\langle S^{(n)}\rangle = \sum_{x_1,\ldots x_n} p_{x_1}p_{x_2}\cdots p_{x_n} S(\boldsymbol{\rho}_{x_1}\otimes\boldsymbol{\rho}_{x_2}\otimes\cdots\otimes\boldsymbol{\rho}_{x_n}). \tag{5.181}$$

Используя аддитивность энтропии произведения состояний и $\sum_x p_x = 1$, мы получаем

$$\langle S^{(n)}\rangle = n\sum_x p_x S(\boldsymbol{\rho}_x) \equiv n\langle S\rangle. \tag{5.182}$$

При больших n энтропия кодового слова с большой вероятностью близка к ее среднему значению, более того, велика вероятность того, что собственные значения $\boldsymbol{\rho}_{x_1}\otimes\cdots\otimes\boldsymbol{\rho}_{x_n}$ близки к $2^{-n\langle S\rangle}$. Другими словами, типичное кодовое слово $\boldsymbol{\rho}_{x_1}\otimes\cdots\otimes\boldsymbol{\rho}_{x_n}$ имеет носитель в типичном подпространстве размерности $2^{n\langle S\rangle}$.

Это утверждение является близким аналогом высказывания (ключевого в доказательстве теоремы Шеннона о кодировании для канала связи с шумом) о том, что если через классический канал связи с шумом послано типичное сообщение, то количество типичных сообщений, которые могут быть получены, равно $2^{nH(Y|X)}$.

Далее доказательство следует знакомым путем. Для каждого типичного сообщения $x_1x_2\ldots x_n$ Боб может построить «декодирующее подпространство» размерности $2^{n(\langle S\rangle+\delta)}$ с уверенностью, что почти все носители сообщения Алисы принадлежат этому подпространству. Его ПОЗМ будет предназначена для определения, в каком декодирующем подпространстве находится сообщение Алисы. Ошибки декодирования будут маловероятны, если малы перекрытия типичных декодирующих подпространств.

Несмотря на то, что на самом деле Боба интересует только оценка декодирующего подпространства (и, следовательно, $x_1x_2\ldots x_n$), предположим, что он выполняет полное ДХИ, определяемое всеми векторами, образующими линейную оболочку всех типичных подпространств кодовых

слов Алисы. (При больших n это ДХИ будет стремиться к ортогональному измерению, пока число кодовых слов не слишком велико.) Он получает конкретный результат, который, вероятно, находится в типичном подпространстве размерности $2^{nS(\rho)}$, определяемом источником $\rho \otimes \rho \otimes \cdots \otimes \rho$. Более того, этот результат, вероятно, находится в декодирующем подпространстве сообщения, которое Алиса на самом деле послала. Поскольку результаты измерения Боба однородно распределены в пространстве размерности 2^{nS}, а ансамбль чистых состояний, определяемый частным декодирующим подпространством, имеет размерность $2^{n(\langle S\rangle+\delta)}$, то среднее перекрытие вектора, определенного результатом Боба, с типичным декодирующим подпространством равно:

$$\frac{2^{n(\langle S\rangle+\delta)}}{2^{nS}} = 2^{-n(S-\langle S\rangle-\delta)} = 2^{-n(\chi-\delta)}. \tag{5.183}$$

Если Алиса выбирает 2^{nR} кодовых слов, то средняя вероятность ошибки декодирования будет

$$2^{nR}2^{-n(\chi-\delta)} = 2^{-n(\chi-R-\delta)}. \tag{5.184}$$

Мы можем выбрать R любым, меньшим χ, тогда эта вероятность ошибки будет очень мала при $n \to \infty$.

Эти доводы показывают, что усредненная по случайным кодам и кодовым словам вероятность ошибки мала. Как обычно, мы выбираем конкретный код и отбрасываем некоторые кодовые слова, чтобы получить малую вероятность ошибки для каждого кодового слова. Более того, конкретный код может быть выбран типичным, так что частный ансамбль каждого кодового слова стремится к $\mathcal{E}$ при $n \to \infty$. Мы приходим к выводу, что отнесенная к одной букве доступная информация χ *асимптотически достижима*.

По своей структуре это доказательство близко к аналогичному доказательству соответствующей классической теоремы о кодировании. В частности, величина χ здесь играет такую же роль, как I в теореме Шеннона. В то время как 2^{-nI} является вероятностью того, что конкретная типичная последовательность лежит в определенной сфере декодирования, $2^{-n\chi}$ представляет собой перекрытие конкретного типичного состояния с определенным декодирующим подпространством.

5.4.5. Емкость канала связи

Комбинируя границу Холево с выводом о том, что достижимы χ битов на одну букву, можно получить выражение для *классической* емкости

квантового канала связи. (Но с предупреждением: мы уверены, что эту «емкость» нельзя превысить, если только мы отказываемся от использования запутанных кодовых слов.)

Алиса готовит n-буквенные сообщения и посылает их Бобу через квантовый канал связи с шумом, описываемый супероператором \$. Предположим, что упомянутый супероператор \$ действует на каждую букву независимо (квантовый канал *без памяти*). Боб выполняет ПОЗМ, которая оптимизирует получаемую им информацию относительно того, что приготовила Алиса.

Фактически окажется, что лучше всего Алиса готовит сообщения из чистых состояний (это следует из субаддитивности энтропии). Если конкретная буква приготовлена как чистое состояние $|\varphi_x\rangle$, то Боб получит

$$|\varphi_x\rangle\langle\varphi_x| \to \$(|\varphi_x\rangle\langle\varphi_x|) \equiv \boldsymbol{\rho}_x. \tag{5.185}$$

А если Алиса посылает чистое состояние $|\varphi_{x_1}\rangle\cdots|\varphi_{x_n}\rangle$, то Боб получит смешанное состояние $\boldsymbol{\rho}_{x_1}\otimes\cdots\otimes\boldsymbol{\rho}_{x_n}$. Таким образом, ансамбль кодовых слов Алисы определяет ансамбль смешанных состояний $\tilde{\mathcal{E}}^{(n)}$, получаемых Бобом. Следовательно, оптимальное количество получаемой Бобом информации по определению равно $\mathrm{Acc}(\tilde{\mathcal{E}}^{(n)})$, что удовлетворяет границе Холево:

$$\mathrm{Acc}(\tilde{\mathcal{E}}^{(n)}) \leqslant \chi(\tilde{\mathcal{E}}^{(n)}). \tag{5.186}$$

Теперь ансамблем Боба является

$$\{\boldsymbol{\rho}_{x_1}\otimes\cdots\otimes\boldsymbol{\rho}_{x_n},\, p(x_1,x_2,\ldots,x_n)\}, \tag{5.187}$$

где $p(x_1,x_2,\ldots,x_n)$ совершенно произвольное распределение вероятностей кодовых слов Алисы. Вычислим χ для этого ансамбля. Заметим, что

$$\begin{aligned}
&\sum_{x_1,\ldots,x_n} p(x_1,x_2,\ldots,x_n)S(\boldsymbol{\rho}_{x_1}\otimes\cdots\otimes\boldsymbol{\rho}_{x_n}) = \\
&= \sum_{x_1,\ldots,x_n} p(x_1,x_2,\ldots,x_n)\Big[S(\boldsymbol{\rho}_{x_1}) + S(\boldsymbol{\rho}_{x_2}) + \ldots + S(\boldsymbol{\rho}_{x_n})\Big] = \\
&= \sum_{x_1} p_1(x_1)S(\boldsymbol{\rho}_{x_1}) + \sum_{x_2} p_2(x_2)S(\boldsymbol{\rho}_{x_2}) + \ldots + \\
&\qquad + \sum_{x_n} p_n(x_n)S(\boldsymbol{\rho}_{x_n}).
\end{aligned} \tag{5.188}$$

где, например, $p_1(x_1) = \sum_{x_2,\ldots,x_n} p(x_1, x_2, \ldots, x_n)$ — частное распределение вероятностей для первой буквы. Более того, из субаддитивности энтропии мы имеем:

$$S(\tilde{\boldsymbol{\rho}}^{(n)}) \leqslant S(\tilde{\boldsymbol{\rho}}_1) + S(\tilde{\boldsymbol{\rho}}_2) + \ldots + S(\tilde{\boldsymbol{\rho}}_n), \tag{5.189}$$

где $\tilde{\boldsymbol{\rho}}_i$ — приведенная матрица плотности i-ой буквы. Комбинируя (5.188) и (5.189), мы находим, что

$$\chi(\tilde{\mathcal{E}}^{(n)}) \leqslant \chi(\tilde{\mathcal{E}}_1) + \chi(\tilde{\mathcal{E}}_2) + \ldots + \chi(\tilde{\mathcal{E}}_n), \tag{5.190}$$

где $\tilde{\mathcal{E}}_i$ — маргинальный ансамбль, управляющий i-ой буквой, полученной Бобом. Неравенство (5.190) применимо к любому ансамблю факторизуемых состояний.

Теперь для канала, описываемого супероператором $, определим *емкость канала* по отношению к факторизуемым состояниям

$$C(\$) = \max_{\mathcal{E}} \chi(\$(\mathcal{E})). \tag{5.191}$$

Следовательно, для каждого слагаемого в (5.190) $\chi(\tilde{\mathcal{E}}_i) \leqslant C$ и мы получаем

$$\chi(\tilde{\mathcal{E}}^{(n)}) \leqslant nC, \tag{5.192}$$

где $\tilde{\mathcal{E}}^{(n)}$ — произвольный ансамбль факторизуемых состояний. В частности, из границы Холево мы приходим к заключению, что количество получаемой Бобом информации ограничено сверху величиной nC. Но мы видели, что для любого $\mathcal{E}$ можно асимптотически достичь $\chi(\$(\mathcal{E}))$ битов на одну букву сообщения при правильном выборе кода и декодирующей наблюдаемой. Следовательно, C представляет собой оптимальное количество битов на одну букву, которое может быть передано через канал с шумом с исчезающе малой вероятностью ошибки *при условии*, что сообщения, которые готовит Алиса, представляют собой факторизуемые состояния.

Мы оставили открытой возможность того, что емкость относительно факторизуемых состояний $C(\$)$ может быть превышена, если позволить Алисе готовить *запутанные* состояния из ее n букв. Неизвестно (на январь 1998 г.), существуют ли квантовые каналы, более высокая пропускная способность которых может быть достигнута при использовании запутанных сообщений. Это один из многих интересных открытых вопросов теории квантовой информации[1].

[1] Обзор некоторых результатов относительно пропускной способности квантового канала при использовании запутанных сообщений см. в книге А. С. Холево, *Введение в квантовую теорию информации*, МЦНМО, М., (2002). — *Прим. ред.*

5.5. Плотность запутывания

Прежде чем завершить наш обзор теории квантовой информации, обратимся к еще одной теме, в которой центральную роль играет энтропия фон Неймана: количественное определение запутывания.

Рассмотрим два бинарных чистых состояния. Одно из них — *максимально* запутанное состояние двух кубитов

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big). \tag{5.193}$$

Другое — *частично* запутанное состояние двух *кутритов*

$$|\Psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{2}|11\rangle + \frac{1}{2}|22\rangle. \tag{5.194}$$

Какое из них более запутано?

Непосредственно не видно, что ответ на этот вопрос имеет глубокий смысл. Почему можно найти однозначный способ упорядочения всего континуума бинарных состояний в соответствии со степенью их запутывания? Можем ли мы сравнивать пару кубитов с парой кутритов иным способом, нежели яблоко с апельсином?

Определяющим свойством запутывания является то, что оно не может быть создано локальными операциями. В частности, если Алиса и Боб делят бинарное чистое состояние, то они не могут увеличить его число Шмидта никакими локальными операциями — никаким унитарным преобразованием или ПОЗМ, выполняемыми Алисой или Бобом, даже если они обмениваются классическими сообщениями о своих действиях и результатах измерений. Таким образом, число, используемое для количественного определения запутывания, должно обладать тем свойством, что локальные операции его не увеличивают. Очевидным кандидатом является число Шмидта, но после некоторого размышления оно уже не кажется достаточно удовлетворительным. Рассмотрим состояние

$$|\Psi_\varepsilon\rangle = \sqrt{1-2|\varepsilon|^2}|00\rangle + \varepsilon|11\rangle + \varepsilon|22\rangle, \tag{5.195}$$

имеющее число Шмидта, равное трем при любом $|\varepsilon| > 0$. Можем ли мы на самом деле сказать, что $|\Psi_\varepsilon\rangle$ «более запутано», чем $|\phi^+\rangle$? В конце концов, запутывание может рассматриваться как ресурс — мы можем планировать использовать его, например, для телепортации. Кажется очевидным, что $|\Psi_\varepsilon\rangle$ (при $|\varepsilon| \ll 1$) менее ценный, чем $|\phi^+\rangle$, ресурс.

Тем не менее оказывается, что существует естественный и разумный способ количественного описания запутывания *любого* бинарного чистого состояния. Чтобы сравнить два состояния, выполним локальные операции, чтобы обменять их запутывание на тот «валютный эталон», который можно сравнивать непосредственно. Таким «валютным эталоном» является максимально запутанное состояние.

Точное утверждение о взаимозаменяемости (посредством локальных операций) различных форм запутывания неизбежно будет *асимптотическим*. То есть для того чтобы дать точное количественное описание запутывания конкретного бинарного чистого состояния $|\psi\rangle_{AB}$, представим, что мы хотим приготовить n идентичных копий этого состояния. В нашем распоряжении имеется большой запас максимально запутанных *пар Белла*, поделенных между Алисой и Бобом. Они должны использовать k пар Белла $\left(|\phi\rangle_{AB}\right)^k$ и с помощью локальных операций и классических средств связи приготовить n копий требуемого состояния $\left(|\psi\rangle_{AB}\right)^n$. Чему равно минимальное число $k_{\min}$ пар Белла, необходимое для решения этой задачи?

А теперь предположим, что n копий $|\psi\rangle_{AB}$ уже приготовлены. Алиса и Боб должны выполнить локальные операции, которые преобразуют запутывание $\left(|\psi_{AB}\rangle\right)^n$ назад к стандартной форме; то есть они должны извлечь k' пар Белла $\left(|\phi\rangle_{AB}\right)^{k'}$. Чему равно максимальное количество $k'_{\max}$ пар Белла, которые могут быть выделены (локальным образом) из $\left(|\psi\rangle_{AB}\right)^n$?

Поскольку невозможность порождения запутывания с помощью локальных операций является непушимым принципом, то несомненно, что

$$k'_{\max} \leqslant k_{\min}. \tag{5.196}$$

Однако можно показать, что

$$\lim_{n\to\infty} \frac{k_{\min}}{n} = \lim_{n\to\infty} \frac{k'_{\max}}{n} \equiv E(|\psi\rangle_{AB}). \tag{5.197}$$

В этом смысле локальное преобразование n копий бинарного чистого состояния $|\psi\rangle_{AB}$ в k' запутанных пар является асимптотически *обратимым* процессом. Так как n копий $|\psi\rangle_{AB}$ можно заменить k парами Белла и наоборот, то отсюда следует, что $\frac{k}{n}$ пар Белла однозначно характеризуют степень запутывания, переносимого состоянием $|\psi\rangle_{AB}$. Будем называть отношение k/n (в пределе $n \to \infty$) *запутыванием* E состояния $|\psi\rangle_{AB}$. Величина E измеряет, какие требуются затраты (в парах Белла), чтобы создать $|\psi\rangle_{AB}$,

а также значение $|\psi\rangle_{AB}$ как ресурса (то есть количество кубитов, которые можно телепортировать с помощью $|\psi\rangle_{AB}$).

Чему равно E для данного конкретного чистого состояния $|\psi\rangle_{AB}$? Можете ли вы угадать ответ? Это

$$E = S(\rho_A) = S(\rho_B); \tag{5.198}$$

запутывание является энтропией фон Неймана матрицы плотности Алисы ρ_A (или матрицы плотности Боба ρ_B). Это, очевидно, правильный ответ в том случае, когда $|\psi\rangle_{AB}$ является произведением k пар Белла. В этом случае ρ_A (или ρ_B) равна $\frac{1}{2}\mathbf{1}$ для каждого имеющегося в распоряжении Алисы кубита

$$\rho_A = \frac{1}{2}\mathbf{1} \otimes \frac{1}{2}\mathbf{1} \otimes \cdots \otimes \frac{1}{2}\mathbf{1} \tag{5.199}$$

и

$$S(\rho_A) = kS\left(\frac{1}{2}\mathbf{1}\right) = k. \tag{5.200}$$

Теперь нужно понять, почему $E = S(\rho_A)$ является правильным ответом для любого бинарного чистого состояния.

Прежде всего, мы хотим показать, что если Алиса и Боб делят $k = n[S(\rho_A) + \delta]$ пар Белла, то они могут (с помощью локальных операций) с высокой точностью воспроизведения приготовить $\left(|\psi\rangle_{AB}\right)^n$. Они могут решить эту задачу, комбинируя квантовую телепортацию со сжатием Шумахера. Во-первых, локальным образом манипулируя находящейся в ее распоряжении бинарной системой AC, Алиса конструирует состояние $|\psi\rangle_{AC}$ (n его копий). Таким образом, состояние системы C можно рассматривать как чистое состояние, извлеченное из ансамбля, описываемого матрицей плотности ρ_C, где $S(\rho_C) = S(\rho_A)$. Затем Алиса выполняет сжатие Шумахера над ее n копиями C, сохраняя хорошую точность воспроизведения, несмотря на то, что типичные состояния из $(\mathcal{H}_C)^n$ сжимаются в пространство $\tilde{\mathcal{H}}_C^{(n)}$ с

$$\dim \tilde{\mathcal{H}}_C^{(n)} = 2^{n[S(\rho_A)+\delta]}. \tag{5.201}$$

Теперь Алиса и Боб могут использовать $n[S(\rho_A) + \delta]$ поделенных между ними пар Белла, чтобы телепортировать сжатое состояние из пространства Алисы $(\tilde{\mathcal{H}}_C)^n$ в пространство Боба $(\tilde{\mathcal{H}}_B)^n$. Телепортация, которая в принципе имеет идеальную точность воспроизведения, требует только локальных операций и классических средств связи, если Алиса и Боб делят необходимое количество пар Белла. Наконец, Боб развертывает (декомпрессия

Шумахера) полученное им состояние; тогда Алиса и Боб делят $\left(|\psi\rangle_{AB}\right)^n$ (со сколь угодно высокой точностью воспроизведения при $n \to \infty$).

Предположим теперь, что Алиса и Боб приготовили состояние $\left(|\psi\rangle_{AB}\right)^n$. Так как $|\psi\rangle_{AB}$, вообще говоря, *частично* запутанное состояние, то поделенное между ними запутывание находится в разбавленном виде. Алиса и Боб хотят *концентрировать* поделенное между ними запутывание, сжимая его до минимально возможного гильбертова пространства; то есть они хотят конвертировать его в максимально запутанные пары. Мы покажем, что с высокой вероятностью успеха Алиса и Боб могут «выпарить» из $\left(|\psi\rangle_{AB}\right)^n$ как минимум

$$k' = n[S(\boldsymbol{\rho}_A) - \delta] \tag{5.202}$$

пар Белла.

Поскольку мы знаем, что Алиса и Боб не могут локальным образом порождать запутывание, то они не могут с помощью локальных операций превратить k пар Белла в $k' > k$ пар, по крайней мере не с высокими точностью воспроизведения и вероятностью успеха. Тогда отсюда следует, что $nS(\boldsymbol{\rho}_A)$ является минимальным количеством пар Белла, необходимым для создания n копий $|\psi\rangle_{AB}$, и что $nS(\boldsymbol{\rho}_A)$ — максимальное количество пар Белла, которое может быть извлечено из n копий $|\psi\rangle_{AB}$. Если бы мы могли более эффективно создавать $|\psi\rangle_{AB}$ из пар Белла или более эффективно извлекать пары Белла из $|\psi\rangle_{AB}$, тогда у Алисы и Боба был бы способ, которым они увеличили бы свой запас пар Белла с помощью локальных операций, что, как известно, невозможно. Следовательно, если мы можем найти способ выделить $k' = n[S(\boldsymbol{\rho}_A) - \delta]$ пар Белла из n копий $|\psi\rangle_{AB}$, то мы знаем, что $E = S(\boldsymbol{\rho}_A)$.

Чтобы проиллюстрировать плотность запутывания, представим, что Алиса и Боб имеют n копий частично запутанного чистого состояния двух кубитов:

$$|\psi(\theta)\rangle_{AB} = \cos\theta|00\rangle + \sin\theta|11\rangle. \tag{5.203}$$

(Подобным образом можно записать любое бинарное чистое состояние, если выбрать базис Шмидта и подходящее соглашение относительно фазы.) То есть Алиса и Боб делят состояние

$$\left(|\psi(\theta)\rangle_{AB}\right)^n = \left(\cos\theta|00\rangle + \sin\theta|11\rangle\right)^n. \tag{5.204}$$

Пусть теперь Алиса (или Боб) выполняет локальное преобразование ее (его) n кубитов. Алиса измеряет вдоль оси z *полный* спин ее n кубитов

$$\boldsymbol{\sigma}_{3,A}^{\text{total}} = \sum_{i=1}^{n} \boldsymbol{\sigma}_{3,A}^{i}. \tag{5.205}$$

Главной чертой этого измерения является его «грубость». Наблюдаемая $\sigma_{3,A}^{\text{total}}$ является сильно *вырожденной*. Алиса проецирует состояние ее n спинов на одно из больших собственных пространств этой наблюдаемой. Она не измеряет спин каждого кубита; фактически она старается не получить никакой другой информации, кроме значения $\sigma_{3,A}^{\text{total}}$, или, что эквивалентно, количества ориентированных «вверх» спинов.

Если мы разложим (5.204), то получим всего 2^n слагаемых. Среди них – $\binom{n}{m}$ слагаемых, в которых ровно m из имеющихся у Алисы кубитов имеют значение $+1$. Каждое из этих слагаемых содержит коэффициент $(\cos\theta)^{n-m}(\sin\theta)^m$. Таким образом, вероятность того, что измерение Алисы обнаружит m направленных «вверх» спинов, равна

$$P(m) = \binom{n}{m} (\cos^2\theta)^{n-m}(\sin^2\theta)^m. \tag{5.206}$$

Более того, если она получила этот результат, то ее измерение приготовило *равновзвешенную* суперпозицию всех $\binom{n}{m}$ состояний, имеющих m ориентированных «вверх» спинов. (Конечно, поскольку спины Алисы и Боба идеально скоррелированы, то если бы Боб измерил $\sigma_{3,B}^{\text{total}}$, то он получил бы точно такой же, как и Алиса, результат. Альтернативно о своем результате Алиса могла бы доложить Бобу в классическом сообщении и тем самым избавить его от хлопот выполнять измерение самому.) Независимо от результата измерения Алиса и Боб теперь делят новое состояние $|\psi'\rangle_{AB}$, в котором все ненулевые собственные значения ρ'_A (и ρ'_B) равны.

При большом n распределение вероятностей $P(m)$ в (5.206) имеет резкий пик – вероятность близка к единице, когда m/n близко к $\sin^2\theta$ и

$$\binom{n}{m} \sim \binom{n}{n\sin^2\theta} \sim 2^{nH(\sin^2\theta)}, \tag{5.207}$$

где $H(p) = -p\log p - (1-p)\log(1-p)$ – функция энтропии. То есть с вероятностью, большей чем $1-\varepsilon$, поделенное теперь между Алисой и Бобом запутанное состояние имеет число Шмидта $\binom{n}{m}$, удовлетворяющее

$$2^{n[H(\sin^2\theta)-\delta]} < \binom{n}{m} < 2^{n[H(\sin^2\theta)+\delta]}. \tag{5.208}$$

Теперь Алиса и Боб хотят превратить поделенное ими запутывание в стандартную пару Белла $|\phi^+\rangle$. Это было бы просто, если бы число Шмидта их максимально запутанного состояния оказалось степенью двойки. Тогда Алиса и Боб могли бы выполнить унитарное преобразование, которое

перевело бы 2^k-мерный носитель его/ее матрицы плотности в гильбертово пространство k кубитов, а затем они могли бы отбросить остаток их кубитов. Тогда оставленные ими k пар были бы максимально запутанными.

Конечно, $\binom{n}{m}$ не обязано быть близким к степени двух. Но если Алиса и Боб разделили множество партий из n копий частично запутанного состояния, то они могут концентрировать запутывание в каждой партии. После такой обработки ℓ партий они получат максимально запутанное состояние с числом Шмидта

$$N_{\text{Schm}} = \binom{n}{m_1}\binom{n}{m_2}\binom{n}{m_3}\cdots\binom{n}{m_\ell}, \tag{5.209}$$

где каждое m_i, как правило, близко к $n\sin^2\theta$. Для любого $\varepsilon > 0$ найдется некоторое ℓ такое, что это число Шмидта в конечном счете окажется близким к степени двух

$$2^{k_\ell} < N_{\text{Schm}} < 2^{k_\ell}(1+\varepsilon). \tag{5.210}$$

При этих условиях Алиса или Боб могут выполнить измерение, которое пытается проецировать носитель размерности $2^{k_\ell}(1+\varepsilon)$ его/ее матрицы плотности на подпространство размерности 2^{k_ℓ}, достигая цели с вероятностью $1-\varepsilon$. *Затем* они переводят носитель в гильбертово пространство k_ℓ кубитов и отбрасывают остаток их кубитов. Обычно k_ℓ близко к $n\ell H(\sin^2\theta)$, так что с близкой к единице вероятностью успеха они выделят около $H(\sin^2\theta)$ максимально запутанных пар из каждого частично запутанного состояния.

Конечно, несмотря на то, что количество ориентированных вверх спинов, которые Алиса (или Боб) находит в своем измерении, обычно близко к $n\sin^2\theta$, оно может флуктуировать около этого значения. Иногда, если повезет, они смогут выделить больше, чем $H(\sin^2\theta)$, пар Белла на одну копию $|\psi(\theta)\rangle_{AB}$. Но вероятность такого существенно лучшего результата пренебрежимо мала при $n \to \infty$.

Эти рассуждения легко распространяются на бинарные чистые состояния в более широких гильбертовых пространствах. Бинарное чистое состояние с числом Шмидта s может быть представлено в базисе Шмидта как

$$|\psi(a_1, a_2, \ldots, a_s)\rangle_{AB} = a_1|11\rangle + a_2|22\rangle + \ldots a_s|ss\rangle. \tag{5.211}$$

Тогда Алиса (или Боб) может измерить полное число векторов $|1\rangle$, полное число векторов $|2\rangle$ и так далее в имеющемся в ее (его) распоряжении состоянии $\left(|\psi\rangle_{AB}\right)^n$. Если она (он) находит m_1 векторов $|1\rangle$, m_2 векторов $|2\rangle$

и т. д., тогда ее (его) измерение готовит максимально запутанное состояние с числом Шмидта

$$N_{\text{Schm}} = \frac{n!}{(m_1)!(m_2)!\cdots(m_s)!}. \tag{5.212}$$

При больших m Алиса обычно будет находить

$$m_i \sim |a_i|^2 n \tag{5.213}$$

и, следовательно,

$$N_{\text{Schm}} \sim 2^{nH}, \tag{5.214}$$

где

$$H = -\sum_i |a_i|^2 \log |a_i|^2 = S(\rho_A). \tag{5.215}$$

Таким образом, из n копий состояния $|\psi\rangle_{AB}$ асимптотически при $n \to \infty$ может быть выделено близкое к $S(\rho_A)$ количество пар Белла.

5.5.1. Запутывание смешанного состояния

Мы нашли хорошо мотивированный и однозначный способ количественного описания запутывания бинарного чистого состояния $|\psi\rangle_{AB}$: $E = S(\rho_A)$, где

$$\rho_A = \text{tr}_B \left(|\psi\rangle_{AB\ AB}\langle\psi|\right). \tag{5.216}$$

Значительный интерес также представляет количественное описание запутывания бинарного смешанного состояния. К сожалению, запутывание смешанного состояния не так хорошо понято, как запутывание чистого состояния, и является предметом текущих исследований.

Допустим, что ρ_{AB} — поделенное Алисой и Бобом смешанное состояние и что они имеют n идентичных копий этого состояния. Предположим также, что, используя локальные операции и классические средства связи, Алиса и Боб могут приготовить $(\rho_{AB})^n$ из k пар Белла с асимптотически (при $n \to \infty$) хорошей точностью воспроизведения и высокой вероятностью успеха. Определим F — *запутывание формирования* ρ_{AB} как

$$F(\rho_{AB}) = \lim_{n\to\infty} \frac{k_{\min}}{n}. \tag{5.217}$$

Далее, предположим, что Алиса и Боб могут использовать локальные операции и классическую связь, чтобы выделить k' пар Белла из n копий ρ_{AB}.

Определим D — *запутывание выделения* ρ_{AB} как

$$D(\rho_{AB}) = \lim_{n\to\infty} \frac{k'_{\max}}{n}. \tag{5.218}$$

Для чистых состояний мы нашли $D = E = F$. Но для смешанных состояний явные общие формулы для D или F неизвестны. Поскольку запутывание не может создаваться локально, мы знаем, что $D \leqslant F$, но (по состоянию на январь 1998 г.) не известно, выполняется ли равенство $D = F$. Однако имеются сильные подозрения, что для смешанных состояний $D < F$. Чтобы приготовить смешанное состояние $(\rho_{AB})^n$ из чистого $(|\phi^+\rangle_{AB}\langle\phi^+|)^k$, мы должны пренебречь некоторой квантовой информацией. Было бы удивительно, если бы этот процесс оказался (асимптотически) обратимым.

Полезно различать два разных типа запутывания выделения. D_1 обозначает количество пар Белла, которые могут быть выделены, если разрешена только односторонняя классическая связь (например, Алиса *может посылать* сообщения Бобу, но *не может получать* сообщения от него). $D_2 = D$ обозначает запутывание выделения, если классическая связь ничем не ограничена. Известно, что для некоторых смешанных состояний $D_1 < D_2$ и, следовательно, $D_1 < F$ (тогда как для чистых состояний $D_1 = D_2 = F$).

Одной из причин интереса к запутыванию смешанных состояний (и, в частности, к D_1) является его связь с передачей квантовой информации через квантовые каналы с шумом. Если в квантовом канале связи, описываемом супероператором \$, уровень шума не *слишком* высок, то можно построить n-буквенный блоковый код такой, что квантовая информация может быть закодирована, послана через канал $(\$)^n$, декодирована и воспроизведена со сколь угодно высокой точностью при $n \to \infty$. Оптимальное количество закодированных кубитов на одну букву, которое может быть послано через канал, называется емкостью квантового канала $C(\$)$. Оказывается, что $C(\$)$ может быть связана с D_1 частного смешанного состояния, связанного с каналом, — но мы пока отложим дальнейшее обсуждение емкости квантового канала.

5.6. Резюме

Энтропия Шеннона и сжатие классических данных. *Энтропия Шеннона* ансамбля $X = \{x, p(x)\}$ равна $H(X) \equiv -\langle \log p(x)\rangle$; она количественно определяет сжимаемость классической информации. Сообщение

длиной в n букв, каждая буква которого независимо извлекается из X, может быть сжато до $H(X)$ на одну букву (но не более) и, несмотря на это, все же может быть декодировано со сколь угодно высокой точностью при $n \to \infty$.

Взаимная информация и емкость классического канала связи. *Взаимная информация* $I(X;Y) = H(X) + H(Y) - H(X,Y)$ количественно определяет, насколько коррелированы ансамбли X и Y; если мы узнаем значение y, то приобретаем (в среднем) $I(X;Y)$ битов информации об x. Емкость классического шумящего канала связи без памяти равна $C = \max_{\{p(x)\}} I(X;Y)$. Это максимальное количество битов на одну букву, которое может быть передано через канал (используя наилучший возможный код) с пренебрежимо малой вероятностью ошибки при $n \to \infty$.

Энтропия фон Неймана, информация Холево и сжатие квантовых данных. *Энтропия фон Неймана* матрицы плотности ρ равна

$$S(\rho) = -\operatorname{tr} \rho \log \rho, \tag{5.219}$$

а *информация Холево* ансамбля $\mathcal{E} = \{\rho_x, p_x\}$ квантовых состояний равна

$$\chi(\mathcal{E}) = S\left(\sum_x p_x \rho_x\right) - \sum_x p_x S(\rho_x). \tag{5.220}$$

Энтропия фон Неймана количественно определяет сжимаемость ансамбля чистых квантовых состояний. Сообщение длиной в n букв, каждая буква которого независимо извлекается из ансамбля $\{|\varphi_x\rangle, p_x\}$, может быть сжато до $S(\rho)$ кубитов на одну букву (но не более) и, несмотря на это, все же может быть декодировано со сколь угодно высокой точностью воспроизведения при $n \to \infty$. Если буквы извлекаются из ансамбля $\mathcal{E}$ смешанных квантовых состояний, то невозможно сжатие с высокой точностью воспроизведения до менее чем $\chi(\mathcal{E})$ кубитов на одну букву.

Доступная информация. *Доступная информация* ансамбля $\mathcal{E}$ квантовых состояний представляет собой максимальное количество битов информации, которое (в среднем) можно получить о приготовлении состояния с помощью наилучшего возможного измерения. Доступная информация не может превысить информацию Холево ансамбля. Можно построить такой n-буквенный код, что частный ансамбль каждой буквы будет близок к $\mathcal{E}$, а доступная информация на одну букву – к $\chi(\mathcal{E})$. Емкость квантового канала связи \$ по отношению к факторизуемым состояниям равна

$$C(\$) = \max_{\mathcal{E}} \chi(\$(\mathcal{E})). \tag{5.221}$$

Это максимальное количество классических битов на одну букву, которое может быть передано через квантовый канал с пренебрежимо малой вероятностью ошибки при $n \to \infty$ при условии, что каждое кодовое слово является тензорным произведением букв-состояний.

Плотность запутывания. *Запутывание* E бинарного чистого состояния $|\psi\rangle_{AB}$ равна $E = S(\boldsymbol{\rho}_A)$, где $\boldsymbol{\rho}_A = \mathrm{tr}_B\left(|\psi\rangle_{AB\,AB}\langle\psi|\right)$. С помощью локальных операций и классических средств связи можно приготовить n копий $|\psi\rangle_{AB}$ из nE (но не из чуть меньшего количества) пар Белла, а также можно выделить nE (но не больше) пар Белла из n копий $|\psi\rangle_{AB}$ (асимптотически при $n \to \infty$).

5.7. Упражнения

5.1. Различимость неортогональных состояний. Алиса приготовила один кубит в одном из двух (неортогональных) состояний

$$|u\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |v\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}, \tag{5.222}$$

где $0 < \theta < \pi$. Бобу известно значение θ, но он не знает, приготовила Алиса $|u\rangle$ или $|v\rangle$, и ему нужно выполнить измерение, чтобы как можно больше узнать о приготовленном Алисой состоянии.

Боб рассматривает три возможных измерения.

a) Ортогональное измерение с

$$\mathbf{E}_1 = |u\rangle\langle u|, \qquad \mathbf{E}_2 = \mathbf{1} - |u\rangle\langle u|. \tag{5.223}$$

(В этом случае, если Боб получит результат 2, то он будет знать, что Алиса должна была приготовить $|v\rangle$.)

b) ПОЗМ с тремя исходами

$$\begin{aligned} \mathbf{F}_1 &= A(\mathbf{1} - |u\rangle\langle u|), \quad \mathbf{F}_2 = A(\mathbf{1} - |v\rangle\langle v|), \\ \mathbf{F}_3 &= (1 - 2A)\mathbf{1} + A\big(|u\rangle\langle u| + |v\rangle\langle v|\big), \end{aligned} \tag{5.224}$$

где A имеет максимальное совместимое с положительностью $\mathbf{F}_3$ значение. (В этом случае Боб однозначно определит приготовленное состояние, если получит результат 1 или 2, но ничего не узнает из результата 3.)

c) Ортогональнос измерение с

$$\mathbf{E}_1 = |w\rangle\langle w|, \qquad \mathbf{E}_2 = \mathbf{1} - |w\rangle\langle w|, \tag{5.225}$$

где

$$|w\rangle = \begin{pmatrix} \cos\left[\frac{1}{2}\left(\frac{\theta}{2} + \frac{\pi}{2}\right)\right] \\ \sin\left[\frac{1}{2}\left(\frac{\theta}{2} + \frac{\pi}{2}\right)\right] \end{pmatrix}. \tag{5.226}$$

(В этом случае $\mathbf{E}_1$ и $\mathbf{E}_2$ представляют собой проекторы на спиновые состояния, ориентированные в плоскости OXZ перпендикулярно оси, направленной вдоль биссектрисы угла между $|u\rangle$ и $|v\rangle$.)

Найдите среднюю информацию $I(\theta)$, приобретаемую Бобом (взаимную информацию между приготовленным состоянием и результатом измерения) во всех трех случаях, и изобразите графики всех их, как функций θ. Какое измерение следует выбрать Бобу?

5.2. Относительная энтропия. *Относительная энтропия* $S(\boldsymbol{\rho}|\boldsymbol{\sigma})$ двух матриц плотности $\boldsymbol{\rho}$ и $\boldsymbol{\sigma}$ определяется соотношением

$$S(\boldsymbol{\rho}|\boldsymbol{\sigma}) = \operatorname{tr} \boldsymbol{\rho}\big(\log \boldsymbol{\rho} - \log \boldsymbol{\sigma}\big). \tag{5.227}$$

Покажите, что $S(\boldsymbol{\rho}|\boldsymbol{\sigma})$ неотрицательна, и выведите некоторые следствия этого свойства.

a) Дифференцируемая вещественная функция вещественной переменной называется *вогнутой*, если для всех x и y

$$f(y) - f(x) \leqslant (y - x) f'(x). \tag{5.228}$$

Покажите, что если a и b — наблюдаемые, а f - вогнутая, то

$$\operatorname{tr}\big[f(\mathbf{b}) - f(\mathbf{a})\big] \leqslant \operatorname{tr}\big[(\mathbf{b} - \mathbf{a}) f'(\mathbf{a})\big]. \tag{5.229}$$

b) Покажите, что $f(x) = -x \log x$ является вогнутой функцией при $x > 0$.

c) Используя (a) и (b), покажите, что $S(\boldsymbol{\rho}|\boldsymbol{\sigma}) \geqslant 0$ для любых двух матриц плотности $\boldsymbol{\rho}$ и $\boldsymbol{\sigma}$.

d) Используя неотрицательность $S(\boldsymbol{\rho}|\boldsymbol{\sigma})$, покажите, что если $\boldsymbol{\rho}$ имеет носитель в пространстве размерности D, то

$$S(\boldsymbol{\rho}) \leqslant \log D. \tag{5.230}$$

e) Используя неотрицательность относительной энтропии, докажите *субаддитивность* энтропии

$$S(\rho_{AB}) \leqslant S(\rho_A) + S(\rho_B). \tag{5.231}$$

[**Указание**. Рассмотрите относительную энтропию $\rho_A \otimes \rho_B$ и ρ_{AB}.]

f) Используя субаддитивность, докажите *вогнутость* энтропии

$$S\left(\sum_i \lambda_i \rho_i\right) \geqslant \sum_i \lambda_i S(\rho_i), \tag{5.232}$$

где λ_i — вещественные положительные числа, сумма которых равна единице. [**Указание**. Примените свойство субаддитивности к

$$\rho_{AB} = \sum_i \lambda_i (\rho_i)_A \otimes (|e_i\rangle\langle e_i|)_B.] \tag{5.233}$$

g) Используя свойство субаддитивности, докажите *неравенство треугольника* (также называемое неравенством Араки – Либа):

$$S(\rho_{AB}) \geqslant |S(\rho_A) - S(\rho_B)|. \tag{5.234}$$

[**Указание**. Рассмотрите очищение ρ_{AB}: то есть постройте чистое состояние $|\psi\rangle_{ABC}$ такое, что $\rho_{AB} = \mathrm{tr}_C\left(|\psi\rangle_{ABC\,ABC}\langle\psi|\right)$. Затем примените свойство субаддитивности к ρ_{BC}.]

5.3. Монотонность Линдблада – Ульмана. Согласно теореме, доказанной Линдбладом и Ульманом, относительная энтропия на $\mathcal{H}_A \otimes \mathcal{H}_B$ обладает свойством, называемым *монотонностью*

$$S(\rho_A|\sigma_A) \leqslant S(\rho_{AB}|\sigma_{AB}). \tag{5.235}$$

Относительная энтропия двух матриц плотности системы AB не может быть меньше, чем редуцированная относительная энтропия подсистемы A.

a) Используя монотонность Линдблада — Ульмана, докажите свойство сильной субаддитивности энтропии фон Неймана. [**Указание**. Рассмотрите относительную энтропию $\rho_A \otimes \rho_{BC}$ и ρ_{ABC} в тройной системе ABC.]

b) Используя монотонность Линдблада – Ульмана, покажите, что действие супероператора не может увеличить относительную энтропию, то есть

$$S(\$\rho|\$\sigma) \leqslant S(\rho|\sigma), \tag{5.236}$$

где $\$$ — произвольный супероператор (вполне положительное отображение). [**Указание**. Вспомните, что произвольный супероператор имеет унитарное представление.]

c) Покажите, что из (b) вытекает, что супероператор не может увеличивать информацию Холево ансамбля $\mathcal{E} = \{\rho_x, p_x\}$ смешанных состояний:

$$\chi(\$(\mathcal{E})) \leqslant \chi(\mathcal{E}), \tag{5.237}$$

где

$$\chi(\mathcal{E}) = S\Big(\sum_x p_x \rho_x\Big) - \sum_x p_x S(\rho_x). \tag{5.238}$$

5.4. ПОЗМ Переса – Вутерса. Рассмотрите источник информации Переса – Вутерса, описанный в § 5.4.2. Он приготавливает одно из трех состояний

$$|\Phi_a\rangle = |\varphi_a\rangle|\varphi_a\rangle, \qquad a = 1, 2, 3, \tag{5.239}$$

каждое из которых появляется с априорной вероятностью $1/3$, где состояния $|\varphi_a\rangle$ определены в (5.149).

a) Выразите матрицу плотности

$$\rho = \frac{1}{3}\sum_a |\Phi_a\rangle\langle\Phi_a| \tag{5.240}$$

в базисе Белла максимально запутанных состояний $\{|\phi^{\pm}\rangle, |\psi^{\pm}\rangle\}$ и вычислите $S(\rho)$.

b) Для трех векторов $|\Phi_a\rangle$, $a = 1, 2, 3$ постройте определенное в (5.162) «достаточно хорошее измерение». (Вновь разложите векторы $|\Phi_a\rangle$ в базисе Белла.) В этом случае ДХИ является ортогональным измерением. Выразите элементы базиса ДХИ в базисе Белла.

c) Вычислите взаимную информацию результатов ДХИ и приготовления состояний.

Глава 6

Квантовые вычисления

6.1. Классические (вычислительные) схемы

В первой главе было введено понятие квантового компьютера. Здесь мы более строго определим модель квантовых вычислений и отметим ее некоторые основные свойства. Но прежде чем объяснять, что делает квантовый компьютер, возможно, следовало бы поговорить о том, что делает классический компьютер.

6.1.1. Универсальные вентили

Классический (детерминистский) компьютер вычисляет функции: по данным n битам на входе он производит m битов на выходе, которые однозначно определены входом. То есть он находит значение

$$f\colon \{0,1\}^n \to \{0,1\}^m \tag{6.1}$$

для определенного, состоящего из n битов, аргумента. Функция, имеющая m-битовое значение, эквивалентна m функциям с однобитовым значением каждая, поэтому вполне можно сказать, что основной задачей, выполняемой классическим компьютером, является вычисление

$$f\colon \{0,1\}^n \to \{0,1\}. \tag{6.2}$$

Нетрудно подсчитать количество таких функций. Существует 2^n возможных входов, и для каждого из них имеется два возможных выхода. Итак, всего имеется 2^{2^n} функций, переводящих n битов в один.

Вычисление любой такой функции можно свести к последовательности элементарных логических операций. Разделим возможные значения входа

$$x = x_1\, x_2 \ldots x_n \tag{6.3}$$

на множество значений, для которых $f(x) = 1$, и дополнительное ему множество, для которого $f(x) = 0$. Для каждого $x^{(a)}$ такого, что $f(x^{(a)}) = 1$, рассмотрим функцию $f^{(a)}$ такую, что

$$f^{(a)}(x) = \begin{cases} 1, & x = x^{(a)}, \\ 0, & x \neq x^{(a)}. \end{cases} \tag{6.4}$$

Тогда

$$f(x) = f^{(1)}(x) \vee f^{(2)}(x) \vee f^{(3)}(x) \vee \dots . \tag{6.5}$$

f — логическое OR ($\vee$) всех функций $f^{(a)}$. В двоичной арифметике двухбитовая операция $\vee$ может быть представлена как

$$x \vee y = x + y - x \cdot y; \tag{6.6}$$

она имеет значение 0, если x и y оба равны нулю, и значение 1 в противном случае.

Рассмотрим вычисление $f^{(a)}$. В том случае, когда $x^{(a)} = 111\dots1$, мы можем записать

$$f^{(a)}(x) = x_1 \wedge x_2 \wedge x_3 \wedge \dots \wedge x_n; \tag{6.7}$$

это логическое AND ($\wedge$) всех n битов. В двоичной арифметике AND представляет собой произведение

$$x \wedge y = x \cdot y. \tag{6.8}$$

Для любого другого $x^{(a)}$ функция $f^{(a)}$ вновь строится как логическое AND n битов, в котором к каждому равному нулю $x_i^{(a)}$ предварительно применяется операция логического NOT ($\neg$), например:

$$f^{(a)}(x) = (\neg x_1) \wedge x_2 \wedge x_3 \wedge (\neg x_4) \wedge \dots, \tag{6.9}$$

если

$$x^{(a)} = 0110\dots. \tag{6.10}$$

Логическая операция NOT в двоичной арифметике представляется как

$$\neg x = 1 - x. \tag{6.11}$$

Мы построили функцию $f(x)$ из трех элементарных логических отношений: NOT, AND, OR. Полученное выражение (6.5) называется «дизъюнктивной нормальной формой» $f(x)$. Мы также неявно использовали еще одну операцию, COPY, превращающую один бит в два:

$$\text{COPY} : x \to xx. \tag{6.12}$$

Операция COPY необходима, поскольку каждая $f^{(a)}$ в разложении f по дизъюнктивным нормальным формам требует свою собственную копию x, на которую она будет действовать.

Фактически мы можем сократить набор элементарных логических отношений до меньшего. Определим операцию NAND («NOT-AND») соотношением

$$x \uparrow y = \neg(x \wedge y) = (\neg x) \vee (\neg y). \tag{6.13}$$

В двоичной арифметике операцией NAND служит

$$x \uparrow y = 1 - x \cdot y. \tag{6.14}$$

Если мы можем выполнять COPY, то NAND можно использовать для выполнения операции NOT:

$$x \uparrow x = 1 - x^2 = 1 - x = \neg x. \tag{6.15}$$

(Или если мы можем приготовить константу $y = 1$, тогда $x \uparrow 1 = 1 - x = \neg x$.) Аналогично

$$(x \uparrow y) \uparrow (x \uparrow y) = \neg(x \uparrow y) = 1 - (1 - x \cdot y) = x \cdot y = x \wedge y, \tag{6.16}$$

а

$$(x \uparrow x) \uparrow (y \uparrow y) = (\neg x) \uparrow (\neg y) = 1 - (1 - x) \cdot (1 - y) = x + y - x \cdot y = x \vee y. \tag{6.17}$$

Итак, если мы можем выполнять COPY, то NAND выполняет AND и OR. Таким образом, одного логического отношения NAND вместе с COPY достаточно для вычисления любой функции f. [Вы можете убедиться в том, что возможной альтернативой в выборе универсального логического отношения является NOR («NOT-OR»):[1]

$$x \downarrow y = \neg(x \vee y) = (\neg x) \wedge (\neg y).] \tag{6.18}$$

Если мы можем приготовить постоянный бит ($x = 0$ или $x = 1$), то количество элементарных операций может быть сокращено с двух до одной. Операция NAND/NOT

$$(x, y) \rightarrow (1 - x, 1 - x \cdot y) \tag{6.19}$$

[1] Обратим внимание на то, что вторые равенства в (6.13) и (6.18) фактически являются следствиями известного в теории множеств *принципа двойственности*: (1) Дополнение пересечения равно сумме дополнений и (2) Дополнение суммы равно пересечению дополнений. См., например, А.Н. Колмогоров, С.В. Фомин, *Элементы теории функций и функционального анализа*, М.: Наука, 1976. — *Прим. ред.*

вычисляет NAND (если мы игнорируем первый выходящий бит) и снимает копию (если возьмем в качестве второго входящего бита $y = 1$, а затем применим NOT к обоим выходящим битам)[1]. Таким образом, можно сказать, что NAND/NOT является универсальным логическим вентилем. Если в нашем распоряжении имеется запас постоянных битов, а вентили NAND/NOT могут применяться к любой выбранной паре входящих битов, то мы можем выполнить последовательность операций NAND/NOT для вычисления любой функции $f\colon \{0,1\}^n \to \{0,1\}$ при любом значении входа $x = x_1 x_2 \ldots x_n$.

Этими соображениями мотивируется модель вычислительной схемы. Компьютер имеет несколько основных компонентов, которые могут выполнять элементарные операции с битами или парами битов, такие как COPY, NOT, AND, OR. Он также может готовить постоянные биты или входящие переменные биты. Вычисление представляет собой конечную последовательность таких операций, *схему*, применяемую к точно определенной строке входящих битов[2]. Результатом вычисления является конечное значение всех битов, оставшихся после выполнения всех элементарных операций.

То, что для вычисления любой функции, зависящей от конечного входа, достаточно лишь нескольких элементарных операций, является фундаментальным результатом теории вычислений. Он означает, что с помощью очень простых аппаратных средств можно выполнять сколь угодно сложные вычисления.

До сих пор мы обсуждали вычисления, применяющиеся к частному фиксированному входу, но можно рассматривать и *семейства* схем, действующих на входы переменной длины. Семейства схем предоставляют полезную модель для анализа и классификации *сложности* вычислений, которая будет естественным образом обобщена, когда мы обратимся к квантовым вычислениям.

6.1.2. Сложность схем

Исследуя сложность, мы часто будем интересоваться функциями с однобитовым выходом

$$f : \{0,1\}^n \to \{0,1\}. \tag{6.20}$$

О такой функции f можно сказать, что она кодирует решение «проблемы принятия решения» — функция проверяет вход и выдает ответ ДА или НЕТ.

[1] Можно предложить более простую реализацию операции COPY путем последовательного применения вентиля NAND/NOT: $(x,0) \to (1-x,1) \to (x, 1-(1-x)\cdot 1) = (x,x)$. — *Прим. ред.*

[2] Схемы должны быть *апериодическими*, в том смысле, что *полностью* замкнутые циклы в них недопустимы.

Часто оказывается, что вопрос, который не хотелось бы формулировать словесно как вопрос, имеющий ответ ДА/НЕТ, может быть «переформулирован» как проблема принятия решения. Например, функция, определяющая FACTORING-проблему (задача факторизации):

$$f(x,y) = \begin{cases} 1, & \text{если целое } x \text{ имеет делитель, меньший чем } y, \\ 0 & \text{в противном случае;} \end{cases} \tag{6.21}$$

знание $f(x,y)$ для всех $y < x$ эквивалентно знанию *наименьшего* нетривиального множителя x. Другим важным примером проблемы принятия решения служит HAMILTONIAN-проблема (задача нахождения гамильтонова обхода): рассмотрим ℓ-вершинный граф, представленный $\ell \times \ell$-матрицей смежности (равенство единице ее ij-элемента означает, что существует ребро, связывающее вершины i и j); функция

$$f(x) = \begin{cases} 1, & \text{если граф } x \text{ имеет гамильтонов обход,} \\ 0 & \text{в противном случае.} \end{cases} \tag{6.22}$$

(Обход называется гамильтоновым, если он проходит через каждую вершину графа только один раз.)

Мы хотим оценивать трудность проблемы, количественно определяя необходимые для ее решения ресурсы. Сложность проблемы принятия решения разумно измерять *размером* минимальной схемы, вычисляющей соответствующую функцию $f\colon \{0,1\}^n \to \{0,1\}$. Под размером мы понимаем количество элементарных операций или компонентов, которые нужно объединить в схему, чтобы вычислить f. Также можно интересоваться тем, какого *времени* требует вычисление, если многие операции могут выполняться параллельно. *Глубина* схемы представляет собой необходимое количество шагов при условии, что операции, действующие на различные биты, могут выполняться одновременно (то есть глубиной является максимальная длина прямого пути от входа схемы до ее выхода). *Ширина* схемы — это максимальное количество операций, выполняемых на любом из ее этапов.

Мы хотели бы разделить проблемы принятия решения на два класса: легкие и трудные. Но где следует провести границу? Рассмотрим с этой целью бесконечные семейства проблем принятия решения с переменным размером входа; то есть количество битов на входе может быть любым целым n. Тогда можно проверить, как соизмеряется с n размер схемы, решающей проблему.

Однако в использовании масштабного поведения семейства схем в качестве характеристики сложности задачи имеется одна тонкость. Было бы

обманом скрывать сложность проблемы в *конструкции* схемы. Следовательно, мы должны ограничиться семействами, обладающими приемлемыми свойствами «однородности» — должно быть «просто» построить схему с $n+1$-битовым входом, коль скоро схема с n-битовым входом уже построена.

Пусть с данным семейством функций $\{f_n\}$ (где f_n имеет n-битовый вход) сопоставлены вычисляющие их схемы $\{C_n\}$. Мы говорим, что $\{C_n\}$ является семейством схем «полиномиального размера», если размер C_n растет с n не быстрее некоторой степени n:

$$\text{size}\,(C_n) \leqslant \text{poly}(n), \tag{6.23}$$

где poly обозначает полином. Тогда определим:

$$P = \left\{ \begin{array}{l} \text{проблема принятия решения, решаемая семействами схем} \\ \text{полиномиального размера} \end{array} \right\}$$

(P означает разрешимость за «полиномиальное время»). Проблемы принятия решения, принадлежащие P, являются «простыми», остальные — «сложными». Заметим, что C_n вычисляет $f_n(x)$ для любого возможного n-битового входа и, следовательно, если проблема принятия решения принадлежит P, то даже «в худшем случае» мы можем найти ответ, используя схему, размер которой не превышает $\text{poly}(n)$. (Как отмечалось выше, мы неявно предполагаем, что семейство схем «однородно», так что проблема *разработки* схемы сама может быть решена с помощью алгоритма, требующего полиномиального времени. В этом предположении разрешимость за полиномиальное время с помощью семейства схем эквивалентна разрешимости за полиномиальное время с помощью универсальной машины Тьюринга.)

Конечно, чтобы определить размер схемы, вычисляющей f_n, мы должны знать, что представляют собой ее элементарные компоненты. К счастью, принадлежность проблемы к P не зависит от выбора набора вентилей, до тех пор пока они универсальны, их множество конечно, а каждый вентиль действует на ограниченное множество битов. Один универсальный набор вентилей может *моделироваться* другим.

Огромное большинство семейств функций $f\colon \{0,1\}^n \to \{0,1\}$ не принадлежит P. Для большей части функций выход существенно случаен, и нет лучшего способа «вычислить» $f(x)$, чем обратиться к таблице ее значений. Но поскольку имеется 2^n n-битовых входов, то эта таблица имеет *экспоненциальный* размер, такой же размер должна иметь и схема, кодирующая эту таблицу. Проблемы из P принадлежат очень частному классу —

они имеют такую структуру, что функция f может быть эффективно вычислена.

Особый интерес представляют проблемы принятия решения, на которые можно ответить, приведя легко проверяемый пример. Пусть, например, для данных x и $y < x$ трудно (в худшем случае) определить, имеет ли x множитель, меньше чем y. Но если кто-нибудь любезно предоставит такое $z < y$, на которое делится x, то нам просто проверить, что z действительно является множителем x. Аналогично трудно определить, имеет ли граф гамильтонов обход, но если кто-нибудь нам его показал, то легко убедиться в том, что он и в самом деле гамильтонов.

Эту идею, что проблема может быть трудно разрешимой, но ее решение, коль скоро оно найдено, легко проверяемо, можно формализовать с помощью понятия «недетерминированной» схемы. Ассоциированная с $C_n(x^{(n)})$ недетерминированная схема $\tilde{C}_{n,m}(x^{(n)}, y^{(m)})$ обладает свойством[1]:

$$C_n(x^{(n)}) = 1, \quad \text{если} \quad \tilde{C}_{n,m}(x^{(n)}, y^{(m)}) = 1, \quad \text{для } \textit{некоторого } y^{(m)} \tag{6.24}$$

(где $x^{(n)}$ представляет n битов, а $y^{(m)}$ — m битов). Таким образом, если $y^{(m)}$ оказалось удачно подобранным для некоторого $x^{(n)}$, то мы можем использовать $\tilde{C}_{n,m}$ для *проверки* того, что $C_n(x^{(n)}) = 1$. Определим

$$NP = \left\{ \begin{array}{l} \text{проблемы принятия решения, допускающие семейство } \textit{недете-} \\ \textit{рминированных} \text{ схем полиномиального размера} \end{array} \right\}$$

(NP означает разрешимость за «недетерминированное полиномиальное время»). Если проблема принадлежит NP, то нет гарантии, что она простая. Это означает лишь то, что ее решение легко проверить, если мы располагаем правильной информацией. Очевидно, что $P \subseteq NP$. Подобно P, NP-проблемы образуют малый подкласс всех проблем принятия решения.

---

[1] Согласно одному из определений *проблема принадлежит* классу сложности NP, если существует схема полиномиального по n размера, проверяющая предложенное решение за полиномиальное время. Проверяющая схема [в данном случае это $\tilde{C}(x^{(n)}, y^{(m)})$] является *детерминированной*. Термин *недетерминированный* происходит от того, что согласно другому определению (см. следующее чуть ниже определение в основном тексте) NP-проблема допускает решение за полиномиальное время на так называемой *недетерминированной* машине Тьюринга, способной в некоторых состояниях выбирать различные варианты вычисления. См. А. Китаев, А. Шень, М. Вялый, Классические и квантовые вычисления, М., МЦНМО, ЧеРо (1999). — *Прим. ред.*

Многое в теории сложности опирается на фундаментальное предположение

$$\text{Предположение:} \quad P \neq NP; \tag{6.25}$$

существуют сложные проблемы принятия решения, решения которых легко проверяемы. К сожалению, эта важная гипотеза все еще ждет своего доказательства. Однако после 30-ти лет попыток показать обратное – большинство специалистов в теории сложности твердо уверены в ее справедливости[1].

Важным примером NP-проблемы является CIRCUIT-SAT[2]. В этом случае вход представляет собой состоящая из n вентилей схема C с m-битовым входом и однобитовым выходом. Проблема состоит в том, чтобы найти, существует ли *какой-нибудь* m-битовый вход, для которого выход равен единице. Функция, которую необходимо вычислить, представляет собой

$$f(C) = \begin{cases} 1, & \text{если существует } x^{(m)} \text{ } \textit{такое, что } C(x^{(m)}) = 1, \\ 0 & \text{в противном случае.} \end{cases} \tag{6.26}$$

Это NP-проблема, поскольку *данную* схему легко *смоделировать* и вычислить ее выход для любого частного входа.

Я перехожу к формулировке некоторых важных результатов теории сложности, которые будут для нас важны. Здесь не будет времени для доказательств. Вы можете почерпнуть больше, обратившись к одному из многих учебников по этому предмету. [3]

Многие идеи, порожденные теорией сложности, вытекают из теоремы Кука (1971). Она утверждает, что *любая* NP-проблема *полиномиально приводима* к CIRCUIT-SAT. Это означает, что для любой PROBLEM $\in NP$ существует семейство схем полиномиального размера, которое отображает «ситуацию» $x^{(n)}$ для PROBLEM на «ситуацию» $y^{(m)}$ для CIRCUIT-SAT, то есть

$$\text{CIRCUIT-SAT}(y^{(m)}) = 1, \quad \text{если} \quad \text{PROBLEM}(x^{(n)}) = 1. \tag{6.27}$$

Отсюда следует, что если бы в нашем распоряжении имелось магическое устройство, которое могло бы эффективно решать CIRCUIT-SAT

[1] Проблема соотношения классов сложности P и NP остается нерешенной и в настоящее время (начало 2007г.). Более того она входит в список важнейших задач математики XXI в. – *Прим. ред.*

[2] CIRCUIT-SAT (circuit-satisfyability) problem – проблема выполнимости схемы. (перев.)

[3] Одной из лучших является книга M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Competeness*, русский перевод М. Гэри, Д. Джонсон, *Вычислительные машины и труднорешаемые задачи*, М.: Мир (1982). [Краткий, но достаточно полный, обзор классов сложности и их иерархии можно найти в первой части книги: А. Китаев, А. Шень, М. Вялый, *Классические и квантовые вычисления*, М.: МЦНМО, ЧеРо (1999). – *Прим. ред.*]

(CIRCUIT-SAT «оракул»), то с помощью полиномиальной редукции мы могли бы связаться с ним, чтобы эффективно решить PROBLEM. Из теоремы Кука следует, что если вдруг окажется, что CIRCUIT-SAT $\in P$, то $P = NP$.

Проблема, которая, подобно CIRCUIT-SAT, обладает тем свойством, что к ней полиномиально приводится любая проблема из NP, называется NP-полной (NPC). После Кука было найдено множество других примеров NPC-проблем. Чтобы показать, что PROBLEM $\in NP$ является NP-полной, достаточно найти другую, полиномиально приводимую к ней проблему, о которой уже известно, что она NP-полная. Например, можно продемонстрировать полиномиальную приводимость CIRCUIT-SAT к HAMILTONIAN. Тогда из теоремы Кука следует, что HAMILTONIAN тоже NP-полна.

Если мы предположим, что $P \neq NP$, то отсюда следует, что в NP существуют проблемы промежуточной трудности (класс NPI). Это ни P, ни NPC.

Другой важный класс сложности называется со-NP. Эвристически NP-проблемами разрешимости являются те, на которые мы можем ответить, приведя пример, если ответом является ДА, в то время как на со-NP-проблему можно ответить *контрпримером*, если ответом является НЕТ. Более формально

$$\{C\} \in NP\colon C(x) = 1, \quad \text{если} \quad C(x, y) = 1 \quad \text{для некоторого } y; \tag{6.28}$$

$$\{C\} \in \text{со-}NP\colon C(x) = 1, \quad \text{если} \quad C(x, y) = 1 \quad \text{для всех } y. \tag{6.29}$$

Очевидно, что между классами NP и со-NP существует симметрия — рассматриваем ли мы проблему из NP или из со-NP зависит от того, как был сформулирован вопрос. (Проблема «Существует ли гамильтонов обход?» принадлежит классу NP. Проблема «Действительно ли, что гамильтонова обхода не существует?» принадлежит классу со-NP.) Однако интересен вопрос: существует ли проблема ($\notin P$), одновременно принадлежащая *обоим* классам: NP и со-NP. Если да, то мы можем легко проверить ответ (коль скоро имеем подходящий пример) независимо от того является им ДА или НЕТ. Считается (хотя и не доказано), что $NP \neq$ со-NP. (Приведя пример, мы можем показать, что граф имеет гамильтонов обход, но мы не знаем, как подобным образом показать, что он *не* имеет гамильтоновых обходов!) При условии, что $NP \neq$ со-NP, существует теорема, которая утверждает, что ни одна со-NP-проблема не принадлежит NPC. Следовательно, проблемы, принадлежащие пересечению NP и со-NP и при этом не входящие в P, являются хорошими кандидатами для включения их в NPI.

Фактически такой проблемой, принадлежащей $NP \cap$co-NP, относительно которой считается, что она не входит в P, является FACTORING-проблема. Как уже отмечалось, FACTORING-проблема принадлежит классу NP, поскольку мы можем легко проверить правильность предоставленного множителя числа x. Но она также принадлежит и co-NP, поскольку известно, что если нам дано простое число, то (по крайней мере в принципе) мы можем эффективно проверить его простоту. Таким образом, если некто сообщает нам простые множители числа x, то мы можем эффективно проверить, что разложение на простые множители правильно, и можем *исключить*, что любое целое, меньшее y, является делителем числа x. Следовательно, похоже на то, что FACTORING-проблема принадлежит классу NPI[1].

Мы пришли к грубой (гипотетической) картине структуры $NP \cap$ co-NP. Классы NP и co-NP не совпадают, но имеют нетривиальное пересечение. Класс P принадлежит пересечению $NP \cap$ co-NP (поскольку $P =$ co-P), но кроме этого оно содержит проблемы, не входящие в P (подобные FACTORING-проблеме). Ни NPC ни co-NPC не пересекаются с $NP \cap$ co-NP.

Можно было бы гораздо больше рассказать о теории сложности, но мы ограничимся здесь упоминанием еще одного элемента, связанного с обсуждением квантовой сложности. Иногда полезно рассматривать *вероятностные* схемы, которые имеют доступ к генератору случайных чисел. Например, вентиль в вероятностной схеме может действовать одним из двух способов и «подбрасывает монету», чтобы решить, какое *действие* выполнить. При одном фиксированном входе такая схема может избрать один из множества вычислительных путей. Об алгоритме, выполняемом вероятностной схемой, говорят как о «рандомизированном».

Если мы приступаем к проблеме принятия решения, используя вероятностный компьютер, то получаем распределение вероятностей результатов. Таким образом, мы не будем всегда получать обязательно правильный ответ. Но если для любого возможного входа вероятность получения правильного ответа больше, чем $\frac{1}{2} + \delta$ ($\delta > 0$), то такая машина полезна. Фактически мы можем многократно повторить вычисление и, учитывая

[1] В 2002 г. группой индийских математиков предложен полиномиальный детерминированный алгоритм проверки целых чисел на простоту. Его асимптотическая сложность $O((\log n)^k)$, где n — проверяемое число, k — целое число ~ 10. Таким образом, проблема распознавания простоты целого числа принадлежит классу сложности P. M. Agrawal, N. Kayal, N. Saxena, *Primes is in P*, Annals of Math., **160**, 781–793 (2004), http://www.math.princeton.edu/~annals/ . См. также Л. Ю. Бараш, *Алгоритм AKS проверки чисел на простоту и поиск констант генераторов псевдослучайных чисел*, Безопасность информационных технологий, **2**, 27–38 (2005). — *Прим. ред.*

«голос большинства», достичь вероятности ошибки, меньшей ε. Более того, количество необходимых для этого повторений вычисления всего лишь полилогарифмически зависит от ε^{-1}[1].

Если проблема допускает решение с помощью семейства вероятностных схем полиномиального размера, которые всегда дают правильный ответ с вероятностью, большей $\frac{1}{2} + \delta$ (для любого входа и при фиксированном $\delta > 0$), то мы говорим, что она принадлежит классу BPP («вероятностное полиномиальное время с ограниченной ошибкой»). Очевидно, что

$$P \subset BPP, \tag{6.30}$$

но соотношение между NP и BPP не известно. В частности, не доказано, что BPP содержится в NP.

6.1.3. Обратимые вычисления

Разрабатывая модель квантового компьютера, мы обобщим модель классических вычислительных схем. Но нашими квантовыми логическими вентилями будут унитарные и, следовательно, обратимые преобразования, тогда как классические логические вентили типа NAND необратимы. Прежде чем обсуждать квантовые схемы, полезно рассмотреть некоторые особенности классических обратимых вычислений.

Помимо их связи с квантовыми вычислениями, другие побудительные причины для изучения обратимых классических вычислений обсуждались в первой главе. Как заметил Ландауэр, поскольку необратимые логические элементы стирают информацию, они с необходимостью диссипативны и, следовательно, требуют постоянного расхода энергии. Но если компьютер оперирует обратимо, то в принципе не должно быть никакой потребности в энергии. Мы можем вычислять даром!

Обратимый компьютер вычисляет обратимую функцию, преобразуя n битов в другие n битов:

$$f\colon \{0,1\}^n \to \{0,1\}^n; \tag{6.31}$$

функция должна быть обратимой, то есть для каждого выхода существует единственный вход; тогда мы в принципе в состоянии пройти вычисление в обратном порядке и, зная выход, воспроизвести вход. Поскольку это взаимно-однозначная функция, ее можно рассматривать как перестановку 2^n строк из n битов — всего таких функций $(2^n)!$.

[1] То есть количество необходимых повторений ограничено сверху полиномом $\mathrm{poly}(\log(\varepsilon^{-1}))$. — *Прим. ред.*

Конечно, любое необратимое вычисление можно «оформить» как вычисление обратимой функции. Например, для любой $f : \{0,1\}^n \to \{0,1\}^m$ мы можем сконструировать $\tilde{f} : \{0,1\}^{n+m} \to \{0,1\}^{n+m}$ такую, что

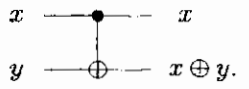
$$\tilde{f}(x; 0^{(m)}) = (x; f(x)) \tag{6.32}$$

(где $0^{(m)}$ обозначает m битов, первоначально положенных равными нулю). Поскольку $\tilde{f}$ преобразует каждое $(x; 0^{(m)})$ к своему, отличному от других, результату, она может быть расширена до обратимой функции от $n + m$ битов. Следовательно, для любой f, преобразующей n битов в m, существует обратимая функция $\tilde{f}$, преобразующая $n + m$ битов в $n + m$ битов, которая вычисляет $f(x)$, действуя на $(x; 0^{(m)})$.

Как теперь построить сложные обратимые вычисления из элементарных компонентов — то есть что образует набор универсальных вентилей? Мы увидим, что одно- и двухбитовых обратимых вентилей не достаточно; для универсальных обратимых вычислений нам понадобятся трехбитовые вентили.

Из четырех 1-бит $\to$ 1-бит вентилей обратимы два — тривиальный и NOT. Из $(2^4)^2 = 256$ возможных 2-бит $\to$ 2-бит вентилей обратимы $4! = 24$. Одним из особенно интересных является вентиль контролируемое NOT или XOR, с которым мы уже сталкивались в четвертой главе:

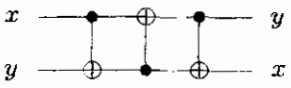
$$XOR\colon (x, y) \to (x, x \oplus y), \tag{6.33}$$



Этот вентиль инвертирует второй бит, если первый равен единице, и ничего не делает, если первый бит равен нулю (отсюда его название: контролируемое NOT). Его квадрат является тривиальным, то есть он обратен по отношению к самому себе. Конечно, этот вентиль выполняет операцию NOT на втором бите, если первый бит положить равным единице, и выполняет операцию копирования, если начальным значением y является нуль:

$$XOR\colon (x, 0) \to (x, x), \tag{6.34}$$

С помощью схемы



построенной из трех XOR-ов, мы можем поменять местами два бита[1]:

$$(x, y) \to (x, x \oplus y) \to (y, x \oplus y) \to (y, x). \tag{6.35}$$

С помощью этих обменов можно тасовать биты внутри схемы, собирая их вместе, если мы хотим подействовать на них конкретным компонентом в фиксированном положении.

Чтобы увидеть, что одно- и двухбитовые вентили неуниверсальны, заметим, что все они *линейны*. Каждый обратимый двухбитовый вентиль действует по правилу

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x' \\ y' \end{pmatrix} = \mathcal{M} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}, \tag{6.36}$$

где константа $\binom{a}{b}$ принимает одно из четырех возможных значений, а $\mathcal{M}$ — одна из шести обратимых матриц

$$\mathcal{M} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}. \tag{6.37}$$

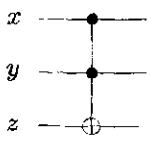
[Все суммирования в (6.36) выполняются по модулю 2.] Комбинируя шесть вариантов $\mathcal{M}$ с четырьмя возможными константами, мы получаем 24 различных вентиля, которыми исчерпывается набор всех обратимых $2 \to 2$ вентилей.

Поскольку линейные преобразования замкнуты относительно композиции, любая схема, скомбинированная из обратимых $2 \to 2$ (и $1 \to 1$) вентилей, будет вычислять линейную функцию

$$x \to \mathcal{M}x + a. \tag{6.38}$$

Однако при $n \geqslant 3$ существуют нелинейные обратимые функции n битов. Важным примером служит *вентиль Тоффоли* $\theta^{(3)}$ (или дважды контролируемое NOT)

$$\theta^{(3)} : (x, y, z) \to (x, y, z \oplus x \cdot y); \tag{6.39}$$



[1] В двоичной арифметике сложение по модулю 2 определяется как $x \oplus y = x + y - 2x \cdot y$. Эта операция, очевидно, коммутативна $x \oplus y = y \oplus x$ и ассоциативна $(x \oplus y) \oplus z = x \oplus (y \oplus z) = x \oplus y \oplus z$. — *Прим. ред.*

он инвертирует третий бит, если два первых равны единице, и ничего не делает в противном случае. Подобно XOR-вентилю он обратен по отношению к самому себе.

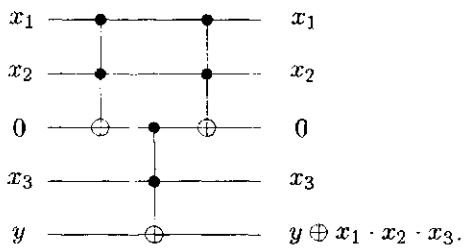
В отличие от обратимых двухбитовых вентилей, $\theta^{(3)}$ является универсальным вентилем булевой логики, если мы можем фиксировать некоторые входящие биты и игнорировать некоторые выходящие биты. Если начальное значение z равно единице, то на третьем выходе возникает $x \uparrow y = 1 - x \cdot y$ — мы можем выполнить NAND. Если же мы фиксируем $x = 1$, то вентиль Тоффоли функционирует подобно XOR-вентилю и может использоваться для копирования.

Вентиль Тоффоли $\theta^{(3)}$ универсален в том смысле, что, используя только его, можно построить схему, вычисляющую любую обратимую функцию (при условии, что можно фиксировать входящие биты и игнорировать выходящие). Полезно показать это непосредственно, не опираясь на наше более раннее доказательство того, что NAND/NOT является универсальным для вычисления любой булевой функции. Фактически можно показать следующее: из вентилей NOT и Тоффоли $\theta^{(3)}$ мы можем построить универсальную функцию n битов при условии, что мы имеем доступ к одному дополнительному биту в памяти.

В качестве первого шага покажем, что из трехбитовых вентилей Тоффоли $\theta^{(3)}$ можно построить n-битовый вентиль Тоффоли $\theta^{(n)}$, действующий по правилу

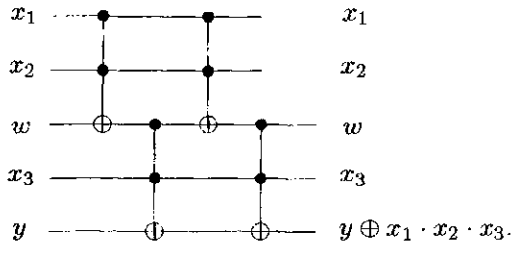
$$(x_1, x_2, \ldots, x_{n-1}, y) \to (x_1, x_2, \ldots, x_{n-1}, y \oplus x_1 x_2 \ldots x_{n-1}). \qquad (6.40)$$

Эта конструкция требует дополнительного бита из вспомогательного пространства. Например, мы конструируем $\theta^{(4)}$ из вентилей $\theta^{(3)}$ с помощью схемы



Цель последнего $\theta^{(3)}$-вентиля — вернуть вспомогательному биту его начальное значение, равное нулю. В действительности, добавив еще один $\theta^{(3)}$-вен-

тиль, можно реализовать $\theta^{(4)}$, который работает независимо от начального значения вспомогательного бита:



Снова мы можем *исключить* последний вентиль, если нас не беспокоит изменение значения вспомогательного бита[1].

Мы можем убедиться в том, что вспомогательный бит действительно необходим. Поскольку $\theta^{(4)}$ представляет собой нечетную перестановку (фактически подстановку) 16-ти четырехбитовых строк, он переставляет 1111 и 1110. Однако $\theta^{(3)}$, действуя на любые три бита из четырех, осуществляет четную перестановку; например, действуя на последние три бита, он переставляет 0111 с 0110 или 1111 с 1110[2]. Поскольку произведение четных перестановок также является четным, мы не можем получить $\theta^{(4)}$ как произведение вентилей $\theta^{(3)}$, действующее только на четыре бита.

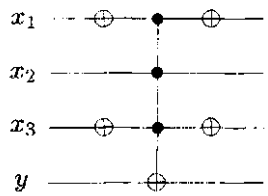
Конструкцию $\theta^{(4)}$ из четырех $\theta^{(3)}$ можно непосредственно обобщить на конструкцию $\theta^{(n)}$ из двух $\theta^{(n-1)}$ и двух $\theta^{(3)}$ (только расширив x_1 в приведенной выше диаграмме до нескольких битов). Итерируя эту конструкцию, мы получим $\theta^{(n)}$-схему, состоящую из $2^{n-2} + 2^{n-3} - 2$ вентилей $\theta^{(3)}$. Более того, для этого достаточно только одного вспомогательного бита[3]. (Если нам необходимо построить $\theta^{(n)}$, то для этого подойдет любой доступный вспомогательный бит, так как схема возвращает ему его начальное значение.) Следующим шагом заметим, что, соединяя $\theta^{(n)}$ с вентилями NOT, мы можем в действительности менять значение контрольной строки,

[1]На самом деле *восстановление* начального значения вспомогательного бита w, осуществляемое в этой схеме вторым верхним $\theta^{(3)}$-вентилем, является *необходимой* операцией. Именно благодаря этому на выходе второго нижнего $\theta^{(3)}$-вентиля получается $[y \ominus (wx_3 \oplus x_1x_2x_3)] \oplus wx_3 = y \oplus x_1x_2x_3$. Последнее равенство легко проверяется с учетом коммутативности и ассоциативности сложения по модулю 2. — *Прим. ред.*

[2]По-видимому автор имеет ввиду, что $\theta^{(4)}$-вентиль осуществляет нечетное количество нетривиальных подстановок (одну), а $\theta^{(3)}$-вентиль — четное (две). — *Прим. ред.*

[3]С большим вспомогательным пространством можно значительно эффективнее построить $\theta^{(n)}$ из $\theta^{(3)}$ вентилей (см. упражнения).

«управляющей» вентилем. Например, схема



инвертирует значение y, если $x_1x_2x_3 = 010$, и действует тривиально в противном случае. Таким образом, эта схема переставляет две строки: 0100 и 0101. Подобным образом, с помощью вентилей $\theta^{(n)}$ и NOT можно придумать схему, которая переставляет любые две n-битовые строки, отличающиеся только одним битом. (Расположение бита, которым они отличаются, выбирается в качестве *цели* вентиля $\theta^{(n)}$.)

Но фактически подстановка, обменивающая любые две n-битовые строки, может быть представлена как произведение подстановок, которые обменивают строки, отличающиеся только одним битом. Если a_0 и a_s — две строки, расстояние Хемминга между которыми равно s (отличаются s позициями), то существует цепь

$$a_0,\ a_1,\ a_2,\ a_3,\ \ldots,\ a_s \tag{6.41}$$

такая, что каждая строка в этой последовательности удалена от ближайших соседей на равное единице расстояние Хемминга. Следовательно, каждая из подстановок

$$(a_0, a_1),\ (a_1, a_2),\ (a_2, a_3),\ \ldots,\ (a_{s-1}, a_s) \tag{6.42}$$

может быть реализована как вентиль $\theta^{(n)}$, соединенный вентилями NOT. Комбинируя подстановки, находим

$$\begin{aligned}(a_0, a_s) = {} & (a_{s-1}, a_s)(a_{s-2}, a_{s-1})\ldots \\ & \ldots(a_2, a_3)(a_1, a_2)(a_0, a_1)(a_1, a_2)(a_2, a_3)\ldots \\ & \ldots(a_{s-2}, a_{s-1})(a_{s-1}, a_s);\end{aligned} \tag{6.43}$$

мы можем сконструировать подстановку с равным s расстоянием Хемминга из $2s-1$ подстановки с единичным расстоянием Хемминга. Отсюда следует, что мы можем построить (a_0, a_s) из вентилей $\theta^{(n)}$ и NOT.

Наконец, поскольку каждая перестановка является произведением подстановок, мы показали, что любая обратимая функция n-битов (каждая перестановка n-битовых строк) представляет собой произведение вентилей $\theta^{(3)}$ и вентилей NOT, использующее только один бит вспомогательного пространства.

Конечно, операция NOT может быть выполнена с помощью вентиля $\theta^{(3)}$, если мы фиксируем два входящих бита равными единицам. Таким образом, вентиль Тоффоли $\theta^{(3)}$ является универсальным для обратимых вычислений, если мы можем фиксировать входящие биты и отбрасывать выходящие.

6.1.4. Компьютер бильярдных шаров

Двухбитовых вентилей достаточно для универсальных необратимых вычислений, но для универсальных обратимых вычислений необходимы трехбитовые вентили. Возникает соблазн сказать, что для их реализации необходимы «трехчастичные взаимодействия», так что создание обратимого аппаратного обеспечения является более сложной проблемой, чем создание необратимого. Однако это утверждение в какой-то мере может быть обманчивым.

Фредкин описал, как можно построить универсальный обратимый компьютер, в котором фундаментальным взаимодействием является упругое столкновение между двумя бильярдными шарами. Шары радиуса $1/\sqrt{2}$ движутся по квадратной решетке, период которой равен единице. В каждый целочисленный момент времени центр каждого шара находится в узле решетки; наличие или отсутствие шара в данном узле (в этот момент времени) кодирует бит информации. В течение каждого единичного интервала времени каждый (подвижный) шар проходит единичное расстояние вдоль одного из направлений решетки. Иногда, в целочисленные моменты времени, происходит упругое рассеяние на 90° двух шаров, занимающих узлы, удаленные друг от друга на расстояние $\sqrt{2}$ (соединенные диагональю ячейки решетки).

Прибор программируется закреплением части шаров в некоторых узлах, которые действуют как идеальные отражатели. После задания начальных положений и направлений движения подвижных шаров система выполняет программу, эволюционируя в течение конечного интервала времени в соответствии с механикой Ньютона. Результат считывается путем наблюдения конечных положений подвижных шаров. Столкновения недиссипативны, так что после обращения всех скоростей вычисление может быть пройдено в обратном порядке.

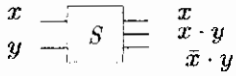
Чтобы показать, что эта машина является универсальным обратимым компьютером, мы должны объяснить, как в ней реализовать универсальный вентиль. Удобно рассмотреть трехбитовый *вентиль Фредкина*

$$(x, y, z) \to (x, xz + \bar{x}z, xz + \bar{x}y), \tag{6.44}$$

который меняет местами y и z, если $x = 0$ (мы ввели обозначение $\bar{x} = \neg x$). Вы можете убедиться в том, что вентиль Фредкина может моделировать вентиль NAND/NOT, если мы фиксируем входы и игнорируем выходы.

Мы можем построить вентиль Фредкина из более примитивного объекта, *вентиля-переключателя*. Переключатель преобразует два бита в три, действуя по правилу

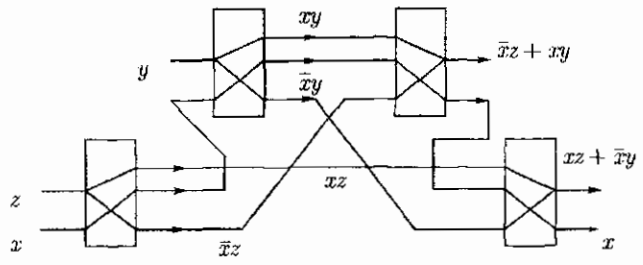
$$(x, y) \to (x, x \cdot y, \bar{x} \cdot y). \tag{6.45}$$



Этот вентиль «обратим» в том смысле, что его можно пройти в обратном направлении, действуя на ограниченный трехбитовый вход, принимающий одно из четырех значений:

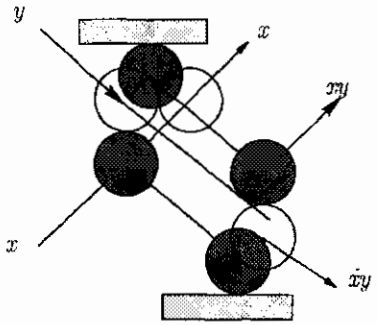
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \tag{6.46}$$

Более того, переключатель сам по себе универсален; фиксируя входы и игнорируя выходы, он может выполнить операцию NOT ($y = 1$, третий выход), AND (второй выход) и COPY ($y = 1$, первый и второй выходы). Тогда не удивительно, что, комбинируя переключатели, можно построить универсальный $3 \to 3$ *вентиль*. Действительно, схема



образует вентиль Фредкина из четырех переключателей (два из них проходятся в прямом направлении, два других — в обратном). Время задержки, необходимое для синхронизации, явно не показано.

В компьютере бильярдных шаров переключатель строится из таких двух отражателей, чтобы (в случае $x = y = 1$) два движущихся шара столкнулись дважды. В этом случае траектории шаров имеют вид



Шар, помеченный как x, вылетает из вентиля вдоль той же траектории (и в то же самое время) независимо от наличия или отсутствия другого шара. Однако при $x = 1$ положение другого шара (если он имеется) смещается вниз по сравнению с его конечным положением при $x = 0$ — это и есть переключатель. Коль скоро мы можем построить переключатель, то можем построить и вентиль Фредкина и, таким образом, реализовать универсальную обратимую логику на компьютере бильярдных шаров.

Очевидной слабостью схемы бильярдных шаров является то, что начальные ошибки положений и скоростей шаров будут быстро накапливаться и в конце концов компьютер окажется несостоятельным. Как отмечалось в первой главе (а Ландауэр настоятельно подчеркивал это), подобным недостатком будет страдать любая предлагаемая схема недиссипативных вычислений. Чтобы контролировать ошибки, мы должны быть в состоянии сжимать фазовое пространство прибора, что с необходимостью будет диссипативным процессом.

6.1.5. Экономия пространства

Но кроме проблемы контроля ошибок, есть еще один ключевой вопрос, касающийся обратимых вычислений. Как распорядиться вспомогательным пространством, необходимым для того, чтобы сделать вычисление обратимым?

Обсуждая универсальность вентиля Тоффоли, мы видели, что в принципе можно выполнить любое обратимое вычисление, используя очень малое вспомогательное пространство. Но на практике может оказаться невероятно сложно понять, как выполнить конкретное вычисление, используя минимальное пространство, и во всяком случае экономия пространства может обернуться непомерным расходом времени.

Существует общая стратегия моделирования необратимого вычисления на обратимом компьютере. Каждый необратимый вентиль, NOT или COPY, можно моделировать вентилем Тоффоли, фиксируя входы и игнорируя выходы. Мы накапливаем и сохраняем весь «мусор» выходящих битов, которые необходимы, чтобы обратить этапы вычисления. Вычисление выполняется вплоть до завершения, после чего делается копия выхода. (Эта COPY-операция логически обратима.) Затем вычисление производится в обратном порядке, чтобы избавиться от «мусора» и вернуть все регистры в их начальные состояния. С помощью этой процедуры обратимая схема осуществляется примерно дважды, до тех пор, пока не будет выполнена моделируемая необратимая схема, а весь генерируемый при этом мусор – выброшен без какой-либо диссипации и, следовательно, энергетических затрат.

Эта процедура работает, но требует огромного пространства памяти – ее необходимый объем растет линейно с продолжительностью T моделируемого необратимого вычисления. Фактически пространство можно использовать гораздо более эффективно (лишь с минимальным замедлением), так что его необходимый объем растет как $\log T$ вместо T. (То есть существует универсальная схема, требующая пространства $\propto \log T$; конечно, моделируя конкретное вычисление, можно добиться даже лучшего результата.)

Чтобы эффективнее использовать пространство, разделим вычисление на более мелкие шаги приблизительно одинакового размера и, когда это возможно, будем обращать их в процессе вычисления. Однако, подобно тому как мы не в состоянии выполнить k-ый шаг вычисления до тех пор, пока не завершен $k-1$-ый шаг, мы не сможем *обратить* k-ый шаг, если предварительно был обращен $k-1$-ый шаг[1]. Необходимый объем пространства (чтобы хранить наш мусор) будет расти как максимальное значение числа шагов вперед за вычетом количества выполненных шагов назад.

Проблему, с которой мы столкнулись, можно сравнить с *обратимой иг-*

[1] Мы скромно предполагаем, что не настолько прозорливы, чтобы предвидеть, какая часть выхода $k-1$-го шага может потребоваться позже. Следовательно, мы сохраняем полную запись состояния машины после $k-1$-го шага, которая не должна удаляться до тех пор, пока не будет обновлена запись после завершения следующего шага.

рой камешками[1]. Выполняемые шаги образуют одномерный ориентированный граф с узлами, пронумерованными как $1, 2, 3, \ldots, T$. Выполнение k-го шага моделируется помещением камешка в k-ый узел графа, а выполнение k-го шага в обратном направлении моделируется удалением камешка из k-го узла. В начале игры нет узлов, занятых камешками, а с каждым ходом мы их добавляем или удаляем. Однако мы не можем поместить камешек в k-ый узел (за исключением $k = 1$) до тех пор, пока не заполнен $k-1$-ый, а также мы не можем удалить камешек из k-го узла (за исключением $k = 1$), если свободен $k-1$-ый узел. Задача в том, чтобы заполнить узел T (завершить вычисление), не используя большего, чем это необходимо, количества камешков (генерируя минимальный объем мусора).

Фактически с помощью n камешков мы можем достичь узла $T = 2^n - 1$, но продвинуться дальше не сможем.

Можно построить рекурсивную процедуру, позволяющую добраться до $T = 2^{n-1}$-го узла с помощью n камешков, оставляя в игре только один камешек. Пусть $F_1(k)$ обозначает помещение камешка в k-й узел, а $F_1^{-1}(k)$ — удаление камешка из k-го узла. Тогда[2]

$$F_2(1,2) = F_1(1)F_1(2)F_1^{-1}(1) \tag{6.47}$$

оставляет камешек в узле $k = 2$, используя максимум два камешка на промежуточных этапах. Аналогично

$$F_3(1,4) = F_2(1,2)F_2(3,4)F_2^{-1}(1,2) \tag{6.48}$$

достигает узла $k = 4$, используя максимум три камешка, а

$$F_4(1,8) = F_3(1,4)F_3(5,8)F_3^{-1}(1,4) \tag{6.49}$$

достигает узла $k = 8$, используя четыре камешка. Очевидно, можно построить процедуру $F_n(1, 2^{n-1})$, которая использует максимум n камешков и оставляет в игре один. [Программа

$$F_n(1,2^{n-1})F_{n-1}(2^{n-1}+1, 2^{n-1}+2^{n-2})\ldots F_1(2^n-1) \tag{6.50}$$

оставляет в игре все n камешков и позволяет заполнить максимально удаленный узел $k = 2^n - 1$.]

---

[1] Как было отмечено Беннетом. Относительно последнего обсуждения см.: M. Li and P. Vitanyi, *Reversibility and Adiabatic Computation: Trading Time and Space for Energy*, Proc. R. Soc. London, **A 452**, 769–789 (1996); quant-ph/9703022.

[2] Правые части (6.47) – (6.50) следует читать слева направо. Именно в этом порядке выполняются описываемые ими действия. – *Прим. ред.*

Понимаемая как программа для выполнения $T = 2^{n-1}$ шагов вычисления, эта стратегия игрока в камешки представляет собой моделирование, требующее роста пространства как $n \sim \log T$. Насколько продолжительным может быть это моделирование? На каждом этапе описанной выше рекурсивной процедуры два шага вперед заменялись двумя шагами вперед и одним назад. Следовательно, $T_{\rm irr} = 2^n$ шагов необратимого вычисления моделируются $T_{\rm rev} = 3^n$ шагами обратимого вычисления или

$$T_{\rm rev} = \left(T_{\rm irr}\right)^{\log 3/\log 2} = \left(T_{\rm irr}\right)^{1,58}; \tag{6.51}$$

мы имеем умеренный степенной закон замедления.

В действительности мы можем уменьшить замедление до

$$T_{\rm rev} \sim \left(T_{\rm irr}\right)^{1+\varepsilon} \tag{6.52}$$

при любом $\varepsilon > 0$. Вместо того чтобы заменять два шага вперед двумя шагами вперед и одним назад, заменим ℓ шагов вперед ℓ шагами вперед и $\ell - 1$-им шагом назад. Состоящая из n этапов рекурсивная процедура достигает узла ℓ^n, используя максимум $n(\ell-1)+1$ камешков. Теперь мы имеем $T_{\rm irr} = \ell^n$, а $T_{\rm rev} = (2\ell-1)^n$, так что

$$T_{\rm rev} \sim \left(T_{\rm irr}\right)^{\log(2\ell-1)/\log \ell}; \tag{6.53}$$

показатель степени замедления равен

$$\frac{\log(2\ell-1)}{\log \ell} = \frac{\log 2\ell + \log\left(1-\frac{1}{2\ell}\right)}{\log \ell} \simeq 1 + \frac{\log 2}{\log \ell}, \tag{6.54}$$

а требуемое пространство растет как

$$S \simeq n\ell \simeq \ell \frac{\log T}{\log \ell}. \tag{6.55}$$

Таким образом, для любого фиксированного $\varepsilon > 0$ мы можем добиться S, растущего как $\log T$, и замедления, не большего чем $\left(T_{\rm irr}\right)^{1+\varepsilon}$. (Для игры в камешки это не оптимальный способ, если наша цель — продвинуться как можно дальше, используя минимально возможное количество камешков. Мы используем больше камешков, чтобы добраться до T-го шага, зато делаем это быстрее.)

Итак, мы видим, что обратимая схема может успешно моделировать схему, построенную из необратимых вентилей, не требуя нереальных ресурсов памяти и не вызывая неразумно большого замедления. Почему это важно? Вас может беспокоить, что поскольку обратимое вычисление «труднее» необратимого, то классификация сложности зависит от того, какими вычислениями мы пользуемся, обратимыми или необратимыми. Однако это не так, поскольку необратимый компьютер легко моделируется обратимым.

6.2. Квантовые схемы

Теперь мы готовы сформулировать математическую модель квантового компьютера. Мы обобщим модель классической вычислительной схемы на модель квантовой вычислительной схемы.

Классический компьютер оперирует битами. Он оснащен конечным набором вентилей, которые могут применяться к множеству битов. Квантовый компьютер оперирует кубитами. Будем предполагать, что он тоже оснащен дискретным набором фундаментальных компонентов, называемых *квантовыми вентилями*. Каждый квантовый вентиль представляет собой унитарное преобразование, действующее на определенное число кубитов. В квантовых вычислениях конечное количество n кубитов первоначально полагаются имеющими значение $|00\ldots0\rangle$. Выполняемая схема построена из конечного числа квантовых вентилей, действующих на эти кубиты. Наконец, выполняется измерение фон Неймана всех кубитов (или некоторого подмножества кубитов), проецирующее каждый из них на базис $\{|0\rangle, |1\rangle\}$. Результат этого измерения является результатом вычисления.

Некоторые особенности этой модели требуют комментария.

(1) Неявно подразумевается, но очень важно, что гильбертово пространство прибора имеет естественное разложение на тензорное произведение пространств более низкой размерности, в данном случае — двумерных пространств кубитов. Конечно, вместо этого мы могли бы рассматривать тензорное произведение, допустим, кутритов. Но в любом случае мы считаем, что существует естественное разложение на подсистемы, которое соответствует квантовым вентилям, действующим одновременно только на несколько подсистем. С математической точки зрения это свойство вентилей является решающим для формулировки хорошо определенного понятия квантовой сложности. С физической точки зрения фундаментальной причиной естественного разложе-

ния на подсистемы является *локальность*; реальные квантовые вентили должны действовать в ограниченной области пространства, то есть компьютер разбивается на подсистемы, взаимодействующие только со своими ближайшими соседями.

(2) Так как унитарные преобразования образуют континуум, может показаться необязательным постулировать, что машина может выполнять только выбранные из дискретного множества квантовые операции. Тем не менее мы принимаем это ограничение, поскольку не хотим, сталкиваясь с выполнением нового вычисления, всякий раз изобретать его новую физическую реализацию.

(3) Мы могли бы допустить, чтобы наши квантовые вентили были супероператорами, а конечным измерением — ПОЗМ. Но поскольку мы можем просто моделировать супероператор, выполняя унитарное преобразование в расширенной системе, или — ПОЗМ, выполняя измерение фон Неймана в расширенной системе, сформулированная модель обладает достаточной общностью.

(4) Мы могли бы допустить, чтобы заключительное измерение было коллективным измерением или проектором на другой базис. Но любое такое измерение можно реализовать, выполняя подходящее унитарное преобразование после проецирования на стандартный базис $\{|0\rangle, |1\rangle\}^n$. Конечно, сложные коллективные измерения лишь с некоторыми затруднениями можно преобразовать в измерения в стандартном базисе и при характеристике сложности алгоритма эти трудности следует иметь в виду.

(5) Мы могли бы допустить наличие измерений на промежуточных этапах вычислений с последующим выбором квантовых вентилей, обусловленным результатами этих измерений. Но фактически тот же результат всегда может быть достигнут с помощью квантовой схемы, в которой все измерения отложены вплоть до ее окончания. (Хотя в принципе мы можем отложить измерения, на практике может оказаться полезным их выполнение на промежуточных этапах квантового алгоритма.)

Будучи унитарным преобразованием, квантовый вентиль обратим. Фактически классический обратимый компьютер представляет собой частный случай квантового компьютера. Классический обратимый вентиль

$$x^{(n)} \to y^{(n)} = f\left(x^{(n)}\right), \tag{6.56}$$

выполняющий перестановку n-битовых строк, может рассматриваться как унитарное преобразование, действующее на «вычислительный» базис $\{|x_i\rangle\}$ как

$$U : |x_i\rangle \rightarrow |y_i\rangle. \tag{6.57}$$

Это действие унитарно, поскольку все 2^n строк $|y_i\rangle$ взаимно ортогональны. Квантовое вычисление, построенное из таких классических вентилей, преобразует $|0\ldots0\rangle$ в одно из состояний вычислительного базиса, так что конечное измерение является детерминированным.

Имеется три главные проблемы, касающиеся нашей модели, к которым мы хотели бы обратиться. Первой из них является *универсальность*. Самое общее унитарное преобразование, которое может быть выполнено на n кубитах, является элементом $U(2^n)$. Наша модель могла бы оказаться неполной, если бы в $U(2^n)$ существовали такие преобразования, которые мы не могли бы выполнить. На самом деле мы увидим, что существует множество способов выбрать дискретный набор *универсальных квантовых вентилей*. Используя набор универсальных вентилей, можно построить схемы, вычисляющие унитарное преобразование, сколь угодно близкое к любому элементу $U(2^n)$.

Благодаря универсальности, существует также аппаратно-независимое понятие *квантовой сложности*. Мы можем определить новый класс сложности BQP — класс проблем принятия решения, которые с высокой вероятностью могут быть решены с помощью квантовой схемы полиномиального размера. Так как один универсальный квантовый компьютер может эффективно моделироваться другим, то этот класс не зависит от деталей аппаратного обеспечения (от выбранного нами набора универсальных вентилей).

Заметим, что квантовый компьютер может легко моделировать классический вероятностный компьютер: он может приготовить состояние $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, а затем спроецировать его на $\{|0\rangle, |1\rangle\}$, генерируя случайный бит. Следовательно, класс BPP несомненно содержится в BQP. Однако, как обсуждалось в первой главе, представляется достаточно разумным ожидать, что в действительности BQP шире, чем BPP, поскольку классический вероятностный компьютер не может легко моделировать квантовый компьютер. Фундаментальная трудность состоит в том, что гильбертово пространство n кубитов огромно, размерности 2^n, и, следовательно, математическое описание типичного вектора в этом пространстве исключительно сложно.

Вторая проблема — наилучшим образом характеризовать ресурсы, необходимые для моделирования квантового компьютера классическим.

Мы увидим, что, несмотря на обширность гильбертова пространства, классический компьютер может моделировать n-кубитовый квантовый компьютер, даже если его запас памяти ограничен, то есть полиномиален по n. Это означает, что BQP содержится в классе сложности $PSPACE$ проблем принятия решения, которые могут быть решены с использованием пространства полиномиального размера, но могут потребовать для этого экспоненциального времени. [Мы знаем, что NP также содержится в $PSPACE$, так как проверка $\tilde{C}(x^{(n)}, y^{(m)}) = 1$ для всех $y^{(m)}$ может быть выполнена с использованием полиномиального пространства.][1]

Третьей важной проблемой, к которой следует обратиться, является *точность*. Класс BQP формально определен при идеализированном предположении, что квантовые вентили могут выполняться с идеальной точностью. Ясно, что при любой реализации квантового вычисления очень важно ослабить это предположение. Семейство квантовых схем полиномиального размера, которое решает трудную проблему, не представляло бы большого интереса, если бы от используемых в схемах вентилей требовалась экспоненциальная точность. Мы покажем, что на самом деле это не так. Идеализированная квантовая схема из T вентилей с приемлемой точностью может моделироваться вентилями с шумом при условии, что вероятность ошибки на один вентиль пропорциональна $1/T$.

Таким образом, квантовые компьютеры бросают серьезный вызов сильному тезису Черча–Тьюринга, утверждающему, что любая физически разумная модель вычисления может быть смоделирована вероятностными классическими схемами с полиномиальным замедлением в худшем случае. Но до сих пор нет строгого доказательства того, что

$$BQP \neq BPP, \tag{6.58}$$

и в ближайшем будущем оно не предвидится[2]. Действительно, следствием было бы

$$BPP \neq PSPACE, \tag{6.59}$$

что решило бы один из давно стоящих, кардинальных открытых вопросов теории сложности.

Возможно, более реалистично надеяться на доказательство того, что $BPP \neq BQP$ вытекает из другого стандартного предположения теории сложности, такого как $P \neq NP$. Такое доказательство до сих пор

[1] В действительности в иерархии сложности существует еще одна ступенька, которая может разделять BQP и $PSPACE$; можно показать, что $BQP \subseteq P^{\#P} \subseteq PSPACE$, но ниже мы не будем рассматривать $P^{\#P}$.

[2] То есть не следует ожидать «переделывизированного доказательства». Разделение между BPP и BQP «относительно оракула» будет установлено ниже в этой главе.

не найдено. Но хотя мы все еще не в состоянии доказать, что квантовые компьютеры имеют возможности, выходящие далеко за пределы возможностей обычных компьютеров, тем не менее можно привести свидетельства, указывающие на то, что $BPP \neq BQP$. Мы увидим, что существуют проблемы, которые выглядят сложными (для классического вычисления), но тем не менее могут быть успешно решены с помощью квантовых схем.

Таким образом, кажется вероятным то, что классификация сложности будет зависеть от того, какой компьютер для решения задачи используется, классический или квантовый. Если такое разделение действительно существует, то именно квантовая классификация должна рассматриваться как более фундаментальная, поскольку она в большей степени опирается на физические законы, управляющие Вселенной.

6.2.1. Точность

Обсудим проблему точности. Представим, что мы хотим выполнить вычисление, в котором квантовые вентили $\mathbf{U}_1$, $\mathbf{U}_2, \ldots, \mathbf{U}_T$ последовательно применяются к начальному состоянию $|\varphi_0\rangle$. Состояние, приготовленное идеальной квантовой схемой, имеет вид

$$|\varphi_T\rangle = \mathbf{U}_T \mathbf{U}_{T-1} \ldots \mathbf{U}_2 \mathbf{U}_1 |\varphi_0\rangle. \tag{6.60}$$

Но в действительности наши вентили не являются идеально точными. Пытаясь применить унитарное преобразование $\mathbf{U}_t$, мы вместо этого применяем некоторое «близкое» унитарное преобразование $\tilde{\mathbf{U}}_t$. (Конечно, это не самый общий тип ошибки, который можно предположить, — унитарное преобразование $\mathbf{U}_t$ может оказаться замененным *супероператором*. В этом случае применимы рассуждения, подобные следующим ниже, но здесь мы ограничим наше внимание «унитарными ошибками».)

Ошибки приводят к тому, что действительное состояние компьютера удаляется от идеального. Как сильно оно удаляется? Пусть $|\varphi_t\rangle$ обозначает идеальное состояние после применения t квантовых вентилей, так что

$$|\varphi_t\rangle = \mathbf{U}_t |\varphi_{t-1}\rangle. \tag{6.61}$$

Но если мы применяем действительное преобразование $\tilde{\mathbf{U}}_t$, то

$$\tilde{\mathbf{U}}_t |\varphi_{t-1}\rangle = |\varphi_t\rangle + |E_t\rangle, \tag{6.62}$$

где

$$|E_t\rangle = (\tilde{\mathbf{U}}_t - \mathbf{U}_t)|\varphi_{t-1}\rangle \tag{6.63}$$

— ненормированный вектор. Если $|\tilde{\varphi}_t\rangle$ обозначает действительное состояние после t шагов, то

$$\begin{aligned} |\tilde{\varphi}_1\rangle &= |\varphi_1\rangle + |E_1\rangle, \\ |\tilde{\varphi}_2\rangle &= \tilde{\mathbf{U}}_2|\tilde{\varphi}_1\rangle = |\varphi_2\rangle + |E_2\rangle + \tilde{\mathbf{U}}_2|E_1\rangle \end{aligned} \tag{6.64}$$

и так далее; в конечном счете мы получим

$$\begin{aligned} |\tilde{\varphi}_T\rangle = |\varphi_T\rangle &+ |E_T\rangle + \tilde{\mathbf{U}}_T|E_{T-1}\rangle + \tilde{\mathbf{U}}_T\tilde{\mathbf{U}}_{T-1}|E_{T-2}\rangle \\ &+ \ldots + \tilde{\mathbf{U}}_T\tilde{\mathbf{U}}_{T-1}\ldots\tilde{\mathbf{U}}_2|E_1\rangle. \end{aligned} \tag{6.65}$$

Итак, мы представили разность между $|\tilde{\varphi}_T\rangle$ и $|\varphi_T\rangle$ в виде суммы T оставшихся слагаемых. Наихудший случай, дающий наибольшее отклонение $|\tilde{\varphi}_T\rangle$ от $|\varphi_T\rangle$, возникает, если все оставшиеся слагаемые ориентированы в одном направлении, так что ошибки интерферируют конструктивно. Следовательно,

$$\begin{aligned} \||\tilde{\varphi}_T\rangle - |\varphi_T\rangle\| \leqslant \||E_T\rangle\| &+ \||E_{T-1}\rangle\| + \\ &+ \ldots + \||E_2\rangle\| + \||E_1\rangle\|, \end{aligned} \tag{6.66}$$

где учтено, что $\|\mathbf{U}|E_i\rangle\| = \||E_i\rangle\|$ для любого унитарного $\mathbf{U}$.

Пусть $\|\mathbf{A}\|$ обозначает норму оператора $\mathbf{A}$, то есть максимум модуля его собственных значений. Тогда

$$\||E_t\rangle\| = \|(\tilde{\mathbf{U}}_t - \mathbf{U}_t)|\varphi_{t-1}\rangle\| \leqslant \|(\tilde{\mathbf{U}}_t - \mathbf{U}_t)\| \tag{6.67}$$

(поскольку $|\varphi_{t-1}\rangle$ нормирован). Предположим теперь, что при каждом значении t ошибка нашего квантового вентиля ограничена неравенством

$$\|(\tilde{\mathbf{U}}_t - \mathbf{U}_t)\| < \varepsilon. \tag{6.68}$$

Тогда после применения T квантовых вентилей мы имеем

$$\||\tilde{\varphi}_T\rangle - |\varphi_T\rangle\| < T\varepsilon; \tag{6.69}$$

в этом смысле накопление ошибки в состоянии растет пропорционально продолжительности вычисления.

Отклонение, ограниченное неравенством (6.68), может быть представлено в эквивалентной форме $\|\mathbf{W}_t - \mathbf{1}\|$, где $\mathbf{W}_t = \tilde{\mathbf{U}}_t \mathbf{U}_t^\dagger$. Так как оператор $\mathbf{W}_t$ унитарен, каждое его собственное число имеет вид фазы $e^{i\theta}$, а соответствующее собственное значение оператора $\mathbf{W}_t - \mathbf{1}$ имеет модуль

$$|e^{i\theta} - 1| = (2 - 2\cos\theta)^{1/2}, \tag{6.70}$$

так что (6.68) требует, чтобы каждое собственное значение удовлетворяло неравенству

$$\cos\theta > 1 - \frac{\varepsilon^2}{2} \tag{6.71}$$

(или $|\theta| \lesssim \varepsilon$ для малых ε). Природа неравенства (6.69) понятна. В каждый момент времени $|\tilde{\varphi}\rangle$ поворачивается относительно $|\varphi\rangle$ на угол порядка ε (в худшем случае), а расстояние между векторами возрастает максимум на величину порядка ε.

Какая точность является достаточно хорошей? На последнем этапе вычисления мы выполняем ортогональное измерение, а вероятность результата a в идеальном случае равна

$$P(a) = |\langle a|\varphi_T\rangle|^2. \tag{6.72}$$

Вследствие ошибок действительной вероятностью будет

$$\tilde{P}(a) = |\langle a|\tilde{\varphi}_T\rangle|^2. \tag{6.73}$$

Если действительный вектор близок к идеальному, то и распределения вероятностей тоже близки. Если мы просуммируем по ортонормированному базису $\{|a\rangle\}$, то получим

$$\sum_a |\tilde{P}(a) - P(a)| \leqslant 2\||\tilde{\varphi}_T\rangle - |\varphi_T\rangle\|, \tag{6.74}$$

как вы покажете в домашнем упражнении. Следовательно, если при больших T мы сохраняем неизменным (и малым) $T\varepsilon$, то ошибка в распределении вероятностей также остается фиксированной. В частности, если мы разработали квантовый алгоритм, который с вероятностью выше $\frac{1}{2} + \delta$ правильно решает проблему принятия решения (в идеальном случае), тогда с вероятностью, превышающей $\frac{1}{2}$, мы можем добиться успеха и с помощью наших шумящих вентилей, если действие этих вентилей может быть выполнено с точностью $T\varepsilon < O(\delta)$. Семейство квантовых схем может реально решать сложные проблемы в классе BQP до тех пор, пока мы в состоянии улучшать точность выполнения вентилей пропорционально объему вычислений.

6.2.2. $BQP \subseteq PSPACE$

Конечно, классический компьютер может моделировать любую квантовую схему. Но какой объем памяти ему для этого потребуется? Поскольку моделирование n-кубитовой схемы включает в себя манипулирование матрицами размера 2^n, то с наивной точки зрения может показаться, что для этого необходим экспоненциальный по n запас памяти. Однако теперь мы покажем, что с приемлемой точностью (хотя и очень медленно!) моделирование может быть выполнено в пространстве полиномиального размера. Это означает, что класс квантовой сложности BQP содержится в классе $PSPACE$ задач, которые могут быть решены с использованием пространства *полиномиального* размера.

Объектом классического моделирования является вычисление вероятности каждого возможного результата a заключительного измерения

$$\text{Prob}(a) = |\langle a|\mathbf{U}^{(T)}|0\rangle|^2, \tag{6.75}$$

где

$$\mathbf{U}^{(T)} = \mathbf{U}_T\mathbf{U}_{T-1}\ldots\mathbf{U}_2\mathbf{U}_1 \tag{6.76}$$

— произведение T квантовых вентилей. Каждый $\mathbf{U}_t$, действующий на n кубитов, может быть представлен унитарной $2^n \times 2^n$-матрицей, характеризуемой комплексными матричными элементами

$$\langle y|\mathbf{U}_t|x\rangle, \tag{6.77}$$

где $x, y \in \{0, 1, \ldots, 2^n - 1\}$. Явно выписывая произведение матриц, мы имеем

$$\langle a|\mathbf{U}^{(T)}|0\rangle = \sum_{x_t}\langle a|\mathbf{U}_T|x_{T-1}\rangle\langle x_{T-1}|\mathbf{U}_{T-1}|x_{T-2}\rangle\ldots \\ \ldots\langle x_2|\mathbf{U}_2|x_1\rangle\langle x_1|\mathbf{U}_1|0\rangle. \tag{6.78}$$

Уравнение (6.78) представляет собой вариант представления квантового вычисления «интегралом по траекториям» — амплитуда вероятности конечного результата a выражается в виде когерентной суммы амплитуд каждого из огромного количества $2^{n(T-1)}$ возможных вычислительных путей, начинающихся в точке 0 и после T шагов заканчивающихся в a.

Чтобы вычислить $\langle a|\mathbf{U}^{(T)}|0\rangle$, наш классический симулятор должен сложить $2^{n(T-1)}$ комплексных чисел в уравнении (6.78). Первая проблема, с которой мы встречаемся, состоит в том, что классические схемы конечного размера реализуют целочисленную арифметику, тогда как матричные

элементы $\langle y|\mathbf{U}_i|x\rangle$ не обязаны быть рациональными числами. Следовательно, классический симулятор должен выполнять приближенные вычисления с разумной точностью. Каждое из $2^{n(T-1)}$ слагаемых суммы представляет собой произведение T комплексных сомножителей. Накапливаемые ошибки непременно должны быть малыми, если мы выражаем матричные элементы с помощью m точных битов, где m велико по сравнению с $n(T-1)$. Следовательно, мы можем заменить каждый комплексный матричный элемент парой целых чисел определенного знака, принимающих значения $\{0, 1, 2, \ldots, 2^{m-1}\}$. Эти целые числа дают двоичное разложение вещественной и мнимой частей матричного элемента, выраженного с точностью 2^{-m}.

Нашему симулятору потребуется вычислить каждое слагаемое в (6.78) и накопить их полную сумму. Но каждое добавление требует только умеренного объема пространства памяти, и, более того, поскольку для следующего сложения необходимо сохранять только накопленную частичную сумму, не очень большое пространство требуется для суммирования всех слагаемых, даже если их экспоненциально много.

Итак, остается лишь рассмотреть вычисление типичного слагаемого суммы, произведения T матричных элементов. Нам потребуется классическая схема, вычисляющая

$$\langle y|\mathbf{U}_i|x\rangle; \tag{6.79}$$

эта схема принимает $2n$ входящих битов (x, y) и выдает на выходе $2m$-битовое (комплексное) значение матричного элемента. Имея схему, выполняющую эту функцию, легко построить схему, которая перемножает комплексные числа, не используя большого пространства.

Наконец, обратимся к свойствам, которые мы потребовали от набора квантовых вентилей, — это дискретное множество вентилей, каждый из которых действует на ограниченное количество кубитов. Поскольку имеется фиксированное (и конечное) количество вентилей, то существует лишь конечное количество вентилей-подпрограмм, с которым нашему симулятору необходимо уметь обращаться. А поскольку вентили действуют только на несколько кубитов, почти все их матричные элементы исчезают (если n велико), а значение $\langle y|\mathbf{U}|x\rangle$ может быть определено (с требуемой точностью) с помощью простой схемы, требующей незначительной памяти.

Например, в случае однокубитового вентиля, действующего на первый кубит,

$$\langle y_1 y_2 \ldots y_n|\mathbf{U}|x_1 x_2 \ldots x_n\rangle = 0, \quad \text{если} \quad x_2 x_3 \ldots x_n \neq y_2 y_3 \ldots y_n. \tag{6.80}$$

Простая схема может сравнить x_2 с y_2, x_3 с y_3 и так далее и дать на выходе

нуль, если равенство не выполняется. В случае равенства она выдает одно из четырех комплексных чисел

$$\langle y_1|\mathbf{U}_i|x_1\rangle \tag{6.81}$$

с m точными битами. Простая схема может закодировать $8m$ битов этой комплекснозначной 2×2-матрицы. Подобным образом простая схема, требующая пространство лишь полиномиального по n и m размера, может вычислить матричные элементы любого вентиля фиксированного размера.

Таким образом, классический компьютер с пространством, ограниченным сверху размером $\mathrm{poly}(n)$, может моделировать n-кубитовый универсальный квантовый компьютер и, следовательно, BQP $\subseteq$ PSPACE. Конечно, также очевидно, что описанное нами моделирование требует экспоненциального времени, так как нам необходимо вычислить сумму $2^{n(T-1)}$ комплексных чисел. (В действительности большинство слагаемых исчезает, но количество неисчезающих слагаемых остается экспоненциально большим.)

6.2.3. Универсальные квантовые вентили

Мы должны обратиться к еще одному фундаментальному вопросу, касающемуся квантовых вычислений, как построить адекватный набор квантовых вентилей? Другими словами, что образует универсальный квантовый компьютер?

Ответ вам понравится. Для реализации универсальных квантовых вычислений достаточно любого типичного двухкубитового вентиля. То есть, если мы можем применять эти вентили к любой паре кубитов, то любого из них, кроме множества меры нуль унитарных 4×4-матриц, достаточно, чтобы построить n-кубитовую схему, вычисляющую преобразование, сколь угодно близкое к любому элементу $U(2^n)$.

Математически это не особенно глубокий результат, но с физической точки зрения он очень интересен. Это означает, что в квантовом мире, пока мы можем придумывать типичные двухкубитовые взаимодействия и осуществлять их точно между любыми двумя кубитами, мы в состоянии вычислять что угодно, независимо от сложности. Нетривиальные вычисления в квантовой теории встречаются повсюду.

Кроме этого общего результата, интересно продемонстрировать и конкретные наборы универсальных вентилей, которые очень легко могут быть реализованы физически. Обсудим несколько примеров.

Существует несколько основных элементов, входящих в состав любого набора универсальных квантовых вентилей.

(1) Степени типичного вентиля. Рассмотрим «типичный» k-битовый вентиль. Это унитарная $2^k \times 2^k$-матрица $\mathbf{U}$ с собственными значениями $e^{i\theta_1}$, $e^{i\theta_2}, \ldots e^{i\theta_{2^k}}$. Для всех, кроме множества меры нуль, таких матриц каждое θ_i представляет собой иррациональное число, кратное π, и все θ_i несоизмеримы (каждое θ_i/θ_j тоже иррационально). Положительная целая степень $\mathbf{U}^n$ матрицы $\mathbf{U}$ имеет собственные значения

$$e^{in\theta_1}, \quad e^{in\theta_2}, \quad \ldots \quad e^{in\theta_{2^k}}. \tag{6.82}$$

Каждый такой список собственных значений определяет точку на 2^k-мерном торе (произведении 2^k окружностей). Так как n принимает целые положительные значения, эти точки плотно заполняют весь тор, если $\mathbf{U}$ является типичной. Если $\mathbf{U} = e^{i\mathbf{A}}$, то для любого вещественного λ положительные целые степени $\mathbf{U}$ сколь угодно близки к $\mathbf{U}(\lambda) = e^{i\lambda\mathbf{A}}$. Мы утверждаем, что любое $\mathbf{U}(\lambda)$ *достигается* положительными целыми степенями $\mathbf{U}$.

(2) Переключение входов и выходов. Имеется несколько (классических) преобразований, которые можно выполнить, всего лишь переставив метки k кубитов или, другими словами, применяя вентиль $\mathbf{U}$ к кубитам в другом порядке. Из $(2^k)!$ перестановок строк длиной k путем обмена кубитами можно реализовать $k!$ Если вентилем, применяемым к k кубитам в стандартном порядке является $\mathbf{U}$, а P – перестановка, осуществляемая путем обмена кубитами, то мы можем построить вентиль

$$\mathbf{U}' = P\mathbf{U}P^{-1} \tag{6.83}$$

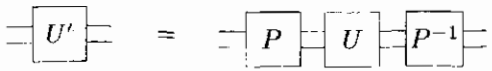
только с помощью переключения входов и выходов исходного вентиля. Например, обмен двумя кубитами осуществляет перестановку

$$P : |01\rangle \leftrightarrow |10\rangle \tag{6.84}$$

или

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{6.85}$$

действующую в базисе $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Переключая входы и выходы, мы получаем вентиль



Мы можем также построить любую целую положительную степень $\mathbf{U}'$: $\left(P\mathbf{U}P^{-1}\right)^n = P\mathbf{U}^nP^{-1}$.

(3) Замкнутая алгебра Ли. Мы уже замечали, что если вентиль $\mathbf{U} = e^{i\mathbf{A}}$ является типичным, то его степени плотны на торе $\left\{e^{i\lambda\mathbf{A}}\right\}$. Мы можем далее доказать, что если $\mathbf{U} = e^{i\mathbf{A}}$ и $\mathbf{U}' = e^{i\mathbf{B}}$ типичные вентили, то для любых вещественных α, β, γ из них можно составить вентиль, сколь угодно близкий к

$$e^{i(\alpha\mathbf{A}+\beta\mathbf{B})} \quad \text{или} \quad e^{-\gamma[\mathbf{A},\mathbf{B}]}. \tag{6.86}$$

Таким образом, «достижимые» преобразования образуют замкнутую *алгебру Ли*. Мы говорим, что $\mathbf{U} = e^{i\mathbf{A}}$ генерируется преобразованием $\mathbf{A}$; тогда если $\mathbf{A}$ и $\mathbf{B}$ являются типичными генераторами достижимых преобразований, то этим же свойством обладают их вещественные линейные комбинации и (умноженный на i) коммутатор.

Сначала заметим, что

$$\lim_{n\to\infty}\left(e^{i\alpha\mathbf{A}/n}e^{i\beta\mathbf{B}/n}\right)^n = \lim_{n\to\infty}\left(1+\frac{i}{n}(\alpha\mathbf{A}+\beta\mathbf{B})\right)^n = e^{i(\alpha\mathbf{A}+\beta\mathbf{B})}. \tag{6.87}$$

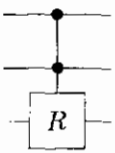
Следовательно, любое преобразование $e^{i(\alpha\mathbf{A}+\beta\mathbf{B})}$ достижимо, если таковыми являются $e^{i\alpha\mathbf{A}/n}$ и $e^{i\beta\mathbf{B}/n}$. Более того,

$$\lim_{n\to\infty}\left(e^{i\mathbf{B}/\sqrt{n}}e^{-i\mathbf{A}/\sqrt{n}}e^{-i\mathbf{B}/\sqrt{n}}e^{i\mathbf{A}/\sqrt{n}}\right)^n = \\ = \lim_{n\to\infty}\left[1-\frac{1}{n}(\mathbf{AB}-\mathbf{BA})\right]^n = e^{-[\mathbf{A},\mathbf{B}]}, \tag{6.88}$$

так что $e^{-[\mathbf{A},\mathbf{B}]}$ также достижимо.

Применяя результаты (1), (2) и (3), можно показать, что типичный двухкубитовый вентиль является универсальным.

1. Вентиль Дойча. Первым, обратившим внимание на существование универсального квантового вентиля, был Дэвид Дойч (1989 г.). Трехкубитовый универсальный вентиль Дойча является квантовым кузеном вентиля Тоффоли. Это дважды контролируемое $\mathbf{R}$-преобразование



которое применяет $\mathbf{R}$ к третьему кубиту, если два первых имеют значение, равное единице; в противном случае — действует тривиально. Здесь

$$\mathbf{R} = -i\mathbf{R}_x(\theta) = -i\exp\left(i\frac{\theta}{2}\boldsymbol{\sigma}_x\right) = -i\left(\cos\frac{\theta}{2} + i\boldsymbol{\sigma}_x\sin\frac{\theta}{2}\right) \tag{6.89}$$

представляет, с точностью до фазы, поворот на θ вокруг оси x, где θ некоторый несоизмеримый с π угол.

n-ая степень вентиля Дойча представляет собой дважды контролируемое $\mathbf{R}^n$. В частности, $\mathbf{R}^4 = \mathbf{R}_x(4\theta)$, так что все однокубитовые преобразования, генерируемые $\boldsymbol{\sigma}_x$, достигаются целыми степенями $\mathbf{R}$. Более того, его $(4n+1)$-ой степенью является преобразование

$$-i\left[\cos\frac{(4n+1)\theta}{2} + i\boldsymbol{\sigma}_x\sin\frac{(4n+1)\theta}{2}\right], \tag{6.90}$$

сколь угодно близкое к $\boldsymbol{\sigma}_x$. Следовательно, вентиль Тоффоли достигается целыми степенями вентиля Дойча, то есть вентиль Дойча является универсальным для классических вычислений.

Действуя на трехкубитовый вычислительный базис

$$\{|000\rangle,\ |001\rangle,\ |010\rangle,\ |011\rangle,\ |100\rangle,\ |101\rangle,\ |110\rangle,\ |111\rangle\}, \tag{6.91}$$

генератор вентиля Дойча переставляет два его последних элемента:

$$|110\rangle \leftrightarrow |111\rangle. \tag{6.92}$$

Изобразим эту 8×8-матрицу как

$$(\boldsymbol{\sigma}_x)_{67} = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & \boldsymbol{\sigma}_x \end{array}\right). \tag{6.93}$$

С помощью вентиля Тоффоли можно выполнить перестановку любых этих восьми элементов, в частности, для любых m и n

$$P = (6m)(7n). \tag{6.94}$$

Следовательно, нам доступно любое преобразование, генерируемое

$$P(\boldsymbol{\sigma}_x)_{67}P^{-1} = (\boldsymbol{\sigma}_x)_{mn}. \tag{6.95}$$

Более того,

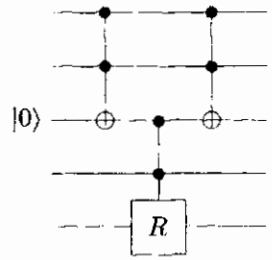
$$\left[(\sigma_x)_{56},(\sigma_x)_{67}\right]=\left[\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}\right]=\begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix}=i(\sigma_y)_{57}. \tag{6.96}$$

Аналогично мы можем реализовать любое унитарное преобразование, генерируемое $(\sigma_y)_{mn}$. Наконец,

$$\left[(\sigma_x)_{mn},(\sigma_y)_{mn}\right]=2i(\sigma_z)_{mn}. \tag{6.97}$$

Следовательно, нам доступно любое преобразование, генерируемое линейной комбинацией матриц $(\sigma_{x,y,z})_{mn}$. Они образуют линейную оболочку алгебры Ли $SU(8)$, следовательно, мы можем генерировать любое трехкубитовое унитарное преобразование (за исключением несущественной общей фазы).

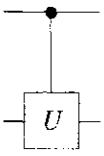
Вспомним теперь, что мы уже обнаружили, что, комбинируя трехбитовые вентили Тоффоли, можно построить n-битовый вентиль Тоффоли. Схема



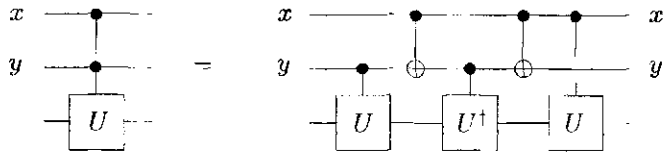
использует один вспомогательный бит, чтобы построить четырехбитовый вентиль Дойча (трижды контролируемое **R**) из одного трехбитового вентиля Дойча и двух трехбитовых вентилей Тоффоли. Аналогичная схема реализует n-битовый вентиль Дойча из одного трехбитового вентиля Дойча и двух $n-1$-битовых вентилей Тоффоли. Коль скоро мы имеем n-битовый вентиль Дойча, а также универсальное классическое вычисление, точно те же аргументы, что и выше, показывают, что можно реализовать любое преобразование из $SU(2^n)$.

2. Универсальные двухкубитовые вентили. Мы видели, что для универсальности классических обратимых вычислений необходимы трехбитовые универсальные вентили. Однако в квантовых вычислениях оказываются адекватными двухкубитовые вентили. Поскольку мы уже знаем, что вентиль Дойча универсален, мы можем установить это, показав, что он может быть образован комбинацией двухкубитовых вентилей.

Фактически, если



обозначает вентиль контролируемое U (унитарное 2×2-преобразование U применяется ко второму кубиту, если первый имеет значение, равное единице; в противном случае вентиль действует тривиально), то вентиль дважды контролируемое U получается с помощью схемы



Степенью U, примененной к третьему кубиту, является

$$y - (x \oplus y) + x - x + y - (x + y - 2xy) = 2xy. \tag{6.98}$$

Следовательно, вентиль Дойча можно построить из вентилей контролируемого U, контролируемого U^{-1} и контролируемого NOT, где

$$\mathrm{U}^2 = -i\mathrm{R}_x(\theta)\,; \tag{6.99}$$

мы можем выбрать

$$\mathrm{U} = e^{-i\pi/4}\mathrm{R}_x\left(\frac{\theta}{2}\right). \tag{6.100}$$

Так как положительные степени U сколь угодно близко приближаются к σ_x и U^{-1}, то вентиль Дойча можно сконструировать из одного лишь контролироуемого U. Следовательно, при иррациональном θ/π контролируемое $e^{-i\pi/4}\mathrm{R}_x\left(\frac{\theta}{2}\right)$ само по себе является универсальным вентилем.

(Заметим, что приведенная выше конструкция показывает, что, несмотря на то, что мы не можем построить вентиль Тоффоли из классических двухбитовых обратимых вентилей, его *можно* сконструировать из контролируемого «квадратного корня из NOT», то есть контролируемого U, квадрат которого $\mathrm{U}^2 = \sigma_x$.)

3. Типичные двухбитовые вентили. Итак, мы нашли конкретные двухбитовые вентили (контролируемые повороты), являющиеся универсальными. Следовательно, для универсальности вполне достаточно, если мы можем строить плотные в $U(4)$ преобразования, действующие на пары кубитов.

Однако на самом деле достаточно любого типичного двухкубитового вентиля, чтобы генерировать все преобразования из $U(4)$. Как мы видели, если $e^{i\mathbf{A}}$ – типичный элемент $U(4)$, то можно реализовать любое преобразование, генерируемое $\mathbf{A}$. Более того, можно реализовать любые преобразования, генерируемые элементом минимальной алгебры Ли, содержащей $\mathbf{A}$ и

$$\mathbf{B} = P\mathbf{A}P^{-1}, \tag{6.101}$$

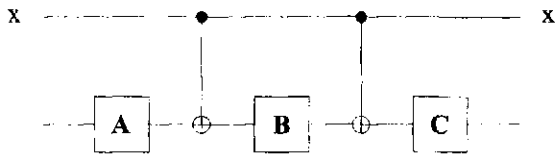
где P – перестановка ($|01\rangle \leftrightarrow |10\rangle$), получаемая переключением входов и выходов.

Рассмотрим теперь общее преобразование $\mathbf{A}$ [разложенное в базисе алгебры Ли $U(4)$], а также рассмотрим конкретную схему построения 16-ти элементов алгебры Ли путем последовательных коммутаций исходя из $\mathbf{A}$ и $\mathbf{B}$. Конструируемые таким образом элементы линейно независимы [а отсюда следует, что любое преобразование в $U(4)$ достижимо], если определитель конкретной 16×16-матрицы не равен нулю. Если этот определитель не обращается в нуль тождественно, то его нули появляются только на подмногообразии меры нуль. Фактически мы можем выбрать, допустим,

$$\mathbf{A} = (\alpha \mathbf{1} + \beta \boldsymbol{\sigma}_x + \gamma \boldsymbol{\sigma}_y)_{23} \tag{6.102}$$

(при несоизмеримых α, β, γ) и с помощью явных вычислений показать, что действительно, начиная с $\mathbf{A}$ и $\mathbf{B}$, последовательными коммутациями можно генерировать всю 16-мерную алгебру Ли. Следовательно, мы приходим к заключению, что неуспех генерирования всей алгебры $U(4)$ нетипичен, и обнаруживаем, что почти все двухкубитовые вентили универсальны.

4. Другие достаточные наборы вентилей. Очевидно также, что универсальные квантовые вычисления можно реализовать с помощью набора вентилей, состоящих из *классических* многокубитовых и квантовых однокубитовых вентилей. Например, можно увидеть, что универсальный набор образуется вентелем XOR, комбинируемым с однокубитовыми вентилями. Рассмотрим схему



применяющую ко второму кубиту преобразование $\mathbf{ABC}$, если $x = 0$, и $\mathbf{A}\sigma_x\mathbf{B}\sigma_x\mathbf{C}$, если $x = 1$. Если мы можем подобрать такие $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, что

$$\begin{aligned} \mathbf{ABC} &= \mathbf{1}, \\ \mathbf{A}\sigma_x\mathbf{B}\sigma_x\mathbf{C} &= \mathbf{U}, \end{aligned} \tag{6.103}$$

тогда эта схема функционирует как вентиль контролируемое $\mathbf{U}$. Фактически для любого унитарного $\mathbf{U}$ с единичным определителем существуют унитарные 2×2-преобразования $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ с такими свойствами (как вы покажете в упражнении). Следовательно, XOR в совокупности с произвольными однокубитовыми преобразованиями образуют универсальный набор. Конечно, двух типичных (некоммутирующих) однокубитовых преобразований достаточно, чтобы добиться чего угодно. В действительности с помощью XOR-а и *единственного* типичного однокубитового поворота мы можем построить второй однокубитовый поворот, не коммутирующий с первым. Таким образом, XOR вместе со всего лишь одним однокубитовым вентилем образует универсальный набор вентилей.

Если мы способны реализовать вентиль Тоффоли, тогда для универсальных вычислений достаточно даже некоторых нетипичных однокубитовых преобразований. Например (еще одно упражнение), вентиль Тоффоли совместно с поворотами на $\pi/2$ вокруг осей x и z представляет собой универсальный набор.

5. Точность. Наше обсуждение универсальности сфокусировалось на *достижимости*, оставив без внимания *сложность*. Мы всего лишь установили, что можем построить квантовую схему, сколь угодно близкую к требуемому элементу из $U(2^n)$, но не рассмотрели размер необходимой нам схемы. Однако с точки зрения теории квантовой сложности универсальность очень важна, поскольку она означает, что с приемлемой точностью и разумным замедлением один квантовый компьютер может моделировать другой.

В действительности до сих пор мы были не очень точны в вопросе о том, что означает для одного унитарного преобразования быть «близким» к другому; для этого следует определить топологию. Одна возможность представляет собой использование той же нормы, что и в предыдущем обсуждении точности. Тогда расстоянием между матрицами $\mathbf{U}$ и $\mathbf{W}$ является $\|\mathbf{U} - \mathbf{W}\|$. Еще одна естественная топология связана с внутренним произведением

$$\langle \mathbf{W}|\mathbf{U}\rangle = \operatorname{tr} \mathbf{W}^\dagger\mathbf{U} \tag{6.104}$$

(если $\mathbf{U}$ и $\mathbf{W}$ являются $N \times N$-матрицами, то это в точности обычное внутреннее произведение в $\mathbb{C}^{N^2}$, в котором $\mathbf{U}$ рассматривается как N^2-компонентный вектор). Тогда мы можем определить квадрат расстояния между матрицами как

$$\|\mathbf{U} - \mathbf{W}\|^2 = \langle \mathbf{U} - \mathbf{W} | \mathbf{U} - \mathbf{W} \rangle. \tag{6.105}$$

Для анализа сложности подходит практически любая разумная топология.

Решающим моментом является то, что, имея любой универсальный набор вентилей, мы можем подойти на расстояние ε к любому желаемому преобразованию, действующему на фиксированное количество кубитов, используя квантовую схему, размер которой ограничен сверху полиномиально по ε^{-1}. Следовательно, один универсальный квантовый компьютер может моделировать другой с точностью ε и не хуже, чем с полиномиальным по ε^{-1} фактором замедления. Теперь нам уже понятно: чтобы иметь высокую вероятность получения правильного ответа при выполнении квантовой схемы размера T, необходимо обеспечить выполнение каждого квантового вентиля с точностью порядка T^{-1}. Следовательно, если вы имеете семейство квантовых схем полиномиального размера, которые выполняет ваш квантовый компьютер, то я могу изобрести семейство схем полиномиального размера, которые выполняет моя машина и с приемлемой точностью эмулирует вашу.

Почему схема $\mathrm{poly}(\varepsilon^{-1})$-размера может достичь данного k-кубитового преобразования $\mathbf{U}$ в пределах расстояния ε? Мы знаем, например, что положительные целые степени типичного k-кубитового $e^{i\mathbf{A}}$ плотны на 2^k-торе $\left\{e^{i\lambda\mathbf{A}}\right\}$. Область тора в пределах расстояния ε до любой заданной точки имеет объем порядка ε^{2^k}. Следовательно, с помощью $\left(e^{i\mathbf{A}}\right)^n$ при некотором целом n порядка ε^{-2^k} мы можем асимптотически (при достаточно малом ε) достичь любого преобразования $\left\{e^{i\lambda\mathbf{A}}\right\}$ с точностью не хуже ε. Нам также известно, что, используя схемы фиксированного размера (независимого от ε), мы можем получить преобразования $\left\{e^{i\mathbf{A}_a}\right\}$, где $\mathbf{A}_a$ образуют линейную оболочку полной алгебры Ли $U(2^k)$. Тогда также с полиномиальной сходимостью мы можем аппроксимировать любое $\exp\left(i\sum_a \alpha_a \mathbf{A}_a\right)$, как в уравнении (6.87).

В принципе мы способны добиться гораздо лучшего результата, достигая желаемого k-кубитового унитарного преобразования с точностью не хуже ε с помощью только $\mathrm{poly}\left(\log(\varepsilon^{-1})\right)$ квантовых вентилей. Так как количество схем размера T, которые мы можем построить, действуя на k кубитов, экспоненциально по T, а схемы заполняют $U(2^k)$ примерно однородно,

то должна существовать схема размера T, достигающая в пределах расстояния порядка ϵ^{-T} любой точки в $U(2^k)$. Однако это может оказаться трудной вычислительной задачей – *классическим способом* разработать схему, экспоненциально близко подходящую к унитарному преобразованию, которого мы пытаемся достичь. Поэтому было бы нечестно опираться на эту более эффективную конструкцию в асимптотическом анализе квантовой сложности.

6.3. Некоторые квантовые алгоритмы

Хотя мы по-прежнему не в состоянии показать, что $BPP \neq BQP$, существует три подхода, которым можно последовать, чтобы изучить различия между возможностями классических и квантовых компьютеров.

(1) Неэкспоненциальное ускорение. Мы можем найти квантовые алгоритмы, которые заметно быстрее лучших классических алгоритмов, но *не экспоненциально* быстрее. Эти алгоритмы не проливают свет на общепринятую классификацию сложности. Но они демонстрируют характер разделения между задачами, которые могут выполнять классические и квантовые компьютеры. Пример: гроверовское квантовое ускорение поиска в неструктурированной базе данных.

(2) «Релятивизированное» экспоненциальное ускорение. Мы можем рассмотреть проблему анализа содержимого «квантового черного ящика». Ящик выполняет *a priori* неизвестное унитарное преобразование. Мы можем приготовить для него входные данные и измерить его результат; наша задача – определить, что делает ящик. Оказывается возможным доказать, что существуют квантовые ящики (специалисты по теории вычислений называют их оракулами[1]), обладающие следующим свойством: загружая ящик квантовыми суперпозициями, можно узнать, что находится внутри него, с *экспоненциальным* ускорением по сравнению с тем, как много времени пришлось бы потратить, если бы нам были разрешены только классические входные данные. Специалист по теории вычислений сказал бы, что $BPP \neq BQP$ «относительно оракула». Пример: саймоновское экспоненциальное квантовое ускорение отыскания периода функции «2 в 1».

[1]Термин «оракул» означает, что ящик отвечает на вопрос *немедленно*; то есть время, затрачиваемое на его работу, не включается в анализ сложности.

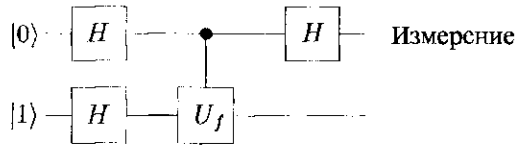
(3) Экспоненциальное ускорение для «по-видимому» трудных задач. Мы можем продемонстрировать квантовый алгоритм, решающий в течение полиномиального времени задачу, которая с классической точки зрения выглядит сложной, то есть серьезно подозревается (хотя и не доказано), что эта задача не принадлежит BPP. Пример: алгоритм факторизации Шора.

1. Проблема Дойча. Мы обсудим примеры из всех трех подходов. Но для начала разомнемся, вспомнив пример простого квантового алгоритма, который предварительно обсуждался в разделе 1.5: алгоритм Дойча для различения между постоянной и сбалансированной функциями $f: \{0,1\} \to \{0,1\}$. Нам предоставлен квантовый черный ящик, вычисляющий $f(x)$; то есть приводящий в действие двухкубитовое унитарное преобразование

$$U_f : |x\rangle|y\rangle \to |x\rangle|y \oplus f(x)\rangle, \tag{6.106}$$

которое инвертирует второй кубит, если $f(\text{первый кубит}) = 1$. Наша задача состоит в том, чтобы определить, выполняется ли $f(0) = f(1)$. Если мы ограничены «классическими» входными данными $|0\rangle$ и $|1\rangle$, то, чтобы получить ответ, нам необходимо обратиться к ящику дважды ($x = 0$ и $x = 1$). Но если нам позволено ввести когерентную суперпозицию этих «классических» состояний, то достаточно одного раза.

Квантовой схемой, решающей эту проблему (обсуждавшуюся в разделе 1.5), является



Здесь H обозначает преобразование Адамара

$$\mathbf{H} : |x\rangle \to \frac{1}{\sqrt{2}} \sum_y (-1)^{xy} |y\rangle \tag{6.107}$$

или

$$\begin{aligned} \mathbf{H} : |0\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ |1\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle); \end{aligned} \tag{6.108}$$

то есть $\mathbf{H}$ представляет собой 2×2-матрицу

$$\mathbf{H} : \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{6.109}$$

Схема преобразует вход $|0\rangle|1\rangle$ в

$$\begin{aligned} |0\rangle|1\rangle &\rightarrow \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) \\ &\rightarrow \frac{1}{2}\Big[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\Big](|0\rangle - |1\rangle) \\ &\rightarrow \frac{1}{2}\left\{\Big[(-1)^{f(0)} + (-1)^{f(1)}\Big]|0\rangle \right. \\ &\qquad \left. + \Big[(-1)^{f(0)} - (-1)^{f(1)}\Big]|1\rangle\right\} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \tag{6.110}$$

Тогда при измерении первого кубита с вероятностью единица будет получен результат $|0\rangle$, если $f(0) = f(1)$ (постоянная функция), и с вероятностью единица — результат $|1\rangle$, если $f(0) \neq f(1)$ (сбалансированная функция).

Квантовый компьютер обладает преимуществом перед классическим компьютером, поскольку он может привлечь *квантовый параллелизм*. Так как мы вводим суперпозицию состояний $|0\rangle$ и $|1\rangle$, выход чувствителен к обоим значениям $f(0)$ и $f(1)$, даже если мы обратились к ящику только один раз.

2. Проблема Дойча–Йожы. Рассмотрим теперь некоторые обобщения проблемы Дойча. По-прежнему будем предполагать, что нам нужно анализировать квантовый черный ящик («квантовый оракул»). Но в надежде узнать что-нибудь о сложности мы будем представлять, что имеем семейство черных ящиков с переменным размером входа. Нас интересует, как время, необходимое для определения того, что происходит внутри ящика, зависит от размера входа (где «время» измеряется тем, сколько раз мы обращаемся к ящику с вопросом).

В *задаче Дойча–Йожы* нам предоставлен квантовый черный ящик, который вычисляет функцию, преобразуя n битов в один:

$$f : \{0,1\}^n \rightarrow \{0,1\}, \tag{6.111}$$

причем у нас есть все основания полагать, что f — постоянная [$f(x) = c$ для всех x] или сбалансированная [$f(x) = 0$ для ровно половины возможных значений входа]. Мы должны решить проблему принятия решения: является ли f постоянной или сбалансированной?

Фактически, используя ту же схему, что и для решения проблемы Дойча (но с x, расширенным от одного до n битов), мы также можем решить и эту проблему, обращаясь к ящику только один раз. Заметим, что если n вентилей Адамара параллельно применяются к n кубитам

$$\mathbf{H}^{(n)} = \mathbf{H} \otimes \mathbf{H} \otimes \ldots \otimes \mathbf{H}, \tag{6.112}$$

то n-кубитовое состояние преобразуется как

$$\mathbf{H}^{(n)}\colon |x\rangle \to \prod_{i=1}^{n} \left[\frac{1}{\sqrt{2}} \sum_{y_i=\{0,1\}} (-1)^{x_i y_i} |y_i\rangle \right] \equiv \frac{1}{2^{n/2}} \sum_{y=0}^{2^n-1} (-1)^{x\cdot y} |y\rangle, \tag{6.113}$$

где x, y представляют n-битовые строки, а $x \cdot y$ обозначает *побитовое* AND (или скалярное произведение по модулю два):

$$x \cdot y = (x_1 \wedge y_1) \oplus (x_2 \wedge y_2) \oplus \ldots \oplus (x_n \wedge y_n). \tag{6.114}$$

Действуя на вход $(|0\rangle)^n|1\rangle$, схема преобразует его следующим образом:

$$\begin{aligned} (|0\rangle)^n|1\rangle &\to \left(\frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle \right) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ &\to \left(\frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle \right) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ &\to \left(\frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} (-1)^{f(x)} (-1)^{x\cdot y} |y\rangle \right) \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \tag{6.115}$$

Теперь вычислим сумму

$$\frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)} (-1)^{x\cdot y}. \tag{6.116}$$

Если f — постоянная функция, то эта сумма равна

$$(-1)^{f(x)} \left(\frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{x\cdot y} \right) = (-1)^{f(x)} \delta_{y,0}; \tag{6.117}$$

она обращается в нуль, за исключением случая, когда $y = 0$. Следовательно, при измерении n-битового выходного регистра с вероятностью единица будет получен результат $|y = 0\rangle \equiv (|0\rangle)^n$. Но если функция f сбалансирована, то при $y = 0$ сумма (6.116) становится равной

$$\frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{f(x)} = 0 \tag{6.118}$$

[поскольку половина слагаемых равны $(+1)$, а другая половина $-(-1)$]. Следовательно, вероятность получения результата измерения $|y = 0\rangle$ равна нулю.

Мы приходим к выводу, что квантовому оракулу достаточно одного вопроса, чтобы со 100% уверенностью различить постоянную и сбалансированную функции. Результат измерения $y = 0$ означает, что f — постоянная, любой другой результат — сбалансированная.

Итак, квантовое вычисление изящно решает эту задачу, но действительно ли это трудная проблема с классической точки зрения? Ограничиваясь вводом классических состояний $|x\rangle$, мы можем задавать вопрос оракулу неоднократно, всякий раз выбирая ввод x случайным образом (без возврата). Как только будут получены различные ответы на два различных вопроса, мы определим, что функция сбалансирована (не постоянная). Но если функция фактически является постоянной, мы не будем *уверены* в том, что это действительно так, до тех пор пока не предложим $2^{n-1} + 1$ вопросов, получая всякий раз один и тот же ответ. В противоположность этому квантовое вычисление дает определенный ответ всего лишь в один прием. В этом смысле (если мы требуем абсолютной определенности) классическое вычисление требует экспоненциального по n количества вопросов, тогда как квантовое вычисление — нет, следовательно, можно говорить об экспоненциальном ускорении.

Но может быть неразумно требовать абсолютной определенности от классического вычисления (в частности, так как любой реальный компьютер подвержен ошибкам, то и квантовый компьютер также будет не способен достигать абсолютной надежности). Допустим, что нас удовлетворяет предположение о сбалансированности или постоянстве с вероятностью успеха

$$P(\text{success}) > 1 - \varepsilon. \tag{6.119}$$

Если функция действительно сбалансирована, то вероятность получения всякий раз одного и того же ответа на k заданных вопросов равна $p = 2^{-(k-1)}$. Если после получения одного и того же ответа k раз под-

ряд мы сделаем предположение, что функция постоянна, быстрый бейесовский анализ показывает, что вероятность того, что наша догадка ошибочна, равна $\frac{1}{2^{k-1}+1}$ (в предположении, что сбалансированность и постоянство *a priori* равновероятны). Итак, если мы высказываем догадку после k вопросов, то вероятность ее ошибочности

$$1 - P(\text{success}) = \frac{1}{2^{k-1}(2^{k-1}+1)}. \tag{6.120}$$

Следовательно, мы можем достичь вероятности успеха $1-\varepsilon$ при $\varepsilon^{-1} = 2^{k-1}(2^{k-1}+1)$ или при $k \sim \frac{1}{2}\log\frac{1}{\varepsilon}$. А так как экспоненциально высокая вероятность успеха достигается с помощью полиномиального количества попыток, то на самом деле незаконно говорить, что проблема является трудной.

3. Задача Бернштейна–Вазирани. Точно такая же схема может быть использована для решения другого варанта задачи Дойча–Йожы. Предположим, что наш квантовый черный ящик вычисляет одну из функций f_a, где

$$f_a(x) = a \cdot x, \tag{6.121}$$

а a представляет собой n-битовую строку. Наше задача — определить a.

Квантовый алгоритм может с определенностью решить эту задачу, получив только один (n-кубитовый) квантовый вопрос. Для этой конкретной функции квантовое состояние в уравнении (6.115) имеет вид

$$\frac{1}{2^n}\sum_{x=0}^{2^n-1}\sum_{y=0}^{2^n-1}(-1)^{a\cdot x}(-1)^{x\cdot y}|y\rangle. \tag{6.122}$$

Но фактически

$$\frac{1}{2^n}\sum_{x=0}^{2^n-1}(-1)^{a\cdot x}(-1)^{x\cdot y} = \delta_{a,y}, \tag{6.123}$$

то есть этим состоянием является $|a\rangle$. Мы можем выполнить схему один раз и измерить n-кубитовый регистр, обнаружив с вероятностью единица n-битовую строку a.

Если разрешены только классические вопросы, то на каждый из них мы получаем только один бит информации и для определения значения a требуется n вопросов. Следовательно, мы имеем четкую границу между квантовой и классической сложностью задачи. Правда, этот пример не

вскрывает соотношения между BPP и BQP, поскольку классическая задача не является трудной. Количество вопросов, необходимых с классической точки зрения, всего лишь линейно, а не экспоненциально по размеру входа.

4. Задача Саймона. Бернштейну и Вазирани удалось сформулировать вариант предыдущей задачи, который *является* классически трудным, и, таким образом, впервые установить «релятивизированную» границу между квантовой и классической сложностью. Мы найдем более поучительным рассмотреть более простой пример, несколько позднее предложенный Даниэлем Саймоном.

Снова нам предоставлен квантовый черный ящик, и на этот раз мы уверены в том, что он вычисляет функцию «2 в 1»

$$f: \{0,1\}^n \rightarrow \{0,1\}^n. \tag{6.124}$$

Более того, функция имеет период, определяемый n-битовой строкой a; то есть

$$f(x) = f(y), \quad \text{если} \quad y = x \oplus a, \tag{6.125}$$

где $\oplus$ — побитовая XOR-операция. [То есть a является периодом, если мы рассматриваем x принимающим значения из $(Z_2)^n$, а не из Z_{2^n}.[1]] Это все, что нам известно об f. Наша задача — определить значение a.

Эта задача классически *трудная*. Нам необходимо обратиться к оракулу экспоненциально большое количество раз, чтобы иметь какую-нибудь разумную вероятность определения a. Мы ничего не узнаем, пока нам не повезет выбрать два вопроса x и y, которые случайно окажутся удовлетворяющими $x \oplus a = y$. Допустим, например, что мы выбираем $2^{n/4}$ вопросов. Количество пар вопросов меньше чем $\left(2^{n/4}\right)^2$, и для каждой пары $\{x, y\}$ вероятность того, что $x \oplus a = y$, равна 2^{-n}. Следовательно, вероятность успешного отыскания a меньше, чем

$$2^{-n}\left(2^{n/4}\right)^2 = 2^{-n/2}; \tag{6.126}$$

даже при экспоненциально большом количестве вопросов вероятность успеха экспоненциально мала.

Если угодно, эту задачу можно сформулировать как проблему принятия решения: функция f является или одно-однозначной (1 в 1), или отображает два в одно (2 в 1) с некоторым случайно выбранным периодом a;

[1] $(Z_2)^n$ - группа, элементами которой являются двоичные строки длины n. Групповая операция представляет собой побитовое сложение по модулю 2. Z_{2^n} — группа остатков от сложения по модулю 2^n. — *Прим. ред.*

обе эти возможности имеют априорные вероятности $1/2$. Нам нужно определить, является ли функция 1 в 1 или 2 в 1. Тогда после $2^{n/4}$ классических вопросов вероятность корректной догадки удовлетворяет неравенству

$$P(\text{success}) < \frac{1}{2} + \frac{1}{2^{n/2}} \tag{6.127}$$

и не удаляется от $1/2$ при больших значениях n.

Но для квантовых вопросов проблема является простой! Используемая нами схема, по существу, та же, что и выше, но теперь *оба* регистра расширены до n кубитов. Мы готовим равновзвешенную суперпозицию всех n-битовых строк (действуя на $|0\rangle$ преобразованием $\mathbf{H}^{(n)}$), а затем обращаемся к оракулу:

$$U_f : \left(\sum_{x=0}^{2^n-1} |x\rangle \right) |0\rangle \rightarrow \sum_{x=0}^{2^n-1} |x\rangle |f(x)\rangle. \tag{6.128}$$

Теперь мы измеряем второй регистр. (Этот этап на самом деле не обязателен, но для ясности изложения я включаю его сюда.) Результатом измерения является значение, случайно выбранное из 2^n равновероятных значений $f(x)$. Допустим, результатом является $f(x_0)$. Тогда, поскольку оба значения, x_0 и $x_0 \oplus a$, и только они отображаются функцией f на $f(x_0)$, мы приготовили состояние

$$\frac{1}{\sqrt{2}}(|x_0\rangle + |x_0 \oplus a\rangle) \tag{6.129}$$

в первом регистре.

Теперь мы хотим извлечь некоторую информацию относительно a. Очевидно, что на этом этапе было бы бесполезно измерять регистр (в вычислительном базисе). Мы получили бы результат x_0 или $x_0 \oplus a$ с вероятностью $1/2$ каждый, но ни тот, ни другой ничего не сказал бы о значении a.

Но представим теперь, что непосредственно перед измерением мы применили к регистру преобразование Адамара $\mathbf{H}^{(n)}$:

$$\begin{aligned}
\mathbf{H}^{(n)} : \frac{1}{\sqrt{2}} (|x_0\rangle + |x_0 \oplus a\rangle) &\rightarrow \\
&\rightarrow \frac{1}{2^{(n-1)/2}} \sum_{y=0}^{2^n-1} \left[(-1)^{x_0 \cdot y} + (-1)^{(x_0 \oplus a) \cdot y} \right] |y\rangle = \\
&= \frac{1}{2^{(n-1)/2}} \sum_{a \cdot y = 0} (-1)^{x_0 \cdot y} |y\rangle.
\end{aligned} \tag{6.130}$$

Если $a \cdot y = 1$, то слагаемые в коэффициенте перед $|y\rangle$ интерферируют деструктивно. Следовательно, в сумме по y выживают только состояния с $a \cdot y = 0$. Тогда результатом измерения является случайным образом выбранное из всех возможных значений y, появляющихся с вероятностью $2^{-(n-1)}$, таких, что $a \cdot y = 0$.

Мы многократно повторяем этот алгоритм, получая всякий раз еще одно значение y, удовлетворяющее $a \cdot y = 0$. Как только мы найдем n таких линейно независимых значений $\{y_1, y_2, y_3, \ldots y_n\}$ [то есть линейно независимых над $(Z_2)^n$], мы можем решить уравнения

$$\begin{aligned} y_1 \cdot a &= 0, \\ y_2 \cdot a &= 0, \\ &\vdots \\ y_n \cdot a &= 0, \end{aligned} \tag{6.131}$$

чтобы определить единственное значение a, и, таким образом, решить поставленную задачу. Нетрудно видеть, что с помощью $O(n)$ повторений мы можем достичь вероятности успеха, экспоненциально близкой к единице.

Итак, наконец-то мы нашли пример задачи, которую можно решить за полиномиальное время, используя квантовые суперпозиции для данного частного типа оракула, тогда как если ограничиться классическими вопросами, то для этого потребуется экспоненциальное время. Специалист по теории вычислений мог бы сказать:

Существует оракул, относительно которого $BQP \neq BPP$

Заметим, что всякий раз, когда мы сравниваем классическую и квантовую сложность относительно оракула, мы рассматриваем квантовый оракул (вопросами и ответами являются состояния в гильбертовом пространстве), но с выделенным ортонормированным базисом. Если мы предлагаем классический вопрос (элемент выделенного базиса), то всегда получаем классический ответ (другой элемент базиса). Проблема в том, можем ли мы достичь существенного ускорения, выбирая более общие, квантовые, вопросы.

6.4. Квантовый поиск в базе данных

Следующий алгоритм, который мы изучим, подобно алгоритму Саймона, также демонстрирует ускорение по отношению к тому, чего мы мо-

жем достичь с помощью классических вычислений. Однако в противоположность экспоненциальному ускорению решения задачи Саймона в этом случае ускорение только квадратично (квантовое время растет как квадратный корень классического времени). Несмотря на это, результат (открытый Л. Гровером) чрезвычайно интересен ввиду большой полезности этого алгоритма[1].

Рассматриваемая эвристически, проблема, к которой мы обратимся, выглядит так: мы столкнулись с очень большой неструктурированной базой данных, содержащей $N \gg 1$ отдельных объектов, а нам необходимо локализовать один конкретный объект, одним словом, найти иголку в стоге сена. С математической точки зрения база данных представлена таблицей или функцией $f(x)$ с $x \in \{0, 1, 2, \ldots N-1\}$. Мы уверены в том, что отдельная запись a появляется в таблице только один раз, то есть что $f(x) = a$ только при одном значении x. Проблема состоит в том, чтобы по данному a отыскать это значение x.

Если база данных подходящим образом *структурирована*, то поиск x прост. Возможно, кто-то был настолько любезен, что записал значения a в возрастающем порядке. Тогда мы можем найти x, просмотрев только $\log_2 N$ отдельных записей в таблице. Предположим, что $N = 2^n$ является степенью двойки. Мы сначала найдем $f(x)$ при $x = 2^{n-1} - 1$ и проверим, больше ли $f(x)$, чем a. Если да, то мы найдем следующее f при $x = 2^{n-2} - 1$ *и так далее*. С каждым взглядом на таблицу мы вдвое сокращаем количество кандидатов среди значений x, так что достаточно n взглядов, чтобы просмотреть все 2^n рассортированных записей. Вы можете использовать этот алгоритм, чтобы отыскать номер в телефонной книге Лос Анжелеса, поскольку в ней имена записаны в алфавитном порядке.

Но допустим, что вы знаете чей-то номер телефона, и вы хотите узнать его *имя*. Если у вас нет возможности заглянуть в обратный справочник, то процедура поиска будет утомительна. Ваши шансы таковы: вам придется проверить порядочное количество отдельных записей в телефонной книге, прежде чем вы наткнетесь на известный вам номер.

Фактически, если N номеров записаны в случайном порядке, то вам необходимо просмотреть $N/2$ номеров, прежде чем с вероятностью $P = 1/2$ найти его номер (и, следовательно, его имя). Обнаруженное Гровером состоит в том, что если вы имеете квантовую телефонную книгу, то, обратившись к ней примерно только $\sqrt{N}$ раз, вы можете с высокой вероятностью узнать интересующее вас имя.

Эта задача тоже может быть сформулирована как проблема оракула

[1] L. K. Grover, *Quantum Mechanics Helps in Searching for a Needle in a Haystack*, Phys. Rev. Lett., **79**, 325–328 (1997); quant-ph/9706033.

или «черного ящика». В этом случае оракулом является телефонная книга или справочная таблица. Мы можем ввести имя (значение x), а оракул – выдать нуль, если $f(x) \neq a$, или единицу, если $f(x) = a$. Наша задача – как можно быстрее найти значение x, при котором

$$f(x) = a. \tag{6.132}$$

Почему эта проблема важна? Возможно, вы никогда не пытались найти в телефонной книге имя, которое соответствует данному номеру, но если бы это не было так трудно, то вы, может быть, гораздо чаще пытались бы делать это. Более широко метод быстрого поиска в неструктурированной базе данных можно было бы привлечь к решению любой задачи из NP. Нашим оракулом может быть подпрограмма, которая опрашивает каждого потенциального «свидетеля» y, который потенциально мог бы подтвердить решение проблемы. Например, если мы сталкиваемся с графом и нам необходимо узнать, существует ли на нем гамильтонов обход, мы можем представить обход «оракулу», а он – быстро ответить, является этот обход гамильтоновым или нет. Если бы нам был известен быстрый способ спросить оракул обо всех возможных обходах, то мы были бы способны эффективно найти гамильтонов обход (если он существует).

6.4.1. Оракул

Итак, «оракулом» кратко называют подпрограмму, которая быстро вычисляет функцию, чтобы проверить предлагаемое решение проблемы принятия решения, однако продолжим рассматривать оракул абстрактно, как «черный ящик». Оракул «знает», что из 2^n возможных строк длины n одна («помеченная» строка или «решение» ω) особенная. Мы предлагаем оракулу вопрос x, а он сообщает нам или $x = \omega$, или нет. Другими словами, он сообщает значение функции .

$$\begin{aligned} f_\omega(x) &= 0, \qquad x \neq \omega, \\ f_\omega(x) &= 1, \qquad x = \omega. \end{aligned} \tag{6.133}$$

Даже более того, это *квантовый оракул*, следовательно, он может отвечать на вопросы, представляющие собой суперпозиции строк. Оракулом является квантовый черный ящик, выполняющий унитарное преобразование

$$\mathbf{U}_{f_\omega} : |x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f_\omega(x)\rangle, \tag{6.134}$$

где $|x\rangle$ – n-кубитовое состояние, а $|y\rangle$ – однокубитовое состояние.

Как мы видели раньше в других контекстах, состояние однокубитового регистра может быть выбрано равным $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, так что оракул действует как

$$\begin{aligned}\mathbf{U}_{f_\omega} \; : \; & |x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\ & \rightarrow (-1)^{f_\omega(x)}|x\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).\end{aligned} \tag{6.135}$$

Теперь мы можем игнорировать второй регистр и получить

$$\mathbf{U}_\omega \; : \; |x\rangle \rightarrow (-1)^{f_\omega(x)}|x\rangle \tag{6.136}$$

или

$$\mathbf{U}_\omega = \mathbf{1} - 2|\omega\rangle\langle\omega|. \tag{6.137}$$

Оракул обращает знак состояния $|\omega\rangle$, но на любое другое состояние, ортогональное $|\omega\rangle$, действует тривиально. Это преобразование имеет простую геометрическую интерпретацию. Действуя на любой вектор в 2^n-мерном гильбертовом пространстве, $\mathbf{U}_\omega$ *отражает* его в гиперплоскости, перпендикулярной $|\omega\rangle$ (он сохраняет компоненты в гиперплоскости и обращает компоненту вдоль $|\omega\rangle$).

Мы знаем, что оракул выполняет это отражение для некоторого частного состояния вычислительного базиса $|\omega\rangle$, но *a priori* нам ничего не известно относительно значения строки ω. Наша задача – обращаясь к оракулу минимальное количество раз, определить ω с максимальной вероятностью.

6.4.2. Итерация Гровера

В качестве первого шага приготовим состояние

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle. \tag{6.138}$$

Равновзвешенная суперпозиция всех состояний вычислительного базиса может быть легко получена применением преобразования Адамара к каждому кубиту начального состояния $|x = 0\rangle$. Хотя нам не известно значение ω, мы знаем, что $|\omega\rangle$ является состоянием из вычислительного базиса, так что независимо от значения ω

$$|\langle\omega|s\rangle| = \frac{1}{\sqrt{N}}. \tag{6.139}$$

Если бы мы измерили состояние $|s\rangle$, проецируя его на вычислительный базис, то мы «нашли» бы маркированное состояние $|\omega\rangle$, с вероятностью, равной всего лишь $1/N$. Однако, следуя алгоритму Гровера, мы можем многократно итерировать преобразование, повышая амплитуду вероятности неизвестного искомого состояния $|\omega\rangle$ и одновременно подавляя амплитуды всех ненужных состояний $|x \neq \omega\rangle$. Сконструируем эту итерацию Гровера, комбинируя выполняемое оракулом неизвестное отражение $\mathbf{U}_\omega$ с известным отражением, которое мы можем выполнить сами. Этим известным отражением является преобразование

$$\mathbf{U}_s = 2|s\rangle\langle s| - \mathbf{1}, \tag{6.140}$$

которое сохраняет $|s\rangle$, но обращает знак любого вектора, ортогонального $|s\rangle$. Геометрически, действуя на произвольный вектор, оно сохраняет его компоненту вдоль $|s\rangle$ и обращает знаки компонент в гиперплоскости, ортогональной $|s\rangle$.

Ниже мы вернемся к проблеме построения схемы, выполняющей $\mathbf{U}_s$; а пока лишь предположим, что можем эффективно выполнять $\mathbf{U}_s$.

Одна итерация Гровера представляет собой унитарное преобразование

$$\mathbf{R}_{\text{grov}} = \mathbf{U}_s\mathbf{U}_\omega, \tag{6.141}$$

в котором наше отражение следует за вопросом оракулу. Рассмотрим, как $\mathbf{R}_{\text{grov}}$ действует в плоскости, натянутой на векторы $|\omega\rangle$ и $|s\rangle$. Проще всего понять это действие, представив его геометрически. Вспомним, что

$$|\langle\omega|s\rangle| = \frac{1}{\sqrt{N}} = \sin\theta, \tag{6.142}$$

так что $|s\rangle$ лежит в плоскости, натянутой на ортогональные векторы $|\omega\rangle$ и $|\omega^\perp\rangle$, и наклонен к последнему из них под углом θ. В этой плоскости $\mathbf{U}_\omega$ отражает вектор относительно оси $|\omega^\perp\rangle$, а $\mathbf{U}_s$ – относительно оси $|s\rangle$. Совместно два этих отражения поворачивают вектор на угол 2θ:

$$U_{s0}U_\omega = 2\theta.$$

Тогда итерация Гровера является ничем иным, как поворотом на угол 2θ в плоскости, определяемой векторами $|s\rangle$ и $|\omega\rangle$.

6.4.3. Поиск одного из четырех

Предположим, например, что в базе данных $N = 4$ объекта, среди которых один маркированный. С помощью классических вопросов маркированный объект может быть найден с 1-го, 2-го, 3-го или 4-го раза; в среднем для достижения цели необходимо $2\frac{1}{2}$ вопроса, а в худшем случае –

четыре[1]. Но так как $\sin\theta = \frac{1}{\sqrt{N}} = \frac{1}{2}$, то $\theta = 30°$, $2\theta = 60°$ и, следовательно, после итерации Гровера $|s\rangle$ поворачивается в направлении, перпендикулярном $|\omega^\perp\rangle$, то есть вдоль оси $|\omega\rangle$. Теперь измерение, проецирующее на вычислительный базис, с *полной определенностью* дает результат $|\omega\rangle$. Достаточно всего одного квантового вопроса, чтобы найти маркированное состояние, заметное улучшение по сравнению с классическим случаем.

Иногда полезно альтернативное представление итерации Гровера как «инверсии относительно среднего». Если разложить состояние $|\psi\rangle$ в вычислительном базисе

$$|\psi\rangle = \sum_x a_x |x\rangle, \tag{6.143}$$

то его внутреннее произведение с $|s\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle$ можно представить в виде

$$\langle s|\psi\rangle = \frac{1}{\sqrt{N}} \sum_x a_x = \sqrt{N}\langle a\rangle, \tag{6.144}$$

где

$$\langle a\rangle = \frac{1}{N} \sum_x a_x \tag{6.145}$$

— средняя амплитуда. Тогда применение $\mathbf{U}_s = 2|s\rangle\langle s| - \mathbf{1}$ к $|\psi\rangle$ дает

$$\mathbf{U}_s|\psi\rangle = \sum_x (2\langle a\rangle - a_x)|x\rangle\,; \tag{6.146}$$

амплитуды преобразуются как

$$\mathbf{U}_s : \; a_x - \langle a\rangle \;\to\; \langle a\rangle - a_x, \tag{6.147}$$

то есть коэффициент перед $|x\rangle$ инвертируется относительно среднего значения амплитуды.

Возвращаясь к случаю $N = 4$, заметим, что в состоянии $|s\rangle$ каждая амплитуда равна $\frac{1}{2}$. Один вопрос оракулу обращает знак амплитуды маркированного состояния и, таким образом, сокращает среднюю амплитуду

[1] Конечно, если мы знаем, что один маркированный объект здесь обязательно присутствует, то четвертый вопрос на самом деле является излишним, так что можно быть точнее и говорить, что необходимо самое большее три вопроса, а в среднем — $2\frac{1}{4}$.

до $\frac{1}{4}$. Тогда инверсия относительно среднего значения переводит амплитуды всех немаркированных состояний от $\frac{1}{2}$ в нуль и увеличивает амплитуду маркированного состояния от $-\frac{1}{2}$ до $+1$. Итак, мы воспроизвели наш вывод о том, что достаточно одного вопроса, чтобы с полной определенностью найти маркированное состояние.

Также легко понять, что одного вопроса достаточно для того, чтобы найти маркированное состояние, если в базе данных имеется N записей и ровно $\frac{1}{4}$ из них маркирована. Тогда, как и выше, один вопрос сокращает среднюю амплитуду от $\frac{1}{\sqrt{N}}$ до $\frac{1}{2\sqrt{N}}$, а *инверсия относительно среднего* сокращает амплитуды немаркированных состояний до нуля.

(Сравнивая количество квантовых и классических вопросов, с которыми нужно обратиться к оракулу, возможно, не совсем справедливо говорить, что в квантовом случае необходим только один вопрос. Если оракул выполняет программу, которая вычисляет функцию, то в процессе вычисления некоторое вспомогательное пространство будет заполнено мусором. Нам будет необходимо удалить мусор, пройдя вычисление в обратном направлении для того, чтобы сохранить квантовую когерентность. Если классическое вычисление необратимо, то нет необходимости возвращать оракул в исходное состояние. В этом смысле, на языке теории сложности, один вопрос квантовому оракулу может быть примерно эквивалентным двум вопросам классическому оракулу.)

6.4.4. Поиск одного из N

Вернемся теперь к случаю, в котором база данных содержит N объектов, среди которых ровно один маркирован. Каждая итерация Гровера поворачивает квантовое состояние в плоскости, определяемой векторами $|s\rangle$ и $|\omega\rangle$; после T итераций состояние оказывается наклоненным к оси $|\omega^{\perp}\rangle$ под углом $\theta + 2T\theta$. Чтобы оптимизировать вероятность обнаружения маркированного состояния при выполнении заключительного измерения, итерировать следует до угла, близкого к 90°, или

$$(2T+1)\theta \simeq \frac{\pi}{2} \;\Rightarrow\; 2T+1 \simeq \frac{\pi}{2\theta}; \tag{6.148}$$

вспомним, что $\sin\theta = \frac{1}{\sqrt{N}}$, или, при больших N,

$$\theta \simeq \frac{1}{\sqrt{N}}. \tag{6.149}$$

Если выбрать

$$T = \frac{\pi}{4}\sqrt{N}[1 + O(N^{-1/2})], \tag{6.150}$$

то вероятность получения $|\omega\rangle$ в качестве результата измерения будет равна

$$\text{Prob}(\omega) = \sin^2\left((2T + 1)\theta\right) = 1 - O\left(\frac{1}{N}\right). \tag{6.151}$$

Таким образом, необходимо лишь около $\frac{\pi}{4}\sqrt{N}$ вопросов, чтобы с высокой вероятностью определить ω, квадратичное ускорение по сравнению с классическим результатом.

6.4.5. Множество решений

Если существует $r > 1$ маркированных состояний и r известно, то количество итераций можно модифицировать так, чтобы вероятность отыскания одного из них оставалась очень близкой к единице. Анализ такой же, как и выше, за исключением того, что теперь оракул индуцирует отражение в гиперплоскости, ортогональной вектору

$$|\tilde{\omega}\rangle = \frac{1}{\sqrt{r}}\sum_{i=1}^{r}|\omega_i\rangle \tag{6.152}$$

— равновзвешенной суперпозиции маркированных состояний вычислительного базиса $|\omega_i\rangle$. Теперь

$$\langle s|\tilde{\omega}\rangle = \sqrt{\frac{r}{N}} \equiv \sin\theta, \tag{6.153}$$

а итерация Гровера поворачивает вектор на угол 2θ в плоскости, натянутой на векторы $|s\rangle$ и $|\tilde{\omega}\rangle$; мы снова приходим к выводу, что после количества итераций

$$T = \frac{\pi}{4\theta} = \frac{\pi}{4}\sqrt{\frac{N}{r}} \tag{6.154}$$

состояние близко к $|\tilde{\omega}\rangle$. Тогда если мы выполним измерение, проецируя на вычислительный базис, то с вероятностью, близкой к единице, найдем одно из маркированных (равновероятных) состояний. (С ростом количества решений время, необходимое для отыскания одного из них, падает как $r^{-1/2}$, в противоположность к r^{-1} в классическом случае.)

Обратим внимание на то, что если продолжить выполнение итераций Гровера, то вектор продолжит поворачиваться, и, таким образом, вероятность отыскания маркированного состояния (в результате заключительного измерения) начнет падать. Алгоритм Гровера подобен вышечке суфле – стоит передержать его в духовке, как оно начнет опадать. Следовательно, если нам ничего неизвестно о количестве маркированных состояний, то поиск одного из них может оказаться безуспешным. Например, $T \sim \frac{\pi}{4}\sqrt{N}$ итераций оптимально при $r = 1$, но при $r = 4$ вероятность отыскания маркированного состояния после этого количества итераций довольно близка к нулю.

Но даже если r *a priori* неизвестно, мы все же можем найти решение с квадратичным, по сравнению с классическими алгоритмами (при $r \ll N$), ускорением. Например, мы можем выбрать количество итераций случайным в интервале от нуля до $\frac{\pi}{4}\sqrt{N}$. Тогда для каждого r, ожидаемая вероятность отыскания маркированного состояния близка к $\frac{1}{2}$. Следовательно, маловероятно, что нам не удастся найти маркированное состояние после нескольких повторений. А при каждом измерении мы можем предлагать оракулу найденное нами состояние в качестве классического вопроса, чтобы получить подтверждение того, является ли оно действительно маркированным.

В частности, если решение не было найдено после нескольких попыток, то вполне возможно, что оно не существует. Таким образом, с высокой вероятностью можно дать корректный ответ ДА/НЕТ на вопрос «Есть ли здесь маркированные состояния?». Следовательно, мы можем принять алгоритм Гровера, в котором оракул проверяет предложенное решение, чтобы решить любую NP-проблему с квадратичным ускорением по сравнению с классическим методом исчерпывающего поиска.

6.4.6. Осуществление отражения

Чтобы выполить итерацию Гровера, необходимо (кроме вопроса оракулу) унитарное преобразование

$$\mathbf{U}_s = 2|s\rangle\langle s| - \mathbf{1}, \tag{6.155}$$

которое отражает вектор относительно оси, определяемой вектором $|s\rangle$. Как эффективно построить это преобразование из квантовых вентилей? Так как $|s\rangle = \mathbf{H}^{(n)}|0\rangle$, где $\mathbf{H}^{(n)}$ – побитовое преобразование Адамара, то можно записать

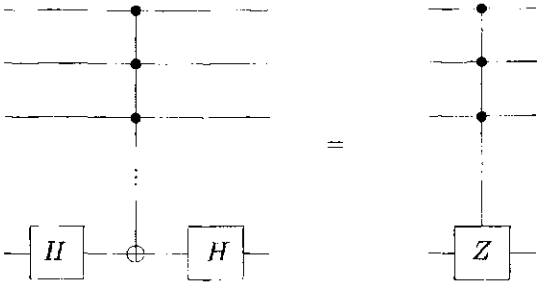
$$\mathbf{U}_s = \mathbf{H}^{(n)}\big(2|0\rangle\langle 0| - \mathbf{1}\big)\mathbf{H}^{(n)}, \tag{6.156}$$

то есть для этого достаточно построить отражение относительно оси $|0\rangle$. Мы легко можем построить это отражение из n-битового вентиля Тоффоли $\theta^{(n)}$.

Вспомним, что

$$\mathbf{H}\sigma_x\mathbf{H} = \sigma_z\,; \tag{6.157}$$

инвертирование бита с адамаровским поворотом базиса эквивалентно обращению относительной фазы векторов $|0\rangle$ и $|1\rangle$. Следовательно:



после сопряжения последнего бита преобразованием $\mathbf{H}$ вентиль $\theta^{(n)}$ становится $(n-1)$-кратно контролируемым σ_z, который обращает фазу вектора $|11\ldots\rangle|1\rangle$ и действует тривиально на все другие состояния вычислительного базиса. Сопрягая с помощью $\mathrm{NOT}^{(n)}$, мы получаем $\mathbf{U}_s$ с точностью до несущественного общего знака минус.

В упражнении вы покажете, что n-битовый вентиль Тоффоли $\theta^{(n)}$ можно построить из $(2n-5)$-ти трехбитовых вентилей Тоффоли $\theta^{(3)}$ (если доступно достаточное вспомогательное пространство). Следовательно, образующая $\mathbf{U}_s$ схема имеет *линейный* по $n = \log N$ размер. Гроверовский поиск в базе данных (при условии, что оракул мгновенно отвечает на вопрос) требует времени порядка $\sqrt{N}\log N$. Если мы рассматриваем оракул как подпрограмму, которая вычисляет функцию за полилогарифмическое время, тогда поиск требует времени порядка $\sqrt{N}\mathrm{poly}\big(\log N\big)$.

6.5. Оптимальность алгоритма Гровера

Гроверовское квадратичное квантовое ускорение поиска в базе данных уже интересно и потенциально важно, но конечно же, действуя более искусно, мы можем добиться лучшего результата, не так ли? Нет, оказывается не можем. Алгоритм Гровера обеспечивает максимально быстрый квантовый поиск в неструктурированной базе данных, если «время» измеряется в соответствии с количеством задаваемых оракулу вопросов.

Рассмотрим случай одного маркированного состояния $|\omega\rangle$, пусть $\mathbf{U}(\omega, T)$ обозначает квантовую схему, T раз обращающуюся к оракулу. Мы не накладываем на эту схему *никаких* ограничений, за исключением количества задаваемых ей вопросов; в частности, мы не ограничиваем количество квантовых вентилей. Эта схема применяется к начальному состоянию $|\psi(0)\rangle$, производя конечное состояние

$$|\psi_\omega(T)\rangle = \mathbf{U}(\omega, T)|\psi(0)\rangle. \tag{6.158}$$

Теперь мы должны выполнить измерение, предназначенное выделить $|\omega\rangle$ среди N его возможных значений. Для того чтобы мы были в состоянии идеально различать возможные состояния $|\psi_\omega(t)\rangle$, они должны быть взаимно ортогональными, а чтобы их можно было корректно различать с высокой вероятностью, они должны быть почти ортогональны.

Если состояния $\{|\psi_\omega\rangle\}$ образуют ортонормированный базис, то для любого фиксированного нормированного вектора $|\varphi\rangle$

$$\sum_{\omega=0}^{N-1} \| |\psi_\omega\rangle - |\varphi\rangle \|^2 \geqslant 2N - 2\sqrt{N}. \tag{6.159}$$

[Сумма минимизируется, если $|\varphi\rangle$ является равновзвешенной суперпозицией всех элементов базиса $|\varphi\rangle = \frac{1}{\sqrt{N}} \sum_\omega |\psi_\omega\rangle$, как вы можете показать, привлекая метод неопределенных множителей Лагранжа нахождения условных экстремумов.] Наша стратегия состоит в подходящем выборе состояния $|\varphi\rangle$, позволяющем с помощью неравенства (6.159) что-нибудь узнать о количестве обращений к оракулу T.

Наша схема с T вопросами образует унитарное преобразование

$$\mathbf{U}(\omega, T) = \mathbf{U}_\omega \mathbf{U}_T \mathbf{U}_\omega \mathbf{U}_{T-1} \ldots \mathbf{U}_\omega \mathbf{U}_1, \tag{6.160}$$

где $\mathbf{U}_\omega$ — преобразование оракула, а $\mathbf{U}_t$ — произвольные преобразования не-оракула. Выберем в качестве $|\varphi(T)\rangle$ результат применения к состоянию $|\psi(0)\rangle$ преобразования $\mathbf{U}(\omega, T)$, в котором каждое $\mathbf{U}_\omega$ заменено на $\mathbf{1}$; то есть результат применения той же схемы, но со всеми вопросами, задаваемыми «пустому оракулу». Следовательно,

$$|\varphi(T)\rangle = \mathbf{U}_T \mathbf{U}_{T-1} \ldots \mathbf{U}_2 \mathbf{U}_1 |\psi(0)\rangle, \tag{6.161}$$

в то время как

$$|\psi_\omega(T)\rangle = \mathbf{U}_\omega \mathbf{U}_T \mathbf{U}_\omega \mathbf{U}_{T-1} \ldots \mathbf{U}_\omega \mathbf{U}_1 |\psi(0)\rangle. \tag{6.162}$$

Чтобы сравнить $|\varphi(T)\rangle$ с $|\psi_\omega(T)\rangle$, воспользуемся нашим предыдущим анализом влияния ошибок на точность схемы, рассматривая ω-оракул как ошибочное применение пустого оракула. Норма вектора ошибки после t-го шага [ср. уравнение (6.63)] равна

$$\||E(\omega,t)\rangle\| = \|(\mathbf{U}_\omega - \mathbf{1})|\varphi(t)\rangle\| = 2\|\langle\omega|\varphi(t)\rangle\|, \tag{6.163}$$

поскольку $\mathbf{U}_\omega = \mathbf{1} - 2|\omega\rangle\langle\omega|$. После T вопросов мы имеем [ср. уравнение (6.66)]

$$\||\psi_\omega(T)\rangle - |\varphi(T)\rangle\| \leqslant 2\sum_{t=1}^{T}|\langle\omega|\varphi(t)\rangle|. \tag{6.164}$$

Из тождества

$$\left(\sum_{t=1}^{T} c_t\right)^2 + \frac{1}{2}\sum_{s,t=1}^{T}(c_s - c_t)^2 = \sum_{s,t=1}^{T}\left(c_t c_s + \frac{1}{2}c_s^2 - c_t c_s + \frac{1}{2}c_t^2\right) = T\sum_{t=1}^{T} c_t^2 \tag{6.165}$$

следует неравенство

$$\left(\sum_{t=1}^{T} c_t\right)^2 \leqslant T\sum_{t=1}^{T} c_t^2, \tag{6.166}$$

применение которого к (6.164) дает

$$\||\psi_\omega(T)\rangle - |\varphi(T)\rangle\|^2 \leqslant 4T\left(\sum_{t=1}^{T}|\langle\omega|\varphi(t)\rangle|^2\right). \tag{6.167}$$

Суммируя по ω, мы находим

$$\sum_\omega \||\psi_\omega(T)\rangle - |\varphi(T)\rangle\|^2 \leqslant 4T\sum_{t=1}^{T}\langle\varphi(t)|\varphi(t)\rangle = 4T^2. \tag{6.168}$$

Привлекая неравенство (6.159), мы приходим к выводу, что

$$4T^2 \geqslant 2N - 2\sqrt{N}, \tag{6.169}$$

если состояния $|\psi_\omega(T)\rangle$ взаимно ортогональны. Следовательно, мы показали, что любой квантовый алгоритм, способный различить все возможные значения маркированного состояния, должен обратиться к оракулу T раз, где

$$T \geqslant \sqrt{\frac{N}{2}} \tag{6.170}$$

(без учета малых при $N \to \infty$ поправок). Алгоритм Гровера находит ω с помощью $\frac{\pi}{4}\sqrt{N}$ вопросов, что превышает эту границу всего примерно на 11%. В действительности можно усовершенствовать доказательство, чтобы улучшить границу до $\frac{\pi}{4}\sqrt{N}(1-\varepsilon)$, что асимптотически насыщается алгоритмом Гровера[1]. Более того, можно показать, что схема Гровера достигает оптимальной вероятности успеха с помощью $T \leqslant \frac{\pi}{4}\sqrt{N}$ вопросов.

Испытываешь приступ разочарования (и одновременно волну восхищения перед Гровером) от осознания того, что алгоритм поиска в базе данных не может быть улучшен. Какое отношение это имеет к квантовой сложности?

Для многих проблем оптимизации в NP-классе не известно лучшего метода, чем исчерпывающий поиск всех возможных решений. Используя *квантовый параллелизм*, можно достичь квадратичного ускорения исчерпывающего поиска. Теперь мы знаем, что квадратичное ускорение является наилучшим, если мы полагаемся на силу явного квантовового параллелизма и не разрабатываем наш квантовый алгоритм, используя специфическую структуру решаемой задачи. Тем не менее при достаточной изобретательности можно добиться и лучшего результата.

Оптимальность алгоритма Гровера может быть истолкована как свидетельство того, что $BQP \not\supseteq NP$. *По крайней мере, если окажется*, что $NP \subseteq BQP$, а $P \neq NP$, то тогда NP-проблемы должна объединять глубокая внутренняя структура (свидетельств которой в настоящее время нет), хорошо подходящая к возможностям квантовых схем.

Даже квадратичное ускорение может оказаться полезным для различных NP-полных проблем оптимизации. Однако квадратичное ускорение, в отличие от экспоненциального, реально не перемещает границу между разрешимостью и сложностью. В один прекрасный день квантовые компьютеры смогут превзойти классические компьютеры в выполнении исчер-

[1] C. Zalka, *Grover's Quantum Searching Algorithm is Optimal*, Phys. Rev., **A60**, 2746–2751 (1999); quant-ph/9711070. [Впервые оптимальность алгоритма Гровера была доказана в работе: C. H. Bennet, E. Bernstein, G. Brassard, U. Vazirani, *Strengths and Weaknesses of Quantum Computing*, SIAM J. Comput., **26(5)**, 1510–1523 (1997); quant-phys/9701001 – *Прим. ред.*

нывающего поиска, но только в том случае, если часы квантовых приборов не будут слишком сильно отставать от их классических прототипов.

6.6. Обобщенный поиск и структурированный поиск

В итерации Гровера мы выполняем преобразование $\mathbf{U}_s = 2|s\rangle\langle s| - \mathbf{1}$, отражение относительно оси, определяемой вектором $|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle$. Почему именно относительно нее? Преимущество состояния $|s\rangle$ состоит в том, что оно имеет одинаковые перекрытия с каждым состоянием вычислительного базиса. Таким образом, перекрытие любого маркированного состояния $|\omega\rangle$ с $|s\rangle$ гарантированно равно $|\langle\omega|s\rangle| = 1/\sqrt{N}$. Следовательно, если нам известно количество маркированных состояний, то мы можем определить, сколько потребуется итераций, чтобы с высокой вероятностью отыскать одно из них – количество необходимых итераций не зависит от того, какое состояние маркировано.

Но, конечно, мы могли бы выбрать отражение относительно другой оси. Если мы можем построить унитарное преобразование $\mathbf{U}$ (с разумной эффективностью), тогда мы можем образовать

$$\mathbf{U}(2|0\rangle\langle 0| - \mathbf{1})\mathbf{U}^\dagger = 2\mathbf{U}|0\rangle\langle 0|\mathbf{U}^\dagger - \mathbf{1} \tag{6.171}$$

преобразование, отражающее относительно оси $\mathbf{U}|0\rangle$.

Предположим, что

$$|\langle\omega|\mathbf{U}|0\rangle| = \sin\theta, \tag{6.172}$$

где $|\omega\rangle$ – маркированное состояние. Тогда, если мы заменим в итерации Гровера $\mathbf{U}_s$ на отражение (6.171), то одна итерация будет выполнять поворот на угол 2θ в плоскости, определяемой векторами $|\omega\rangle$ и $\mathbf{U}|0\rangle$ (в соответствии с теми же аргументами, что мы использовали для $\mathbf{U}_s$). Таким образом, после T итераций с $(2T+1)\theta \simeq \pi/2$ измерение в вычислительном базисе с высокой вероятностью найдет $|\omega\rangle$. Следовательно, если мы заменим в квантовой схеме Гровера $\mathbf{H}^{(n)}$ на $\mathbf{U}$, мы по-прежнему сможем выполнять поиск в базе данных до тех пор, пока $\mathbf{U}|0\rangle$ остается не ортогональным маркированному состоянию[1]. Но если мы не имеем никакой *априорной* информации о том, какое состояние маркировано, то $\mathbf{H}^{(n)}$ является наилучшим выбором не только потому, что $|s\rangle$ имеет известное перекрытие

[1] L. K. Grover, *Quantum Computers Can Search Rapidly By Using Almost Any Transformation*, Phys. Rev. Lett., **80**, 4329–4332 (1998); quant-ph/9712011.

с каждым маркированным состоянием, но также и потому, что оно имеет максимальное *среднее* перекрытие со всеми возможными маркированными состояниями.

Но иногда, выполняя поиск в базе данных, мы *имеем* некоторую информацию о том куда следует заглянуть, а в этом случае может оказаться полезной описанная выше стратегия обобщенного поиска[1].

В качестве примера проблемы с некоторой вспомогательной структурой предположим, что $f(x, y)$ — функция с однобитовым значением, зависящая от двух n-битовых строк x и y, и нам нужно найти единственное решение $f(x, y) = 1$. С помощью алгоритма Гровера мы можем искать среди N^2 возможных значений ($N = 2^n$) пар (x, y) и с высокой вероятностью найти решение (x_0, y_0) после $\frac{\pi}{4}N$ итераций, квадратичное ускорение по сравнению с классическим поиском.

Но предположим далее, что $g(x)$ — функция только от x, и известно, что $g(x) = 1$ при ровно M значениях x, где $1 \ll M \ll N$. Более того, известно, что $g(x_0) = 1$. Следовательно, мы можем воспользоваться функцией g, чтобы помочь в поиске решения (x_0, y_0).

Теперь для консультаций у нас есть два оракула, один выдает значение $f(x, y)$, а другой значение $g(x)$. Наша задача — найти (x_0, y_0), задав минимальное количество вопросов.

С классической точки зрения нам необходимо около NM вопросов для того, чтобы с разумной вероятностью найти решение. Сначала мы вычисляем $g(x)$ для каждого x; затем мы ограничиваем наш поиск решения $f(x, y) = 1$ только такими M значениями x, при которых $g(x) = 1$. Естественно поинтересоваться, существует ли способ выполнить квантовый поиск за время порядка квадратного корня от классического времени. Исчерпывающий поиск, который обращается только к f-оракулу, требует времени $N \gg \sqrt{NM}$ и, следовательно, не решает проблемы. Нам необходимо пересмотреть наш метод квантового поиска, чтобы воспользоваться преимуществом структуры, предоставляемой функцией g.

Лучший метод — сначала применить алгоритм Гровера к $g(x)$. Примерно за $\frac{\pi}{4}\sqrt{\frac{N}{M}}$ итераций мы приготовим состояние, близкое к равновзвешенной суперпозиции из M решений $g(x) = 1$. В частности, состояние $|x_0\rangle$ возникает с амплитудой $\frac{1}{\sqrt{M}}$. Затем мы применяем алгоритм

[1] E. Farhi and S. Gutmann, *Quantum-Mechanical Square Root Speedup in a Structured Search Problem*, quant-ph/9711035; L. K. Grover, *Quantum Search On Structured Problems*, quant-ph/9802035.

Гровера к $f(x,y)$ при фиксированном x. Приблизительно после $\frac{\pi}{4}\sqrt{N}$ итераций состояние $|x_0\rangle|s\rangle$ эволюционирует достаточно близко к состоянию $|x_0\rangle|y_0\rangle$. Следовательно, $|x_0, y_0\rangle$ появляется с амплитудой $\frac{1}{\sqrt{M}}$.

Образованное из $\frac{\pi}{4}\sqrt{N}$ вопросов унитарное преобразование, которое мы до сих пор строили, может рассматриваться как преобразование $\mathbf{U}$, определяющее обобщенный поиск. Более того, нам известно, что

$$\langle x_0, y_0|\mathbf{U}|0,0\rangle \simeq \frac{1}{\sqrt{M}}. \tag{6.173}$$

Следовательно, если мы итерируем обобщенный поиск примерно $\frac{\pi}{4}\sqrt{M}$ раз, то приготовим состояние, достаточно близкое к $|x_0, y_0\rangle$. В совокупности примерно после

$$\left(\frac{\pi}{4}\right)^2 \sqrt{NM} \tag{6.174}$$

вопросов мы можем с высокой вероятностью найти решение. Это действительно квадратичное ускорение по сравнению с классическим поиском.

6.7. Некоторые задачи не допускают ускорения

Пример структурированного квантового поиска иллюстрирует, что квадратичные квантовые ускорения по сравнению с классическими алгоритмами могут быть достигнуты для различных проблем, а не только для исчерпывающего поиска в неструктурированной базе данных. Можно даже надеяться, что квантовый параллелизм позволяет существенно ускорить любой классический алгоритм. Сейчас эта надежда будет разбита — для многих задач квантовое ускорение невозможно.

Продолжим рассматривать задачи с квантовым черным ящиком, оракулом, который вычисляет функцию f, отображающую n битов в один. Но мы немного модифицируем наши обозначения. Функция f может быть представлена строкой из $N = 2^n$ битов:

$$X = X_{N-1}X_{N-2}\ldots X_1X_0, \tag{6.175}$$

где X_i обозначает $f(i)$. Наша задача — вычислить некоторую зависящую от X функцию с однобитовым значением, то есть ответить на ДА/НЕТ-вопрос о свойствах оракула. То, что сейчас будет показано, означает, что некоторые функции от X не могут быть вычислены с низкой вероятностью

ошибки, используя квантовый алгоритм, за исключением алгоритма, обращающегося к оракулу столько раз (или почти столько же раз), сколько требуется классическому алгоритму[1].

Главная идея состоит в том, что булева функция от переменных X_i может быть представлена полиномом от X_i. Более того, распределение вероятностей для квантового измерения может быть выражено через полином от X_i, где степень полинома равна $2T$, если измерение следует после T вопросов оракулу. Проблема в том, может ли полином степени $2T$ обеспечить разумную аппроксимацию интересующей нас булевой функции.

Действие оракула может быть представлено как

$$\mathbf{U}_O\colon |i, y; z\rangle \to |i, y \oplus X_i; z\rangle, \tag{6.176}$$

где i принимает значения из $\{0, 1, \ldots, N-1\}$, $y \in \{0, 1\}$, а z обозначает состояние вспомогательного кубита, не изменяемого оракулом. Следовательно, в каждом 2×2-блоке, натянутом на $|i, 0; z\rangle$ и $|i, 1; z\rangle$, $\mathbf{U}_O$ представляет собой 2×2-матрицу

$$\begin{pmatrix} 1 - X_i & X_i \\ X_i & 1 - X_i \end{pmatrix}. \tag{6.177}$$

Квантовые вентили, в отличие от вопросов к оракулу, не зависят от X. Следовательно, после того как схема из T вопросов подействует на любое начальное состояние, результирующее состояние $|\psi\rangle$ будет иметь амплитуды, которые (по крайней мере) являются полиномами степени T от X. Если мы выполним ПОЗМ на $|\psi\rangle$, то связанная с положительным оператором $\mathbf{F}$ вероятность результата $\langle\psi|\mathbf{F}|\psi\rangle$ может быть выражена через полином от X степени, не меньшей чем $2T$.

Любая булева функция от X_i может быть выражена (единственным образом) через полином степени $\leqslant N$ по X_i. Например, рассмотрим функцию OR от N переменных X_i; это

$$\mathrm{OR}(X) = 1 - (1 - X_0)(1 - X_1)\cdots(1 - X_{N-1}), \tag{6.178}$$

полином степени N.

Допустим, мы хотим применить нашу квантовую схему для того, чтобы *с полной определенностью* вычислить функцию OR. Тогда мы должны

[1] E. Farhi, *et al*, *A Limit on the Speed of Quantum Computation in Determining Parity*, Phys. Rev. Lett., **81**, 5442–5444 (1998); quant-ph/9802045; R. Beals, *et al*, *Quantum Lower Bounds by Polynomials*. In *Proceedings of the 39th Annual Symposium on Foundamentals of Computer Science* (FOCS'98), 352-361, IEEE, Los Alamos, California, November, 1998; quant-ph/9802049.

быть в состоянии выполнить измерение с двумя результатами 0 и 1, где

$$\begin{aligned} \text{Prob}(0) &= 1 - \text{OR}(X), \\ \text{Prob}(1) &= \text{OR}(X). \end{aligned} \tag{6.179}$$

Но эти выражения являются полиномами степени N, которые могут быть вычислены только после как минимум T-кратного обращения схемы к оракулу, где

$$T \geqslant \frac{N}{2}. \tag{6.180}$$

Мы приходим к выводу, что не существует квантовой схемы, которая менее чем за $N/2$ обращений к оракулу может точно вычислить OR. Фактически для этой функции (или любой функции, принимающей значение 0 только для одного из N ее возможных аргументов) существует более сильное заключение (см. упражнение): требуется $T \geqslant N$, чтобы *с полной определенностью* вычислить OR.

С другой стороны, вычисляя функцию OR (отвечая на ДА/НЕТ-вопрос «Имеется ли маркированное состояние?»), именно алгоритм Гровера может достичь количества вопросов порядка $\sqrt{N}$. Таким образом, вывод о том, что для *детерминированного* вычисления OR необходимо N вопросов, хотя и корректен, но несколько обманчив. Мы можем вычислить OR *вероятностным образом* с помощью гораздо меньшего количества вопросов. По-видимому, алгоритм Гровера может построить полином от X, который, имея степень только $O(\sqrt{N})$, тем не менее обеспечивает весьма адекватную аппроксимацию полинома N-ой степени $\text{OR}(X)$.

Однако OR, принимающая значение 1 для каждого значения X, кроме $X = \vec{0}$, является очень простой булевой функцией. Нам следует рассмотреть другие функции, которые могут предложить квантовому компьютеру более серьезные проблемы.

Первое, что приходит в голову — это функция $\text{PARITY}(X)$, принимающая значение 0, если строка X содержит четное количество единиц, и — значение 1, если строка X содержит нечетное количество единиц. Очевидно, чтобы определить четность, классический алгоритм должен обратиться к оракулу N раз. Насколько лучше это можно сделать, предлагая квантовые вопросы? Фактически мы вообще не можем добиться лучшего результата — необходимо по крайней мере $N/2$ квантовых вопросов, чтобы с вероятностью успеха, большей $\frac{1}{2} + \delta$, найти правильное значение $\text{PARITY}(X)$.

При обсуждении PARITY удобно использовать новые переменные

$$\tilde{X}_i = 1 - 2X_i, \tag{6.181}$$

принимающие значения ± 1, так что

$$\mathrm{PARITY}(\tilde{X}) = \prod_{i=0}^{N-1} \tilde{X}_i \tag{6.182}$$

также принимает значения ± 1. Теперь, после того как выполнена квантовая схема со всеми T вопросами оракулу, мы должны выполнить ПОЗМ с двумя возможными исходами $\mathbf{F}_{\mathrm{even}}$ и $\mathbf{F}_{\mathrm{odd}}$; результатом будет наша оценка $\mathrm{PARITY}(\tilde{X})$. Как мы уже отмечали, вероятность получения результата (к примеру) «even» («четный») может быть выражена через полином $P_{\mathrm{even}}^{(2T)}(\tilde{X})$ степени (самое большее) $2T$ по $\tilde{X}_i$:

$$\langle \mathbf{F}_{\mathrm{even}} \rangle = P_{\mathrm{even}}^{(2T)}(\tilde{X}). \tag{6.183}$$

Как часто наша догадка будет верна? Рассмотрим сумму

$$\sum_{\{\tilde{X}\}} P_{\mathrm{even}}^{(2T)}(\tilde{X}) \cdot \mathrm{PARITY}(\tilde{X}) = \sum_{\{\tilde{X}\}} P_{\mathrm{even}}^{(2T)}(\tilde{X}) \prod_{i=0}^{N-1} \tilde{X}_i. \tag{6.184}$$

Поскольку каждое слагаемое полинома $P_{\mathrm{even}}^{(2T)}(\tilde{X})$ содержит самое большее $2T$ из переменных $\tilde{X}_i$, то можно воспользоваться тождеством

$$\sum_{X_i \in \{0,1\}} \tilde{X}_i = 0, \tag{6.185}$$

чтобы понять, что сумма в (6.184) должна обращаться в нуль при $N > 2T$. Таким образом,

$$\sum_{\mathrm{par}(\tilde{X})=1} P_{\mathrm{even}}^{(2T)}(\tilde{X}) = \sum_{\mathrm{par}(\tilde{X})=-1} P_{\mathrm{even}}^{(2T)}(\tilde{X}); \tag{6.186}$$

то есть при $T < N/2$ мы с одинаковой вероятностью можем угадать «even», как в случае, когда на самом деле $\mathrm{PARITY}(\tilde{X})$ является нечетной, так и в том случае, когда она действительно четная (в среднем). Наш квантовый алгоритм ничего не может сообщить о значении $\mathrm{PARITY}(\tilde{X})$; то есть в среднем по возможным (*a priori* равновероятным) значениям X_i мы с равной вероятностью будем как правы, так и неправы.

Мы можем также показать, демонстрируя явный алгоритм (см. упражнение), что для определения PARITY (или вероятностного, или детерминированного) *достаточно* $N/2$ вопросов (при условии, что N четное). В этом смысле мы достигаем двукратного ускорения по сравнению с классическими вопросами. Но это лучшее из того, чего мы можем добиться.

6.8. Поиск в распределенной базе данных

Поучительно взглянуть на квантовый алгоритм поиска в базе данных с новой точки зрения. Представим, что двум участникам, Алисе и Бобу, необходимо договориться о встрече в удобный для обоих день. У Алисы есть календарь, в который внесено $N = 2^n$ дней, каждый из которых отмечен нулем, если в этот день она занята, или единицей, если она свободна. Аналогичный календарь имеется у Боба. Их задача найти день, в который они оба будут свободны.

Алиса и Боб имеют квантовые компьютеры, но они находятся очень далеко друг от друга. (Алиса на Земле, а Боб путешествует по туманности Андромеды.) Следовательно, для них слишком дорого связываться друг с другом. Им нужно срочно определить дату, но они должны экономить на объеме посылаемой туда и обратно информации.

Даже если существует день, когда они оба свободны, может оказаться, что найти его нелегко. Если Алиса и Боб связываются, посылая туда и обратно классические биты, тогда в худшем случае им будет необходимо обменяться порядка $N = 2^n$ записями календаря, чтобы иметь разумный шанс успешно договориться о встрече. Мы спросим: может быть, вместо этого им лучше обмениваться кубитами?[1] (От Земли до Андромеды тщательно спроектирован и построен квантовый информационный хайвей, так что посылать кубиты вместо битов не намного дороже.)

Для знакомого с основами теории квантовой информации этот вопрос выглядит странным. Теорема Холево раз и навсегда сказала, что один кубит может переносить не более одного бита классической информации. Хотя, немного подумав, мы увидим, что теорема Холево фактически не решает проблему. Хотя она ограничивает взаимную информацию приготовления состояния и результата измерения, она не гарантирует (по крайней мере не прямо), что Алисе и Бобу необходимо обменяться таким же количеством

[1] В ранней версии этих лекций я предлагал другой сценарий, в котором Алиса и Боб имели почти идентичные таблицы, но с одной несовпадающей записью; их задачей было найти положение несовпадающего бита. Однако этот пример был неудачен, поскольку задача могла быть решена с помощью всего лишь $\log N$ битов классической связи. (Я благодарен Ричарду Кливу, указавшему на эту ошибку.) Мы хотели, чтобы Алиса узнала адрес (двоичную строку длиной n) одной записи, которой ее таблица отличается от таблицы Боба. Для этого Боб вычисляет четность $N/2$ записей своей таблицы с меткой, принимающей значение 0 в ее самом младшем значащем бите, и посылает Алисе только бит четности. Алиса сравнивает четность тех же записей ее таблицы и находит один бит (самый младший значащий бит) адреса несовпадающей записи. Затем они повторяют то же самое для каждого из оставшихся $n - 1$ битов до тех пор, пока Алиса не узнает полный адрес «ошибки». Всего послано только n битов (и все от Боба к Алисе).

кубитов, что и битов, чтобы сравнить их календари. Тем не менее это приятный сюрприз — узнать, что Алиса и Боб могут выполнить задание, обменявшись $O(\sqrt{N}\log N)$ кубитами по сравнению с $O(N)$ классическими битами[1].

Чтобы добиться этого, Алиса и Боб должны действовать сообща, осуществляя распределенную версию поиска в базе данных. Алиса имеет доступ к оракулу (ее календарь), вычисляющему функцию $f_A(x)$, а Боб имеет оракул (его календарь), вычисляющий $f_B(x)$. Вместе они могут предложить оракулу

$$f_{AB}(x) = f_A(x) \wedge f_B(x). \tag{6.187}$$

Один из них может осуществить отражение $\mathbf{U}_s$, так что они могут выполнить полную итерацию Гровера и произвести исчерпывающий поиск подходящего дня x такого, что $f_{AB}(x) = 1$ (когда Алиса и Боб оба свободны). Если взаимоприемлемый день действительно существует, они достигнут цели в его поиске после порядка $\sqrt{N}$ вопросов.

Но как Алисе и Бобу задать вопрос $f_{AB}(x)$? Опишем, как им это сделать, действуя на любое одно из состояний вычислительного базиса $|x\rangle$. Сначала Алиса выполняет

$$|x\rangle|0\rangle \rightarrow |x\rangle|f_A(x)\rangle, \tag{6.188}$$

а затем посылает $n+1$ кубитов Бобу. Боб выполняет

$$|x\rangle|f_A(x)\rangle \rightarrow (-1)^{f_A(x)\wedge f_B(x)}|x\rangle|f_A(x)\rangle. \tag{6.189}$$

Это преобразование, очевидно, унитарно, и вы можете легко проверить, что Боб может осуществить его, обратившись к своему оракулу. Теперь, как и требовалось, фазовый множитель перед $|x\rangle$ равен $(-1)^{f_{AB}(x)}$, но в другом регистре продолжает храниться $|f_A(x)\rangle$, что будет портить когерентность суперпозиции значений x. Боб не может удалить этот регистр, но это может сделать Алиса. Тогда Боб посылает $n+1$ кубитов обратно Алисе, а она еще раз консультируется со своим оракулом, чтобы выполнить

$$(-1)^{f_A(x)\wedge f_B(x)}|x\rangle|f_A(x)\rangle \rightarrow (-1)^{f_A(x)\wedge f_B(x)}|x\rangle|0\rangle. \tag{6.190}$$

Обменявшись $2(n+1)$ кубитами, они завершили один вопрос из f_{AB} оракулу и, таким образом, могут выполнить одну итерацию Гровера.

---

[1] H. Burhman, *et. al.*, *Quantum vs. Classical Communication and Computation*, In *Proceedings of the 30th Annual ACM Symposium of Theory of Computing*, ACM Press, 1998; quant-ph/9802040.

Допустим, например, что Алиса и Боб знают, что имеется только одна взаимоприемлемая дата, но у них нет *априорной* информации о том, что это за день. После примерно $\frac{\pi}{4}\sqrt{N}$ итераций, требующих обменяться всего

$$Q \simeq \frac{\pi}{4}\sqrt{N} \cdot 2(\log N + 1) \tag{6.191}$$

кубитами, Алиса выполняет измерение, получая удобную дату с вероятностью, близкой к единице.

Таким образом, по крайней мере в этом частном случае, обмен $O(\sqrt{N}\log N)$ кубитами так же хорош, как и обмен $O(N)$ классическими битами. По-видимому, нужно быть осторожнее в интерпретации границы Холево, которая явно утверждает, что кубит имеет не большую способность переносить информацию, чем бит!

Если Алисе и Бобу заранее неизвестно, сколько имеется подходящих дат, то они тем не менее могут выполнить поиск Гровера (как мы отмечали в § 6.4.5) и с разумной вероятностью найти решение. С помощью $2 \cdot \log N$ битов классического сообщения они могут проверить, действительно ли найденная дата устраивает их обоих.

6.8.1. Сложность квантовой связи

В более общем виде можно представить, что каждый из нескольких участников обладает n-битовым входом; им необходимо вычислить функцию, зависящую от всех входов, с тем, чтобы ее значение в конце концов стало известно одному из них. Какой минимальный объем сообщений необходим для вычисления функции (детерминированного или вероятностного)? Хорошо изученный раздел классической теории сложности, которому адресуется этот вопрос, называется *сложностью связи*. То, что мы установили выше, является квадратичной границей между квантовой и классической сложностями связи для частного класса функций двух участников.

Помимо перехода от обмена классическими битами к обмену кубитами существуют другие интересные пути обобщения классической сложности связи. Например, предположим, что участники делят некоторое заранее подготовленное запутанное состояние (пары Белла или многокомпонентное запутывание) и могут использовать его наряду с классической связью, чтобы выполнить вычисление функции. Вновь непосредственно не очевидно, что разделенное запутывание упростит задачу, так как само по себе оно еще не позволяет участникам обмениваться классическими сообщени-

ями. Но оказывается, что запутывание *оказывает* помощь, по крайней мере небольшую[1].

В последнее время анализ сложности связи популярен среди теоретиков в области сложности, но эта дисциплина пока еще не представляется занимающей важное положение в практической технике связи. Возможно, это удивительно, принимая во внимание важность эффективного распределения вычислительной нагрузки в параллельных вычислениях, которые стали обычным делом. Более того, похоже, что в реальной жизни практически вся связь может рассматриваться как форма дистанционных вычислений. На самом деле мне не нужно получать все биты, дошедшие до меня по телефонной линии, особенно потому, что я скорее всего запомню только несколько битов информации, имеющих отношение к звонку в ближайшем будущем (фильм, на который мы решили сходить). Как менее прозаический пример, нам на Земле может быть необходимо связаться с роботом в глубоком космосе, чтобы проинструктировать, выходить ли ему на орбиту вокруг удаленной звездной системы. Так как ширина полосы предельно ограничена, то мы хотели бы вычислить правильный ответ на ДА/НЕТ-вопрос «Выходить ли на орбиту?» с помощью минимального обмена информацией между Землей и роботом.

Возможно, будущая цивилизация будет использовать известное квадратичное разделение между классической и квантовой сложностью связи, обмениваясь, скорее, кубитами, чем битами, со своей флотилией космических сил. А возможно, будет найдено экспоненциальное разделение, по крайней мере в определенных ситуациях, что существенно повысило бы стимул для развития необходимой технологии квантовой связи.

6.9. Периодичность

Проблема Саймона до сих пор является единственным примером, в котором мы нашли экспоненциальное разделение между скоростью квантового алгоритма и скоростью соответствующего классического алгоритма. Алгоритм Саймона использует квантовый параллелизм, чтобы ускорить поиск периода функции. Его успех вдохновляет нас искать другие квантовые алгоритмы, предназначенные для отыскания других разновидностей периода.

[1] R. Cleve, *et. al.*, *Quantum Entanglement and the Communication Complexity of the Inner Product Function*, Lect. Notes Comput. Sci., **1509**, 61–74 (1998), quant-ph/9708019; H. Buhrman, *et al*, ... *Complexity*, Phys. Rev., **A60**, 2737–2741 (1998), quant-ph/9710054.

Саймон изучал периодические функции, принимающие значения в $(Z_2)^n$. Для этой цели мощным инструментом служило n-битовое преобразование Адамара $\mathbf{H}^{(n)}$. Если вместо этого мы хотим изучать периодические функции, принимающие значения в Z_{2^n}, то инструментом сопоставимой силы будет (дискретное) преобразование Фурье.

Урок задачи Саймона в том, что, хотя поиск иголок в стоге сена может быть трудным, отыскание *периодически* распределенных иголок в стоге сена может оказаться гораздо проще. Например, если мы рассеиваем фотон на периодическом массиве иголок, он вероятнее всего рассеется в одном из преимущественных направлений, в котором удовлетворяется условие брэгговского отражения. Эти преимущественные направления зависят от расстояния между иголками. Таким образом, рассеяв только один фотон, мы уже приобретаем некоторую полезную информацию о периоде. При построении эффективных квантовых алгоритмов следует и дальше пользоваться смысловым подтекстом этой метафоры.

Итак, представим квантовый оракул, вычисляющий функцию

$$f\colon \{0,1\}^n \to \{0,1\}^m, \tag{6.192}$$

которая имеет неизвестный период r, где r – положительное целое число, удовлетворяющее

$$1 \ll r \ll 2^n. \tag{6.193}$$

То есть

$$f(x) = f(x + mr), \tag{6.194}$$

где m – произвольное целое число такое, что x и $x + mr$ лежат в $\{0, 1, 2, \ldots, 2^n - 1\}$. Мы должны найти период r. Рассматриваемая классически, эта проблема *трудная*. Если r, скажем, порядка $2^{n/2}$, нам необходимо обратиться к оракулу порядка $2^{n/2}$ раз, прежде чем мы, возможно, найдем два значения x, отображаемых на одно и то же значение $f(x)$, и, следовательно, что-нибудь узнаем о периоде r. Но, как мы увидим, существует квантовый алгоритм, определяющий r за время $\mathrm{poly}(n)$.

Даже если нам известно, как эффективно вычислять функцию $f(x)$, определение ее периода может оказаться трудной задачей. Наш квантовый алгоритм может быть применен для отыскания за $\mathrm{poly}(n)$ время периода любой функции, которую мы умеем вычислять за $\mathrm{poly}(n)$ время. Эффективное отыскание периода позволяет эффективно решать множество (по-видимому) трудных задач, таких как факторизация целого числа или вычисление дискретного логарифма[1].

[1] Дискретный логарифм определяется как минимальный положительный корень x уравне-

Ключевая идея, лежащая в основе квантового поиска периода, заключается в том, что преобразование Фурье может быть вычислено с помощью эффективной квантовой схемы (что было обнаружено Питером Шором). Квантовое преобразование Фурье (QFT) использует мощь квантового параллелизма, чтобы достичь экспоненциального ускорения хорошо известного (классического) быстрого преобразования Фурье (FFT). Поскольку FFT имеет такое широкое поле применений, то, возможно, однажды и QFT станет столь же широко распространенным.

6.9.1. Отыскание периода

QFT является унитарным преобразованием, действующим в вычислительном базисе согласно

$$\mathrm{QFT}\ :\ |x\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} e^{2\pi i x y/N} |y\rangle, \tag{6.195}$$

где $N = 2^n$. Будем пока предполагать, что мы эффективно выполняем QFT, и посмотрим, как это позволяет извлекать период функции $f(x)$.

Эмулируя алгоритм Саймона, мы сначала обратимся к оракулу с $\frac{1}{\sqrt{N}} \sum_x |x\rangle$ (легко приготавливаемым применением $\mathbf{H}^{(n)}$ к $|0\rangle$) и, таким образом, приготовим состояние

$$\frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} |x\rangle |f(x)\rangle. \tag{6.196}$$

Затем *измерим выходной* регистр, получая результат $|f(x_0)\rangle$ при некотором $0 \leqslant x_0 < r$. Это измерение готовит во входном регистре когерентную суперпозицию A значений x, которые отображаются на $f(x_0)$:

$$\frac{1}{\sqrt{A}} \sum_{j=0}^{A-1} |x_0 + jr\rangle, \tag{6.197}$$

где

$$N - r \leqslant x_0 + (A-1)r < N \tag{6.198}$$

ния $a^x = b (\mathrm{mod}\, p)$, где все переменные являются целыми числами. Вычисление дискретного логарифма является сложной вычислительной проблемой, что находит применение в криптографии. Квантовый алгоритм решения этой задачи см. в книге М. Нильсен, И. Чанг, *Квантовые вычисления и квантовая информация*, М., Мир (2006). — *Прим. ред.*

или

$$A - 1 < \frac{N}{r} < A + 1. \tag{6.199}$$

На самом деле измерение выходного регистра не обязательно. Если его пропустить, то состоянием входного регистра будет некогерентная суперпозиция (просуммированная по $x_0 \in \{0, 1, 2, \ldots, r-1\}$) состояний вида (6.197). Остальная часть алгоритма также хорошо работает, действуя на это начальное состояние.

Теперь наша задача состоит в том, чтобы извлечь значение r из состояния (6.197). Если бы мы на этом этапе измерили входной регистр, проецируя его на вычислительный базис, то мы бы ничего не узнали относительно r. Вместо этого (ср. с алгоритмом Саймона) следует сначала выполнить преобразование Фурье, а *уже затем* проводить измерение.

Применяя QFT к состоянию (6.197), получим

$$\frac{1}{\sqrt{NA}} \sum_{y=0}^{N-1} e^{2\pi i x_0 y/N} \sum_{j=0}^{A-1} e^{2\pi i j r y/N} |y\rangle. \tag{6.200}$$

Если мы теперь выполним измерение в вычислительном базисе, то вероятность получения результата y будет равна

$$\mathrm{Prob}(y) = \frac{A}{N} \left| \frac{1}{A} \sum_{j=0}^{A-1} e^{2\pi i j r y/N} \right|^2. \tag{6.201}$$

Это распределение резко выделяет такие значения y, при которых yr/N близко к целому числу. Например, если N/r оказывается целым (и, следовательно, равным A), то

$$\mathrm{Prob}(y) = \frac{1}{r} \left| \frac{1}{A} \sum_{j=0}^{A-1} e^{2\pi i j y/A} \right|^2 = \begin{cases} \frac{1}{r}, & y = A \cdot (\text{целое}), \\ 0 & \text{в противном случае.} \end{cases} \tag{6.202}$$

В более общем случае мы можем просуммировать геометрическую прогрессию

$$\sum_{j=0}^{A-1} e^{i\theta j} = \frac{e^{iA\theta} - 1}{e^{i\theta} - 1}, \tag{6.203}$$

где

$$\theta = \frac{2\pi yr(\mathrm{mod\,N})}{N}. \tag{6.204}$$

Существует точно r значений $y \in \{0, 1, 2, \ldots, N-1\}$, удовлетворяющих

$$-\frac{r}{2} \leqslant yr(\mathrm{mod\,N}) \leqslant \frac{r}{2}. \tag{6.205}$$

(Чтобы убедиться в этом, представим промаркированные кратные r и N в ряду чисел, простирающемся от 0 до $rN-1$. Для каждого кратного N существует кратное r, удаленное от него не более чем на расстояние $r/2$) Для каждого из этих значений соответствующее θ удовлетворяет

$$-\pi\frac{r}{N} \leqslant \theta \leqslant \pi\frac{r}{N}. \tag{6.206}$$

Теперь, поскольку $A-1 < \frac{N}{r}$, то при этих значениях θ все слагаемые в сумме по j в уравнении (6.203) лежат в одной полуплоскости, следовательно, они интерферируют конструктивно и сумма становится значительной.

Мы знаем, что

$$|1 - e^{i\theta}| \leqslant |\theta|, \tag{6.207}$$

поскольку расстояние по прямой от начала координат меньше, чем длина дуги вдоль окружности, а при $A|\theta| \leqslant \pi$

$$|1 - e^{iA\theta}| \geqslant \frac{2A|\theta|}{\pi}, \tag{6.208}$$

так как мы можем увидеть (графически или вычислив его производную), что это расстояние является выпуклой функцией. На самом деле $A < \frac{N}{r} + 1$ и, следовательно, $A\theta < \pi\left(1 + \frac{r}{N}\right)$, но, применяя вышеуказанную границу к

$$\left|\frac{e^{i(A-1)\theta} - 1}{e^{i\theta} - 1} + e^{i(A-1)\theta}\right| \geqslant \left|\frac{e^{i(A-1)\theta} - 1}{e^{i\theta} - 1}\right| - 1, \tag{6.209}$$

мы тем не менее можем прийти к выводу, что

$$\left|\frac{e^{iA\theta} - 1}{e^{i\theta} - 1}\right| \geqslant \frac{2(A-1)|\theta|}{\pi|\theta|} - 1 = \frac{2}{\pi}A - \left(1 + \frac{2}{\pi}\right). \tag{6.210}$$

Пренебрегая возможной поправкой порядка $2/A$, мы находим

$$\text{Prob}(y) \geqslant \left(\frac{4}{\pi^2}\right)\frac{1}{r} \tag{6.211}$$

для каждого из r значений y, удовлетворяющих неравенству (6.205). Следовательно, с вероятностью, не ниже чем $4/\pi^2$ измеренное значение y будет удовлетворять

$$k\frac{N}{r} - \frac{1}{2} \leqslant y \leqslant k\frac{N}{r} + \frac{1}{2}, \tag{6.212}$$

или

$$\frac{k}{r} - \frac{1}{2N} \leqslant \frac{y}{N} \leqslant \frac{k}{r} + \frac{1}{2N}, \tag{6.213}$$

где k — целое число, выбранное из $\{0, 1, 2, \ldots, r-1\}$. Результат вычисления с разумной вероятностью находится не дальше чем на расстоянии $1/2$ от целого кратного числа N/r.

Пусть нам известно, что $r < M \ll N$. Таким образом, N/r — рациональное число со знаменателем, меньше m. Два различных рациональных числа, знаменатель каждого из которых меньше M, не могут быть ближе друг к другу чем на $1/M^2$, так как $\frac{a}{b} - \frac{c}{d} = \frac{ad-bc}{bd}$. Если результат измерения y удовлетворяет неравенству (6.212), то существует *единственное* значение k/r (при $r < M$), определяемое y/N, при условии, что $N \geqslant M^2$. Это значение k/r может быть успешно извлечено из измеренного y/N с помощью метода цепных дробей.

С вероятностью, превышающей $4/\pi^2$, мы нашли значение k/r, где k выбирается (примерно равновероятно) из $\{0, 1, 2, \ldots, r-1\}$. С приемлемой вероятностью k и r являются взаимно простыми (не имеющими общего множителя), так что мы добились успеха в отыскании r. Обратившись к оракулу, мы можем проверить, действительно ли $f(x) = f(x+r)$. Но если $\text{GCD}(k, r) \neq 1$,[1] то мы нашли всего лишь r_1, множитель r.

Если мы не достигли успеха, то мы могли бы проверить некоторые близкие значения y [измеренное значение могло оказаться вблизи интервала $-r/2 \leqslant yr(\text{mod}N) \leqslant r/2$, в действительности не находясь внутри его] или проверить несколько множителей r [значение $\text{GCD}(k, r)$, если оно не равно единице, по-видимому, невелико]. Если не удается и это, то мы прибегнем к повторению квантовой схемы, получая (с вероятностью не ни-

[1] GCD (Greatest Common Divisor) — наибольший общий делитель. — *Прим. перев.*

же $4/\pi^2$) на этот раз значение k'/r. Теперь k' также может иметь общий множитель с r, в таком случае наша процедура вновь определяет r_2, множитель r. Но с достаточно высокой вероятностью $\mathrm{GCD}(k, k') = 1$, в таком случае $r = \mathrm{LCM}(r_1, r_2)$.[1] Действительно, мы можем вычислить вероятность того, что случайно выбранные k и k' являются взаимно простыми, следующим образом. Так как вероятность того, что простое число p делит случайно выбранное число, равна $1/p$, то вероятность того, что p делит без остатка оба k и k', равна $1/p^2$. А k и k' являются взаимно простыми тогда и только тогда, когда не существует простого числа p, делящего их обоих без остатка. Следовательно,

$$\mathrm{Prob}(k, k' \text{ взаимно простые}) = \prod_{\text{простые } p} \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \simeq 0{,}607 \quad (6.214)$$

[где $\zeta(z)$ обозначает дзета-функцию Римана]. Таким образом, после некоторого постоянного (не зависящего от N) количества повторений алгоритма мы наверняка добьемся успеха в поиске периода r.

6.9.2. От FFT к QFT

Рассмотрим теперь реализацию квантового преобразования Фурье. Преобразование Фурье

$$\sum_x f(x)|x\rangle \to \sum_y \left(\frac{1}{\sqrt{N}} \sum_x e^{2\pi i x y/N} f(x)\right) |y\rangle \quad (6.215)$$

представляет собой перемножение унитарных $N \times N$-матриц, где матричный (x, y)-элемент равен $\left(e^{2\pi i/N}\right)^{xy}$. С наивной точки зрения это преобразование требует $O(N^2)$ элементарных операций. Существует, однако, хорошо известная и очень полезная (классическая) процедура, сокращающая количество операций до $O(N \log N)$. Предполагая, что $N = 2^n$, представим x и y в виде двоичных разложений

$$\begin{aligned} x &= x_{n-1} \cdot 2^{n-1} + x_{n-2} \cdot 2^{n-2} + \ldots + x_1 \cdot 2 + x_0, \\ y &= y_{n-1} \cdot 2^{n-1} + y_{n-2} \cdot 2^{n-2} + \ldots + y_1 \cdot 2 + y_0. \end{aligned} \quad (6.216)$$

[1] LCM (Least Common Multiple) — наименьшее общее кратное. — *Прим. перев.*

В произведении x и y можно отбросить любые слагаемые, содержащие n-ю и более высокие степени двух, поскольку они не вносят вклада в $e^{2\pi i xy/2^n}$. Следовательно,

$$\frac{xy}{2^n} \equiv y_{n-1}(.x_0) + y_{n-2}(.x_1x_0) + y_{n-3}(.x_2x_1x_0) + \ldots + \\ + y_1(.x_{n-2}x_{n-3}\ldots x_0) + y_0(.x_{n-1}x_{n-2}\ldots x_0), \tag{6.217}$$

где множители в круглых скобках представляют собой значения соответствующих двоичных разрядов, например:

$$x_2x_1x_0 = \frac{x_2}{2} + \frac{x_1}{2^2} + \frac{x_0}{2^3}. \tag{6.218}$$

Теперь мы можем вычислить

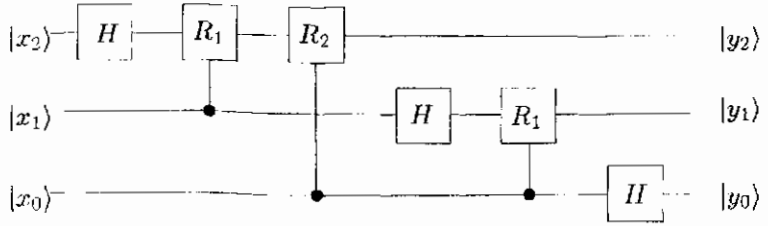
$$\tilde{f}(x) = \frac{1}{\sqrt{N}} \sum_y e^{2\pi i xy/N} f(y) \tag{6.219}$$

для каждого из N значений x. Но сумма по y факторизуется на n сумм по $y_k = 0, 1$, которые могут быть последовательно вычислены за время порядка n.

С помощью квантового параллелизма можно добиться гораздо лучшего результата. Из (6.217) мы получим

$$\begin{aligned} \mathrm{QFT} : |x\rangle \rightarrow & \frac{1}{\sqrt{N}} \sum_y e^{2\pi i xy/N} |y\rangle \\ = & \frac{1}{\sqrt{2^n}} \left(|0\rangle + e^{2\pi i(.x_0)}|1\rangle\right)\left(|0\rangle + e^{2\pi i(.x_1x_0)}|1\rangle\right) \\ & \ldots \left(|0\rangle + e^{2\pi i(.x_{n-1}x_{n-2}\ldots x_0)}|1\rangle\right). \end{aligned} \tag{6.220}$$

QFT преобразует каждое состояние вычислительного базиса в *незапутанное* состояние n кубитов; таким образом, мы ожидаем, что оно может быть эффективно реализовано. Действительно, рассмотрим случай $n = 3$. Нетрудно понять, что эту работу выполняет схема



(но обратим внимание на то, что порядок следования битов на выходе инвертировался). Каждый вентиль Адамара действует как

$$\mathbf{H}\colon |x_k\rangle \to \frac{1}{\sqrt{2}}\left(|0\rangle + e^{2\pi i(.x_k)}|1\rangle\right). \tag{6.221}$$

Другие вклады в относительную фазу векторов $|0\rangle$ и $|1\rangle$ в k-ом кубите обеспечиваются двухкубитовыми условными поворотами, где

$$\mathbf{R}_d = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/2^d} \end{pmatrix}, \tag{6.222}$$

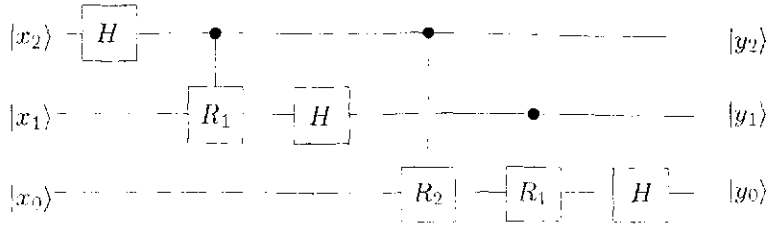
а $d = (k-j)$ – «расстояние» между кубитами.

В случае $n = 3$ QFT строится из трех вентилей $\mathbf{H}$ и трех вентилей контролируемое $\mathbf{R}$. В общем случае очевидное обобщение этой схемы требует n вентилей $\mathbf{H}$ и $\binom{n}{2} = \frac{1}{2}n(n-1)$ контролируемых $\mathbf{R}$. Вновь двухкубитовый вентиль применяется к каждой паре кубитов с контролируемой относительной фазой $\pi/2^d$, где d – «расстояние» между кубитами. Таким образом, семейство схем, осуществляющее QFT, имеет размер порядка $(\log N)^2$.

Мы можем сократить сложность схемы до линейной по $\log N$, если готовы ограничиться реализацией фиксированной точности, поскольку двухкубитовые вентили, действуя на значительно разделенные кубиты, вносят лишь экспоненциально малые фазы. Если мы опустим вентили, действующие на пары, разделенные более чем на m, тогда каждое слагаемое в (6.217) заменяется приближением с m-битовой точностью; полная ошибка в $xy/2^n$ наверняка не хуже, чем $n2^{-m}$, следовательно, мы можем достичь точности ε в $xy/2^n$ при $m \geqslant \log n/\varepsilon$. Если мы сохраним вентили, действующие только на пары кубитов с расстоянием m или менее, то размер схемы будет равен $mn \sim n \log n/\varepsilon$.

Фактически, если мы собираемся измерять в вычислительном базисе непосредственно после выполнения QFT (или его обращения), то возможно дальнейшее упрощение – двухкубитовые вентили не нужны вообще! Заметим, во-первых, что вентиль контролируемое $\mathbf{R}_d$ действует симметричным образом на два кубита — он действует тривиально на $|00\rangle$, $|01\rangle$ и $|10\rangle$ и изменяет фазу $|11\rangle$ на $e^{i\theta_d}$. Таким образом, без модификации вентиля можно менять местами «контрольный» и «целевой» биты. С учетом этой замены наша схема для трехкубитового QFT может быть изображена

заново как



Коль скоро $|y_0\rangle$ измерено, нам *известно* значение бита, который управляет вентилем $\mathbf{R}_1$, действующим на два первых кубита. Следовательно, мы получим то же самое распределение вероятностей результатов измерений, если вместо применения контролируемого $\mathbf{R}_1$ и последующего измерения мы сначала измерим y_0, а затем применим к следующему кубиту поворот $(\mathbf{R}_1)^{y_0}$, обусловленный результатом измерения первого кубита. Аналогично можно заменить вентили контролируемое $\mathbf{R}_1$ и контролируемое $\mathbf{R}_2$ однокубитовым поворотом

$$(\mathbf{R}_2)^{y_0}(\mathbf{R}_1)^{y_1} \tag{6.223}$$

[то есть поворотом с относительной фазой $\pi(.y_1y_0)$], действующим на третий кубит *после* того, как были измерены значения y_1 и y_0.

В общем случае, если мы собираемся измерять после выполнения QFT, то для его осуществления необходимо только n вентилей Адамара и $n-1$ однокубитовых поворотов. Процедура QFT замечательно проста!

6.10. Факторизация

6.10.1. Факторизация как отыскание периода

Что связывает проблему факторизации (поиск простых множителей большого составного положительного целого числа) с периодичностью? Существует хорошо известная (рандомизованная) редукция факторизации к определению периода функции. Хотя эта редукция непосредственно не связана с квантовыми вычислениями, мы обсудим ее здесь для полноты, а также и потому, что перспектива использования квантового компьютера как инструмента факторизации вызывает столь сильное волнение.

Допустим, мы хотим найти множитель n-битового числа N. Выберем псевдослучайным образом $a < N$ и вычислим наибольший общий дели-

тель $\mathrm{GCD}(a, N)$, что можно эффективно [за время порядка $(\log N)^3$] выполнить с помощью алгоритма Евклида. Если $\mathrm{GCD}(a, N) \neq 1$, тогда GCD является нетривиальным множителем числа N и задача решена. Но предположим, что $\mathrm{GCD}(a, N) = 1$.

Отступление: алгоритм Евклида. Чтобы вычислить $\mathrm{GCD}(N_1, N_2)$ (при $N_1 > N_2$) разделим сначала N_1 на N_2, получая остаток R_1. Затем разделим N_2 на R_1, получая остаток R_2. Разделим R_1 на R_2 *и так далее* до тех пор, пока не получим в остатке нуль. Последний ненулевой остаток и есть $R = \mathrm{GCD}(N_1, N_2)$. Чтобы убедиться в том, что алгоритм работает, заметим лишь, что: (1) на R делятся все предыдущие остатки и, следовательно, N_1 и N_2; (2) *любое* число, на которое делятся N_1 и N_2, делит все остатки, включая R. Общий делитель N_1 и N_2, который в свою очередь делится на все остальные общие делители этих чисел, есть не что иное как $\mathrm{GCD}(N_1, N_2)$. Чтобы понять, сколько времени занимает алгоритм Евклида, заметим, что

$$R_j = qR_{j+1} + R_{j+2}, \tag{6.224}$$

где $q \geqslant 1$, а $R_{j+2} < R_{j+1}$; следовательно, $R_{j+2} < \frac{1}{2}R_j$. Два деления сокращают остаток как минимум в два раза, так что требуется не более чем $2\log N_1$ делений, каждое из которых использует $O\big((\log N)^2\big)$ элементарных операций; полное количество операций имеет порядок $O\big((\log N)^3\big)$.

Числа $a < N$, взаимно простые с N (не имеющие общего множителя с N), образуют конечную группу относительно умножения по модулю N. [Почему? Нам нужно установить, что каждый элемент a имеет обратный. Но для данного $a < N$, взаимно простого с N, и каждого $b < N$, пробегающего все взаимно простые с N значения, все произведения $ab \pmod N$ различны[1]. Следовательно, для некоторого b мы должны иметь $ab \equiv 1 \pmod N$; таким образом, число, обратное к a, существует.] Каждый элемент a этой конечной группы имеет конечный *порядок* r, наименьшее положительное целое число такое, что

$$a^r \equiv 1 \pmod N. \tag{6.225}$$

Порядок a по модулю N является периодом функции

$$f_{N,a}(x) = a^x \pmod N. \tag{6.226}$$

[1] Если N — делитель числа $ab - ab'$, то оно должно делителем и $b - b'$.

Мы знаем, что существует эффективный квантовый алгоритм, позволяющий найти период функции; следовательно, если мы можем эффективно вычислить $f_{N,a}$, то также эффективно можем найти и порядок a.

На первый взгляд, вычисление $f_{N,a}$ может показаться трудным, поскольку показатель степени x может быть очень большим. Но если $x < 2^m$ и мы представляем x в виде двоичного разложения

$$x = x_{m-1} \cdot 2^{m-1} + x_{m-2} \cdot 2^{m-2} + \ldots + x_1 \cdot 2 + x_0, \tag{6.227}$$

то

$$a^x (\mathrm{mod}\, N) = \left(a^{2^{m-1}}\right)^{x_{m-1}} \left(a^{2^{m-2}}\right)^{x_{m-2}} \cdots (a)^{x_0} (\mathrm{mod}\, N). \tag{6.228}$$

Каждое a^{2^j} имеет большой показатель степени, тем не менее оно может быть эффективно вычислено *классическим* компьютером путем повторного возведения в квадрат:

$$a^{2^j} (\mathrm{mod}\, N) = \left(a^{2^{j-1}}\right)^2 (\mathrm{mod}\, N). \tag{6.229}$$

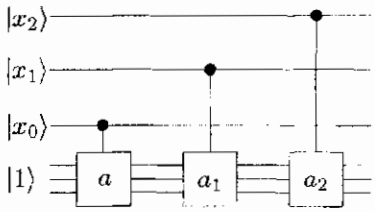
То есть необходимо только $m-1$ (классических) умножений по модулю N, чтобы собрать таблицу всех a^{2^j}.

Вычисление $a^x (\mathrm{mod}\, N)$ выполняет программа

INPUT 1

For $j = 0$ to $m-1$, if $x_j = 1$, MULTIPLY a^{2^j}.

Она требует максимум m умножений по модулю N, каждое из которых требует порядка $(\log N)^2$ элементарных операций[1]. Поскольку $r < N$, у нас будут неплохие шансы успешно выделить период, если выбрать $m \sim 2 \log N$. Следовательно, вычисление $f_{N,a}$ может быть выполнено семейством схем размера $O\left((\log N)^3\right)$, имеющих следующую структуру:



[1] Используя разные трюки для выполнения эффективного перемножения очень больших чисел, количество элементарных операций можно сократить до $O(\log N \log\log N \times \times \log\log\log N)$; таким образом, асимптотически при больших N семейство схем размера $O(\log^2 N \log\log N \log\log\log N)$ может вычислить $f_{N,a}$.

Умножение на a^{2^j} выполняется, если контрольный кубит x_j имеет значение 1.

Предположим, что мы нашли r, период a по модулю N. Тогда, *если* r четное, то

$$N \text{ является делителем числа } \left(a^{r/2}+1\right)\left(a^{r/2}-1\right). \tag{6.230}$$

Очевидно, что N не делит $\left(a^{r/2}-1\right)$; если бы это было так, то порядок a был бы $\leqslant r/2$. Таким образом, *если* наряду с этим N не является делителем $\left(a^{r/2}+1\right)$ или

$$a^{r/2} \neq -1 (\mathrm{mod}\, N), \tag{6.231}$$

тогда N должен иметь нетривиальный общий множитель с каждым из $a^{r/2}\pm 1$. Следовательно, $\mathrm{GCD}(N, a^{r/2}+1) \neq 1$ является множителем N (что можно эффективно определить с помощью классических вычислений), то есть задача решена.

Таким образом, коль скоро найдено r, мы успешно факторизовали N, за исключением случаев, когда (1) r нечетное или (2) r четное и $a^{r/2} \equiv -1 (\mathrm{mod}\, N)$. Насколько вероятен этот успех?

Допустим, что N является произведением двух простых множителей $p_1 \neq p_2$:

$$N = p_1 p_2 \tag{6.232}$$

(это фактически самый неблагоприятный случай). Для каждого $a < p_1 p_2$ существуют единственные $a_1 < p_1$ и $a_2 < p_2$ такие, что

$$\begin{aligned} a &\equiv a_1 (\mathrm{mod}\, p_1), \\ a &\equiv a_2 (\mathrm{mod}\, p_2). \end{aligned} \tag{6.233}$$

Следовательно, случайный выбор $a < N$ эквивалентен случайному выбору $a_1 < p_1$ и $a_2 < p_2$.

Отступление: мы используем **Китайскую теорему об остатках**. Решение a уравнений (6.233) единственно, поскольку если a и b являются решениями, то p_1 и p_2 должны быть делителями $a-b$. Решение существует, поскольку каждое $a < p_1 p_2$ решает уравнения (6.233) при *некоторых* a_1 и a_2. А так как имеется ровно $p_1 p_2$ способов выбрать a_1 и a_2 и ровно $p_1 p_2$ способов выбрать a, то единственность означает, что для каждой пары a_1, a_2 существует соответствующее ей a.

Пусть теперь r_1 обозначает порядок a_1 по модулю p_1, а r_2 – порядок a_2 по модулю p_2. Китайская теорема об остатках утверждает, что $a^r \equiv 1 (\mathrm{mod}\, p_1 p_2)$ эквивалентно тому, что

$$\begin{aligned} a_1^r &\equiv 1 (\mathrm{mod}\, p_1), \\ a_2^r &\equiv 1 (\mathrm{mod}\, p_2). \end{aligned} \tag{6.234}$$

Следовательно, $r = \mathrm{LCM}(r_1, r_2)$. Если r_1 и r_2 оба нечетны, то таковым же является r, и мы проигрываем.

Но если *одно* из двух чисел, r_1 или r_2, четное, то таковым же является r, и мы продолжаем игру. Если

$$\begin{aligned} a_1^{r/2} &\equiv -1 (\mathrm{mod}\, p_1), \\ a_2^{r/2} &\equiv -1 (\mathrm{mod}\, p_2), \end{aligned} \tag{6.235}$$

то $a^{r/2} \equiv -1 (\mathrm{mod}\, p_1 p_2)$, и мы снова проигрываем. Но если или

$$\begin{aligned} a_1^{r/2} &\equiv -1 (\mathrm{mod}\, p_1), \\ a_2^{r/2} &\equiv 1 (\mathrm{mod}\, p_2), \end{aligned} \tag{6.236}$$

или

$$\begin{aligned} a_1^{r/2} &\equiv 1 (\mathrm{mod}\, p_1), \\ a_2^{r/2} &\equiv -1 (\mathrm{mod}\, p_2), \end{aligned} \tag{6.237}$$

то $a^{r/2} \not\equiv -1 (\mathrm{mod}\, p_1 p_2)$, и мы выигрываем. [Конечно, одновременное равенство $a_1^{r/2} \equiv 1 (\mathrm{mod}\, p_1)$ и $a_2^{r/2} \equiv 1 (\mathrm{mod}\, p_2)$ невозможно, что означало бы $a^{r/2} \equiv 1 (\mathrm{mod}\, p_1 p_2)$, то есть r не могло бы быть порядком a.]

Допустим, что

$$\begin{aligned} r_1 &= 2^{c_1} \cdot (\text{нечетное число}), \\ r_2 &= 2^{c_2} \cdot (\text{нечетное число}), \end{aligned} \tag{6.238}$$

где $c_1 > c_2$. Тогда $r = \mathrm{LCM}(r_1, r_2) = 2r_2 \cdot (\text{целое число})$, так что $a^{r/2} \equiv 1 (\mathrm{mod}\, p_2)$, а уравнения (6.236) удовлетворяются – мы победили! Аналогично $c_2 > c_1$ подразумевает уравнение (6.237) – мы снова в выигрыше! Но при $c_1 = c_2$ имеет место $r = r_1 \cdot (\text{нечетное число}) = r_1 \cdot (\text{нечетное число}')$, так что удовлетворяются уравнения (6.235) – в таком случае мы проиграли.

Итак, все сводится к альтернативе: при $c_1 = c_2$ мы проигрываем, а при $c_1 \neq c_2$ – побеждаем. Насколько вероятно, что $c_1 \neq c_2$?

Полезно знать, что мультипликативная группа по модулю p является циклической — она имеет такой образующий элемент порядка $p-1$, что все элементы являются его степенями. [Почему? Множество целых чисел по модулю p образуют конечное *поле*. Если бы группа не была циклической, то максимальный порядок ее элементов был бы $q < p-1$, так что $x^q \equiv 1 (\mathrm{mod}\, p)$ имело бы $p-1$ решений. Но это невозможно: на конечном поле существует не более чем q корней q-ой степени из единицы.]

Допустим, что $p - 1 = 2^k \cdot s$, где s — нечетное, и рассмотрим порядки всех элементов циклической группы порядка $p-1$. Для краткости мы обсудим только самый неблагоприятный для нас случай $k = 1$. Тогда, если b является образующим элементом (имеет порядок $2s$), то четные степени b имеют нечетный порядок, а нечетные — порядок $2 \cdot$ (нечетное число). Тогда в этом случае $r = 2^c \cdot$ (нечетное число), где c с равной вероятностью принимает одно из значений $\{0, 1\}$. Следовательно, если p_1 и p_2 оба такого (неподходящего) типа, а a_1, a_2 выбираются случайно, вероятность того, что $c_1 \neq c_2$, равна $1/2$. Следовательно, раз мы нашли r, то вероятность успешного отыскания множителя как минимум равна $1/2$, если N является произведением двух простых чисел. Если же N имеет больше двух различных простых множителей, то наши нечетные числа даже лучше. Этот метод терпит неудачу, если N является степенью простого числа $N = p^\alpha$, но степени простых чисел успешно факторизуются другими методами.

6.10.2. RSA

Интересует ли кого-нибудь, простой или трудной является факторизация? Некоторых это очень интересует.

Предполагаемая сложность факторизации является основой надежности широко используемой схемы RSA для криптографии с открытым ключом[1], которой вы можете воспользоваться, даже если посылаете через интернет номер своей кредитной карточки.

Идея криптографии с открытым ключом заключается в том, чтобы избежать необходимости обмена секретным ключом (который может быть перехвачен и скопирован) между желающими установить связь с партнерами. Шифрующий ключ общеизвестен. Но его использование для того,

[1] R. L. Rivest, A. Shamir, L. M. Adleman, *A Method of Obtaining Digital Signatures and Public-Key Cryptosystems*, Comm. ACM, **21(2)**, 120–126 (1978).

чтобы извлечь дешифрующий ключ, требует непомерно сложных вычислений. Следовательно, Боб может послать шифрующий ключ Алисе и кому угодно, но только он будет в состоянии декодировать сообщение, которое Алиса (или кто-нибудь другой) закодирует с помощью этого ключа. Кодирование является «односторонней функцией», которую легко вычислить, но очень трудно обратить.

(Конечно, Алиса и Боб могли бы избежать необходимости обмена открытым ключом, если бы они выбрали конфиденциальный ключ во время их предыдущей тайной встречи. Например, они могли бы договориться использовать длинную случайную строку в качестве разового шифра для кодирования и декодирования. Но, возможно, Алиса и Боб никогда не задумывались о том, что однажды им понадобится тайно связаться друг с другом. Или, возможно, ранее они договорились использовать разовый шифр, но уже израсходовали свои тайные ключи и не имеют желания вновь пользоваться ими, опасаясь, что тогда их код смогут взломать шпионы. Сейчас они слишком далеко друг от друга, чтобы безопасно обменяться новым тайным ключом; их самым надежным выбором оказывается криптография с открытым ключом.)

Чтобы построить открытый ключ, Боб выбирает два больших простых числа p и q. Но он не открывает их значения, а вместо этого вычисляет произведение

$$N = pq. \tag{6.239}$$

Поскольку Боб знает разложение N на простые множители, то ему известно и значение функции Эйлера $\varphi(N)$ — количества чисел, меньших чем N, являющихся взаимно простыми с N. В случае произведения двух простых чисел это

$$\varphi(N) = N - p - q + 1 = (p-1)(q-1) \tag{6.240}$$

(только числа, кратные p и q, имеют общий множитель с N). Найти $\varphi(N)$ несложно, если вы знаете разложение N на простые множители, но это трудно, если вам известно только N.

Затем Боб псевдослучайным образом выбирает $e < \varphi(N)$, взаимно простое с $\varphi(N)$. Он открывает Алисе (и любому подслушивающему) значения N и e, но ничего сверх этого.

Алиса преобразует свое сообщение в ASCII[1], число $a < N$. Она кодирует сообщение, вычисляя

$$b = f(a) = a^e (\mathrm{mod}\, N), \tag{6.241}$$

[1] ASCII — American Standard Code for Information Interchange — американский стандарт кода для обмена информацией. — *Прим. перев.*

что можно быстро выполнить путем повторного возведения в квадрат. Как теперь Бобу декодировать это сообщение?

Допустим, что a является взаимно простым с N (что имеет подавляющую вероятность, если p и q очень большие — во всяком случае Алиса может проверить это прежде, чем кодировать). Тогда

$$a^{\varphi(N)} \equiv 1 (\mathrm{mod}\, N) \tag{6.242}$$

(теорема Эйлера). Это так, поскольку числа, меньшие N и взаимно простые с N, образуют группу [порядка $\varphi(N)$] относительно умножения по модулю N. Порядок любого элемента группы должен быть делителем порядка группы (степени a образуют подгруппу). Поскольку $\mathrm{GCD}(e, \varphi(N)) = 1$, то нам известно, что e имеет мультипликативное обратное $d = e^{-1}$ по модулю $\varphi(N)$:

$$ed \equiv 1 (\mathrm{mod}\, \varphi(N)). \tag{6.243}$$

Значение d является строго сохраняемым секретом Боба; он использует его для декодирования, вычисляя

$$\begin{aligned} f^{-1}(b) &= b^d (\mathrm{mod}\, N) = \\ &= a^{ed} (\mathrm{mod}\, N) = \\ &= a\left(a^{\varphi(N)}\right)^{(\text{целое число})} (\mathrm{mod}\, N) = \\ &= a (\mathrm{mod}\, N). \end{aligned} \tag{6.244}$$

Отступление. Как Бобу вычислить $d = e^{-1}$? Мультипликативное обратное является побочным продуктом выполнения алгоритма Евклида вычисления $\mathrm{GCD}(e, \varphi(N)) = 1$. Проследим цепочку остатков снизу вверх, начиная с $R_n = 1$:

$$\begin{aligned} 1 = R_n &= R_{n-2} - q_{n-1} R_{n-1}, \\ R_{n-1} &= R_{n-3} - q_{n-2} R_{n-2}, \\ R_{n-2} &= R_{n-4} - q_{n-3} R_{n-3} \\ &\qquad \text{и так далее} \ldots \end{aligned} \tag{6.245}$$

(где q_j — частные), так что

$$\begin{aligned} 1 &= (1 + q_{n-1} q_{n-2}) R_{n-2} - q_{n-1} R_{n-3}, \\ 1 &= \left(-q_{n-1} - q_{n-3}(1 + q_{n-1} q_{n-2})\right) R_{n-3} + (1 + q_{n-1} q_{n-2}) R_{n-4} \\ &\qquad \text{и так далее} \ldots. \end{aligned} \tag{6.246}$$

Продолжая, мы можем выразить единицу через линейную комбинацию любых двух следующих друг за другом остатков; в конце концов, мы пройдем весь путь вплоть до

$$1 = d \cdot e + q \cdot \varphi(N) \tag{6.247}$$

и идентифицируем d как $e^{-1}(\mathrm{mod}\,\varphi(N))$.

Конечно, если Ева имеет средство сверхбыстрой факторизации, то RSA-схема ненадежна. Она факторизует N, найдет $\varphi(N)$ и быстро вычислит d. На самом деле ей даже не нужно факторизовать N; достаточно вычислить порядок по модулю N закодированного сообщения $a^e(\mathrm{mod}\,N)$. Так как e является взаимно простым с $\varphi(N)$, то порядок $a^e(\mathrm{mod}\,N)$ тот же самый, что и порядок a (оба элемента генерируют одну и ту же *орбиту* или циклическую подгруппу). Как только порядок $\mathrm{Ord}(a)$ становится известен, Ева вычисляет $\tilde{d}$ такое, что

$$\tilde{d}e \equiv 1(\mathrm{mod}\,\mathrm{Ord}(a)), \tag{6.248}$$

так что

$$\left(a^e\right)^{\tilde{d}} \equiv a \cdot \left(a^{\mathrm{Ord}(a)}\right)^{(\text{целое число})} \equiv a(\mathrm{mod}\,N), \tag{6.249}$$

и Ева может дешифровать сообщение. Если нашей единственной целью является взлом RSA, то мы обратимся к алгоритму Шора, чтобы найти $r = \mathrm{Ord}(a^e)$, и нас не должно беспокоить то, можем мы или нет использовать r, чтобы выделить множитель N.

Насколько важна такая перспектива криптографических применений квантовых вычислений? Когда быстрые квантовые компьютеры станут доступной реальностью, заинтересованные стороны могут прекратить использование RSA или могут использовать более длинные ключи, чтобы оставаться на шаг впереди современной техники. Однако люди, имеющие секреты, иногда хотят, чтобы их сообщения до поры (лет 30?) оставались конфиденциальными. Их могут не устроить более длинные ключи, если они не уверены относительно темпов будущих технологических достижений.

А если они избегают RSA, то чем они будут пользоваться вместо этого? Известно не так много подходящих односторонних функций, да и они, а не только RSA, (могут быть) беззащитны перед квантовой атакой. То есть на самом деле многое поставлено на карту. Если станут доступными быстрые большие квантовые компьютеры, то зашифрованная информация станет легко доступной.

Но квантовая теория, одной рукой отбирая, другой дает; квантовые компьютеры могут скомпрометировать схемы открытых ключей, но вместе с этим — предложить альтернативу: обсуждавшееся в четвертой главе безопасное распределение квантового ключа.

6.11. Определение фазы

Существует альтернативный взгляд на алгоритм факторизации (предложенный Китаевым), углубляющий наше понимание того, как он работает: мы можем факторизовать, поскольку можем эффективно и точно измерять собственные значения некоторого унитарного оператора.

Рассмотрим $a < N$, взаимно простое с N. Пусть x принимает значения из $\{0, 1, 2, \ldots, N-1\}$ и пусть $\mathbf{U}_a$ обозначает унитарный оператор

$$\mathbf{U}_a \; : \; |x\rangle \; \rightarrow \; |ax(\mathrm{mod}\, N)\rangle. \tag{6.250}$$

Этот оператор унитарен (перестановка вычислительного базиса), поскольку умножение на a по модулю N обратимо.

Если порядок a по модулю N равен r, то

$$\mathbf{U}_a^r = \mathbf{1}. \tag{6.251}$$

Отсюда следует, что все собственные значения $\mathbf{U}_a$ являются корнями степени r из единицы:

$$\lambda_k = e^{2\pi i k/r}, \quad k \in \{0, 1, 2, \ldots, r-1\}. \tag{6.252}$$

Соответствующими собственными состояниями являются

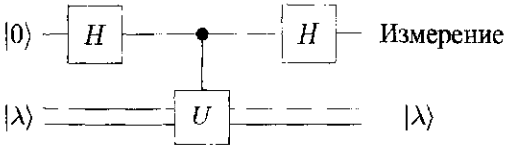
$$|\lambda_k\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} e^{-2\pi i k j/r} |a^j x_0 (\mathrm{mod}\, N)\rangle; \tag{6.253}$$

существует r взаимно ортогональных состояний, связанных с каждой орбитой длины r, генерируемой умножением на a.

Оператор $\mathbf{U}_a$ не эрмитов, но его *фаза* (эрмитов оператор, генерирующий $\mathbf{U}_a$) является наблюдаемой величиной. Предположим, что мы можем выполнить измерение, которое проецирует на базис собственных состояний $\mathbf{U}_a$ и определяет значение λ_k, с равной вероятностью выбираемое из возможных собственных значений. Следовательно, измерение определяет

значение k/r, что делает процедура Шора, и мы можем с приемлемо высокой вероятностью успеха получить множитель N. Но как измерить собственные значения унитарного оператора?

Допустим, что мы можем выполнить унитарное преобразование $\mathbf{U}$, обусловленное контрольным битом, и рассмотрим схему



Здесь $|\lambda\rangle$ обозначает собственное состояние $\mathbf{U}$, соответствующее собственному значению λ ($\mathbf{U}|\lambda\rangle = \lambda|\lambda\rangle$). Тогда действие схемы на контрольный бит имеет вид

$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + \lambda|1\rangle) \rightarrow$$
$$\rightarrow \frac{1}{2}(1+\lambda)|0\rangle + \frac{1}{2}(1-\lambda)|1\rangle). \tag{6.254}$$

Тогда результат измерения контрольного кубита имеет распределение вероятностей

$$\begin{aligned} \text{Prob}(0) &= \left|\frac{1}{2}(1+\lambda)\right|^2 = \cos^2 \pi\phi, \\ \text{Prob}(1) &= \left|\frac{1}{2}(1-\lambda)\right|^2 = \sin^2 \pi\phi, \end{aligned} \tag{6.255}$$

где $\lambda = e^{2\pi i\phi}$.

Как мы обсуждали ранее (например, в связи с проблемой Дойча), эта процедура с уверенностью различает собственные значения $\lambda = 1$ ($\phi = 0$) и $\lambda = -1$ ($\phi = 1/2$). Но также могут быть различимы и другие возможные значения λ, хотя и с меньшей статистической достоверностью. Например, предположим, что состояние, на которое действует $\mathbf{U}$, является суперпозицией его собственных состояний

$$\alpha_1|\lambda_1\rangle + \alpha_2|\lambda_2\rangle. \tag{6.256}$$

Предположим также, что мы n раз выполняем изображенную выше схему с n индивидуальными контрольными битами. Таким образом, мы готовим

состояние

$$\alpha_1|\lambda_1\rangle\left(\frac{1+\lambda_1}{2}|0\rangle+\frac{1-\lambda_1}{2}|1\rangle\right)^{\otimes n}+ \\ +\alpha_2|\lambda_2\rangle\left(\frac{1+\lambda_2}{2}|0\rangle+\frac{1-\lambda_2}{2}|1\rangle\right)^{\otimes n}. \quad (6.257)$$

Если $\lambda_1 \neq \lambda_2$, то при больших n перекрытие между двумя состояниями n контрольных битов экспоненциально мало; измеряя контрольные биты, мы можем, как минимум в отличном приближении, выполнить ортогональное проецирование на базис $\{|\lambda_1\rangle, |\lambda_2\rangle\}$.

Если мы используем достаточно контрольных битов, то в нашем распоряжении имеется достаточно большая выборка, чтобы с приемлемой статистической надежностью измерить $\mathrm{Prob}(0) = \frac{1}{2}(1+\cos 2\pi\phi)$. Выполняя контролируемое $i\mathbf{U}$, мы можем также измерить $\frac{1}{2}(1+\sin 2\pi\phi)$, что достаточно для определения ϕ по модулю целое число.

Однако в алгоритме факторизации нам необходимо измерять фазу $e^{2\pi ik/r}$ с экспоненциальной точностью, что, похоже, требует экспоненциального количества испытаний. Допустим все же, что мы можем эффективно вычислять высокие степени $\mathbf{U}$ (как в случае с $\mathbf{U}_a$), такие как

$$\mathbf{U}^{2^j}. \quad (6.258)$$

С помощью описанной выше процедуры измерения $\mathbf{U}^{2^j}$ мы определяем

$$\exp\left(2\pi i 2^j\phi\right), \quad (6.259)$$

где $e^{2\pi i\phi}$ — собственное значение оператора $\mathbf{U}$. Следовательно, измерение $\mathbf{U}^{2^j}$ с точностью до одного бита эквивалентно измерению j-го бита собственного значения $\mathbf{U}$.

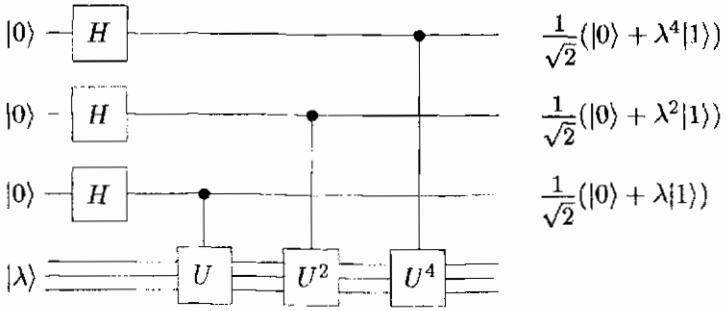
Мы можем использовать эту процедуру определения фазы для отыскания порядка и, следовательно, факторизации. Обратим уравнение (6.253), чтобы получить

$$|x_0\rangle = \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}|\lambda_k\rangle\,; \quad (6.260)$$

каждое состояние вычислительного базиса (при $x_0 \neq 0$) является равновзвешенной суперпозицией r собственных состояний $\mathbf{U}_a$.

Измеряя собственное значение, мы получаем $\lambda_k = e^{2\pi i k/r}$ с k, равновероятно выбирающимся из $\{0, 1, 2, \ldots, r-1\}$. Если $r < 2^n$, то, чтобы определить k/r, мы измеряем с точностью до $2n$ битов. В принципе мы можем выполнять эту процедуру на компьютере, который хранит меньше кубитов, чем это необходимо для вычисления QFT, потому что всякий раз мы можем приступать к определению только одного бита из k/r.

Но поучительно представить, что мы включаем QFT в процедуру определения фазы. Предположим, что схема



действует на собственное состояние $|\lambda\rangle$ унитарного преобразования $\mathbf{U}$. Условный оператор $\mathbf{U}$ готовит состояние $\frac{1}{\sqrt{2}}(|0\rangle + \lambda|1\rangle)$, условный оператор $\mathbf{U}^2$ готовит $\frac{1}{\sqrt{2}}(|0\rangle + \lambda^2|1\rangle)$, а условный оператор $\mathbf{U}^4$ — $\frac{1}{\sqrt{2}}|0\rangle + \lambda^4|1\rangle)$ и так далее. Мы могли бы выполнить преобразование Адамара и измерить каждый из этих кубитов, чтобы получить распределение вероятностей, управляющее j-м битом ϕ, где $\lambda = e^{2\pi i\phi}$. Но целесообразнее заметить, что приготовленное схемой состояние равно

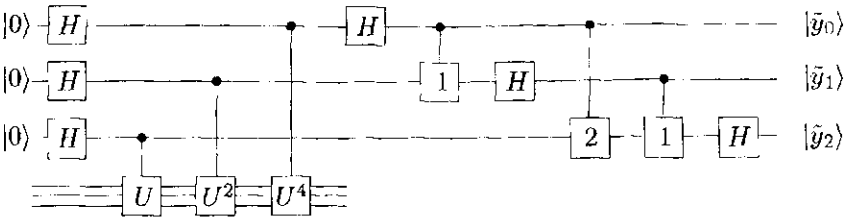
$$\frac{1}{\sqrt{2^m}} \sum_{y=0}^{2^m-1} e^{2\pi i\phi y}|y\rangle. \tag{6.261}$$

Лучший способ узнать значение ϕ — выполнить перед измерением $\mathrm{QFT}^{(m)}$, а не преобразование Адамара $\mathbf{H}^{(m)}$.

Рассмотрим для ясности случай $m = 3$. Схема, которая готовит состояние

$$|x_0\rangle = \frac{1}{\sqrt{8}} \sum_{k=0}^{7} e^{2\pi i\phi y}|y\rangle, \tag{6.262}$$

а затем выполняет его Фурье-анализ, имеет вид



Эта схема почти полностью выполняет описанную выше стратегию определения фазы, но со значительной модификацией. Прежде чем мы выполняем завершающее преобразование Адамара и измеряем $\tilde{y}_1$ и $\tilde{y}_2$, выполняются некоторые условные фазовые повороты. Это те фазовые повороты, которые отличают $\mathrm{QFT}^{(3)}$ от преобразования Адамара $\mathbf{H}^{(3)}$ и резко повышают эффективность, с которой мы можем извлечь значение ϕ.

Мы можем лучше понять, что делают условные повороты, если предположим, что $\phi = k/8$ при $k \in \{0, 1, 2, \ldots, 7\}$; в этом случае нам известно, что преобразование Фурье будет с вероятностью единица генерировать на выходе $\tilde{y} = k$. Мы можем представить k в виде двоичного разложения

$$k = k_2 k_1 k_0 = k_2 \cdot 4 + k_1 \cdot 2 + k_0. \tag{6.263}$$

Фактически схема для самого младшего значащего бита $\tilde{y}_0$ преобразования Фурье является в точности измерительной схемой Китаева, применяемой к унитарному преобразованию $\mathbf{U}^4$, собственные значения которого равны

$$\left(e^{2\pi i\phi}\right)^4 = e^{i\pi k} = e^{i\pi k_0} = \pm 1. \tag{6.264}$$

Измерительная цепь идеально различает собственные значения ± 1, так что $\tilde{y}_0 = k_0$.

Схема для следующего бита $\tilde{y}_1$ почти такая же, как и измерительная схема для $\mathbf{U}^2$ с собственным значением

$$\left(e^{2\pi i\phi}\right)^2 = e^{i\pi k/2} = e^{i\pi(k_1 . k_0)}, \tag{6.265}$$

за исключением того, что добавлен условный фазовый поворот, который умножает фазу на $\exp[-i\pi(.k_0)]$, давая в результате $e^{i\pi k_1}$. Вновь, применяя вслед за измерением преобразование Адамара, мы с определенностью получаем результат $\tilde{y}_1 = k_1$. Аналогично схема для $\tilde{y}_2$ измеряет собственное значение

$$e^{2\pi i\phi} = e^{i\pi k/4} = e^{i\pi(k_2 . k_1 k_0)}, \tag{6.266}$$

за исключением того, что условный поворот удаляет $e^{i\pi(.k_1 k_0)}$, так что результатом с определенностью является $\tilde{y}_2 = k_2$.

Таким образом, QFT наилучшим образом осуществляет программу определения фазы. Сначала мы измеряем значащие биты младших разрядов фазы ϕ и используем приобретенную в измерениях информацию, чтобы улучшить достоверность определения значащих битов следующих разрядов. Имея в виду эту интерпретацию, вы сможете просто запомнить схему для $\mathrm{QFT}^{(n)}$!

6.12. Резюме

Классические схемы. Сложность задачи можно характеризовать размером однородного семейства логических схем, решающих эту задачу. Задача трудная, если размер схемы является суперполиномиальной функцией от размера входа. Один классический универсальный компьютер может эффективно моделировать другой, так что классификация сложности аппаратно-независима. Трехбитовый вентиль Тоффоли является универсальным для классических обратимых вычислений. Обратимый компьютер может моделировать необратимый без значительного замедления и неприемлемых затрат памяти.

Квантовые схемы. Хотя это не доказано, но выглядит правдоподобным, что квантовые схемы полиномиального размера не могут моделироваться классическими вероятностными схемами полиномиального размера ($BQP \neq BPP$); однако для этого достаточно полиномиального пространства ($BQP \subseteq PSPACE$). Квантовая схема с шумом может с приемлемой точностью моделировать идеальную квантовую схему размера T, если каждый квантовый вентиль имеет точность порядка $1/T$. Один универсальный квантовый компьютер может эффективно моделировать другой, так что класс сложности BQP аппаратно независим. Типичный двухкубитовый квантовый вентиль, если он может действовать на любую пару кубитов в приборе, адекватен для универальных квантовых вычислений. Также адекватны вентиль контролируемое NOT плюс типичный однокубитовй вентиль.

Быстрый квантовый поиск. Исчерпывающий поиск маркированного элемента в неструктурированной базе данных из N элементов может быть выполнен с помощью квантового компьютера за время порядка $\sqrt{N}$, но не быстрее. Квадратичное квантовое ускорение также может быть достигнуто для некоторых проблем структурированного поиска, но некоторые проблемы оракула не допускают существенного квантового ускорения. Два участника, каждый из которых имеет в распоряжении таблицу из N

записей, могут локализовать совпадающие строки в их таблицах, обменявшись $O(\sqrt{N}\log N)$ кубитами, явно вступая в конфликт с духом (но не буквой) границы Холево.

Отыскание периода. Используя квантовый параллелизм, можно вычислить квантовое преобразование Фурье в N-мерном пространстве за время порядка $(\log N)^2$ (по сравнению со временем $N\log N$ для классического быстрого преобразования Фурье); если нам нужно сразу после этого выполнять измерение, то для вычисления QFT достаточно однокубитовых вентилей. Таким образом, квантовые компьютеры могут эффективно решать некоторые задачи с периодической структурой, такие как факторизация и задача о дискретном логарифме.

6.13. Упражнения

6.1. Линейное моделирование вентиля Тоффоли. На лекции мы построили n-битовый вентиль Тоффоли $\theta^{(n)}$ из трехбитовых вентилей Тоффоли $\theta^{(3)}$. Схема требовала только одного бита вспомогательного пространства, но количество вентилей было экспоненциально велико по n. Используя более широкое вспомогательное пространство, можно существенно сократить количество вентилей.

a) Найдите вычисляющее $\theta^{(n)}$ семейство схем из $2n-5$ вентилей $\theta^{(3)}$. (Здесь используется $n-3$ вспомогательных битов, которые в конце вычисления возвращаются к своим исходным, равным нулю, значениям.)

b) Найдите вычисляющее $\theta^{(n)}$ семейство схем из $4n-12$ вентилей $\theta^{(3)}$, которое работает независимо от начальных значений вспомогательных битов. (Вновь $n-3$ вспомогательных битов в конце вычисления возвращаются к своим начальным, не обязательно равным нулю, значениям.)

6.2. Универсальный набор квантовых вентилей. Цель этого упражнения — завершить демонстрацию того, что контролируемое NOT и произвольные однокубитовые вентили образуют универсальный набор.

a) Пусть $\mathbf{U}$ — произвольная унитарная 2×2-матрица с единичным определителем. Найдите унитарные преобразования $\mathbf{A}$, $\mathbf{B}$ и $\mathbf{C}$ такие, что

$$\mathbf{ABC} = \mathbf{1}, \tag{6.267}$$

$$\mathbf{A}\sigma_x\mathbf{B}\sigma_x\mathbf{C} = \mathbf{U}. \tag{6.268}$$

[**Указание.** Из конструкции углов Эйлера мы знаем, что

$$\mathbf{U} = \mathbf{R}_z(\psi)\mathbf{R}_y(\theta)\mathbf{R}_z(\phi), \tag{6.269}$$

где, например, $\mathbf{R}_z(\phi)$ обозначает поворот вокруг оси z на угол ϕ. Нам также известно, что, например,

$$\boldsymbol{\sigma}_x\mathbf{R}_z(\phi)\boldsymbol{\sigma}_x = \mathbf{R}_z(-\phi).] \tag{6.270}$$

b) Рассмотрите двухкубитовый *вентиль контролируемой фазы*: он применяет $\mathbf{U} = e^{i\alpha}\mathbf{1}$ ко второму кубиту, если первый кубит имеет значение $|1\rangle$, и действует тривиально в противном случае. Покажите, что фактически он является однокубитовым вентилем.

c) Используя вентили контролируемое NOT и однокубитовые вентили, изобразите схему, реализующую контролируемое U, где U — произвольное унитарное 2×2-преобразование.

6.3. Точность. Цель этого упражнения — установить связь между точностью квантового состояния и точностью соответствующего распределения вероятностей.

a) Пусть $\|\mathbf{A}\|$ обозначает норму оператора $\mathbf{A}$, а

$$\|\mathbf{A}\|_{\mathrm{tr}} = \mathrm{tr}\left[(\mathbf{A}\mathbf{A}^\dagger)^{1/2}\right] \tag{6.271}$$

обозначает *следовую норму*. Покажите, что

$$\|\mathbf{AB}\|_{\mathrm{tr}} \leqslant \|\mathbf{B}\| \cdot \|\mathbf{A}\|_{\mathrm{tr}} \quad \text{и} \quad |\mathrm{tr}\,\mathbf{A}| \leqslant \|\mathbf{A}\|_{\mathrm{tr}}. \tag{6.272}$$

b) Предположим, что $\boldsymbol{\rho}$ и $\tilde{\boldsymbol{\rho}}$ — две матрицы плотности, а $\{|a\rangle\}$ — полный ортонормированный базис. Так что

$$\begin{aligned} P_a &= \langle a|\boldsymbol{\rho}|a\rangle, \\ \tilde{P}_a &= \langle a|\tilde{\boldsymbol{\rho}}|a\rangle \end{aligned} \tag{6.273}$$

— соответствующие распределения вероятностей. Используя (a), покажите, что

$$\sum_a |P_a - \tilde{P}_a| \leqslant \|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}}. \tag{6.274}$$

c) Предположим, что $\boldsymbol{\rho} = |\psi\rangle\langle\psi|$ и $\tilde{\boldsymbol{\rho}} = |\tilde{\psi}\rangle\langle\tilde{\psi}|$ — чистые состояния. Используя (b), покажите, что

$$\sum_a |P_a - \tilde{P}_a| \leqslant 2\||\psi\rangle - |\tilde{\psi}\rangle\|. \tag{6.275}$$

6.4. Поиск в базе данных с непрерывным временем. Квантовая система с n-кубитовым гильбертовым пространством имеет гамильтониан

$$\mathbf{H}_\omega = E|\omega\rangle\langle\omega|, \tag{6.276}$$

где $|\omega\rangle$ — неизвестное состояние вычислительного базиса. Вам нужно найти значение ω с помощью следующей процедуры. Включите не зависящее от времени возмущение $\mathbf{H}'$, так что полным гамильтонианом будет

$$\mathbf{H} = \mathbf{H}_\omega + \mathbf{H}'. \tag{6.277}$$

Приготовьте начальное состояние $|\psi_0\rangle$ и позвольте ему эволюционировать в течение времени T под управлением $\mathbf{H}$. Затем измерьте состояние. По результату измерения вы должны сделать вывод относительно ω.

a) Допустим, что в качестве начального состояния выбрано

$$|s\rangle = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle, \tag{6.278}$$

а в качестве возмущения

$$\mathbf{H}' = E|s\rangle\langle s|. \tag{6.279}$$

Решите зависящее от времени уравнение Шредингера

$$i\frac{d}{dt}|\psi\rangle = \mathbf{H}|\psi\rangle, \tag{6.280}$$

чтобы найти состояние в момент времени T. Каким следует выбрать T, чтобы оптимизировать вероятность успешного определения ω?

b) Теперь предположим, что $|\psi_0\rangle$ и $\mathbf{H}'$ можно выбрать какими угодно, но мы требуем, чтобы спустя время T состоянием системы было $|\omega\rangle$, так чтобы измерение определяло ω с единичной вероятностью успеха. Выведите нижнюю границу, которой должно удовлетворять T, и сравните с вашим результатом в (а). [**Указание.** Как и в нашем анализе на лекции, сравните эволюцию, управляемую $\mathbf{H}$, с эволюцией, управляемой $\mathbf{H}'$ (случай «пустого оракула»), и используйте уравнение Шредингера, чтобы найти, как быстро состояние, эволюционирующее в соответствии с $\mathbf{H}$, отклоняется от состояния, эволюционирующего в соответствии с $\mathbf{H}'$.]

Часть II

Решение упражнений[1]

[1] Решения упражнений выполнены Эндрю Лэндалом и Джимом Харрингтоном (упр. 4.1, 4.2, 4.6).

Решения упражнений к главе 2

2.1. Точность воспроизведения вероятностной гипотезы

Операторы (матрицы) плотности чистого состояния двухуровневой системы находятся во взаимно однозначном соответствии с точками на поверхности сферы Блоха. Благодаря этому соответствию, мы можем выбрать меру Хаара таких матриц плотности равной обычной евклидовой мере на S^2.[1]

$$d\mu = \frac{\sin\theta d\theta d\varphi.}{4\pi}$$

Если мы извлекаем $|\psi\rangle$ из однородного ансамбля чистых состояний и предполагаем, что извлеченным вслед за ним из того же ансамбля состоянием является $|\phi\rangle$, то усредненная ожидаемая точность воспроизведения нашей гипотезы дается (классическим) математическим ожиданием по обоим выборам:

$$\begin{aligned}
\langle F\rangle &= E_{|\psi\rangle}E_{|\phi\rangle}\left[|\langle\phi|\psi\rangle|^2\right] = E_{|\psi\rangle}E_{|\phi\rangle}\left[\langle\phi|\psi\rangle\langle\psi|\phi\rangle\right] = \\
&= E_{|\psi\rangle}E_{|\phi\rangle}\left[\operatorname{tr}\left(|\phi\rangle\langle\phi|\psi\rangle\langle\psi|\right)\right] = \operatorname{tr}\left(E_{|\psi\rangle}E_{|\phi\rangle}\left[|\phi\rangle\langle\phi|\psi\rangle\langle\psi|\right]\right) = \\
&= \operatorname{tr}\left(E_{|\psi\rangle}\left[|\psi\rangle\langle\psi|\right]E_{|\phi\rangle}\left[|\phi\rangle\langle\phi|\right]\right) = \\
&= \operatorname{tr}\left[\left(\int\frac{1}{2}\left(\mathbf{1}+\hat{n}_{|\psi\rangle}\cdot\vec{\boldsymbol{\sigma}}\right)\frac{\sin\theta d\theta d\varphi}{4\pi}\right)\times\right. \\
&\qquad\left.\times\left(\int\frac{1}{2}\left(\mathbf{1}+\hat{n}_{|\phi\rangle}\cdot\vec{\boldsymbol{\sigma}}\right)\frac{\sin\theta d\theta d\varphi}{4\pi}\right)\right] = \\
&= \operatorname{tr}\left[\left(\frac{1}{2}\mathbf{1}\right)\left(\frac{1}{2}\mathbf{1}\right)\right] = \frac{1}{4}\operatorname{tr}\mathbf{1} = \frac{1}{2}.
\end{aligned}$$

То, что усредненная точность воспроизведения равна $\frac{1}{2}$, интуитивно понятно, но мы должны честно выполнить вычисления, чтобы подтвердить правильность нашей интуиции. Особенно работая в квантовом мире!

Прежде чем приступать к решению, я хотел бы объяснить, почему приготовленная измерением матрица плотности может быть записана в виде,

[1] В общем случае отыскание меры Хаара для матриц плотности может оказаться трудной задачей.

данном в условии задачи. Вспомним, что если начальным состоянием квантовой системы является чистое состояние $|\psi\rangle$, то, согласно третьей аксиоме (см. раздел 2.1), измерение наблюдаемой $\mathbf{A}$ с вероятностью $p_n = \langle\psi|\mathbf{P}_n|\psi\rangle$ выбирает проектор $\mathbf{P}_n$ на одно из собственных состояний $\mathbf{A}$ и переводит систему в нормированное чистое состояние

$$|\psi_n\rangle = \frac{\mathbf{P}_n|\psi\rangle}{\langle\psi|\mathbf{P}_n|\psi\rangle^{1/2}}.$$

Это наводит на мысль, что матрица плотности чистого состояния должна трансформироваться в *ансамбль* всех возможных результатов измерения:

$$|\psi\rangle\langle\psi| \to \sum_n p_n|\psi_n\rangle\langle\psi_n|.$$

Но $|\psi_n\rangle\langle\psi_n|$ является именно проектором $\mathbf{P}_n$, так что приготавливаемая измерением матрица плотности с тем же основанием может быть записана в виде

$$|\psi\rangle\langle\psi| \to \sum_n \langle\psi|\mathbf{P}_n|\psi\rangle \boldsymbol{P}_n. \qquad \square$$

Однако следует предостеречь, что такая эволюция верна, если только начальным состоянием системы является чистое состояние. В общем случае, как было показано на лекциях, эволюция матрицы плотности $\boldsymbol{\rho}$ под влиянием измерения (фон Неймана) имеет вид $\boldsymbol{\rho} \to \sum_n \mathbf{P}_n \operatorname{tr}(\boldsymbol{\rho}\mathbf{P}_n)$ (см. раздел 3.1.1).

С данной выше матрицей плотности вычисление точности воспроизведения сравнительно просто:

$$\begin{aligned} F &= \langle\psi|\boldsymbol{\rho}|\psi\rangle = \langle\psi|\big(\mathbf{P}_\uparrow\langle\psi|\mathbf{P}_\uparrow|\psi\rangle + \mathbf{P}_\downarrow\langle\psi|\mathbf{P}_\downarrow|\psi\rangle\big)|\psi\rangle = \\ &= \langle\psi|\mathbf{P}_\uparrow|\psi\rangle\langle\psi|\mathbf{P}_\uparrow|\psi\rangle + \langle\psi|\mathbf{P}_\downarrow|\psi\rangle\langle\psi|\mathbf{P}_\downarrow|\psi\rangle = \\ &= \langle\psi|\mathbf{P}_\uparrow|\psi\rangle^2 + \langle\psi|\mathbf{P}_\downarrow|\psi\rangle^2 = p_\uparrow^2 + p_\downarrow^2 = \\ &= p^2 + (1-p)^2 \qquad\qquad (p \equiv p_\uparrow) \\ &= 2p^2 - 2p + 1. \end{aligned}$$

Чтобы найти *усредненную* точность воспроизведения, необходимо вычислить (классическое) математическое ожидание по всем возможным реализациям состояния $|\psi\rangle$. Поскольку состояния $|\psi\rangle$ распределены *однород-*

но, усредненную точность воспроизведения $\langle F\rangle = E_{|\psi\rangle}[F]$ можно найти, *полагая* однородным распределение $p \in [0,1]$. (Длинный путь состоит в замене $|\psi\rangle$ спинорами или проекторами и усреднении по всем углам.) Это дает следующее значение усредненной точности воспроизведения:

$$\langle F\rangle = \int_0^1 (2p^2 - 2p + 1)dp = \frac{2}{3} - \frac{1}{2} + \frac{1}{6}.$$

Таким образом, выполнение измерения увеличивает точность воспроизведения на $\frac{1}{6}$. Эвристически это можно запомнить, заметив, что с вероятностью $\frac{1}{3}$ состояние $|\psi\rangle$ ориентировано вдоль оси измерения, а соответствующий результат измерения дает $|\psi\rangle$ вдоль этой оси имеет вероятность $\frac{1}{2}$. Вероятность того, что гипотеза верна, равна $\frac{1}{3}\cdot\frac{1}{2}$.

Следует заметить, что, хотя выражения для точности воспроизведения в задачах 2.1 и 2.2 выглядят различными, на самом деле оба они являются частными случаями общего выражения

$$F = \mathrm{tr}\,\rho_1\rho_2.$$

В задаче 2.1 обе ρ_1 и ρ_2 были чистыми состояниями, а в задаче 2.2 чистое только ρ_1. В общем случае F описывает, насколько подобны два квантовых состояния. Непосредственно видно, что точность воспроизведения является не самой лучшей метрикой, так как, например, $\mathrm{tr}(\rho)^2 = 1$ только в частном случае, когда ρ является проектором. Позже в этом курсе мы найдем более хорошие меры точности воспроизведения.

2.3. Разложение Шмидта

2.3.1. Частичные следы. Решение этой части главным образом получается путем чисто механического применения определений. Начальным состоянием системы является:

$$|\Phi\rangle = \frac{1}{\sqrt{2}}|\uparrow\rangle_A\left(\frac{1}{2}|\uparrow\rangle_B + \frac{\sqrt{3}}{2}|\downarrow\rangle_B\right) + \frac{1}{\sqrt{2}}|\downarrow\rangle_A\left(\frac{\sqrt{3}}{2}|\uparrow\rangle_B + \frac{1}{2}|\downarrow\rangle_B\right) =$$

$$= \frac{1}{\sqrt{8}}\Big(|\uparrow\uparrow\rangle + \sqrt{3}|\uparrow\downarrow\rangle + \sqrt{3}|\downarrow\uparrow\rangle + |\downarrow\downarrow\rangle\Big),$$

что, будучи записанным как оператор плотности, представляет собой:

$$|\Phi\rangle\langle\Phi| = \frac{1}{8}\Big(|\uparrow\uparrow\rangle + \sqrt{3}|\uparrow\downarrow\rangle + \sqrt{3}|\downarrow\uparrow\rangle + |\downarrow\downarrow\rangle\Big) \times$$
$$\times\Big(\langle\uparrow\uparrow| + \sqrt{3}\langle\uparrow\downarrow| + \sqrt{3}\langle\downarrow\uparrow| + \langle\downarrow\downarrow|\Big) =$$
$$= \Big(|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\Big)\frac{1}{8}\begin{pmatrix} 1 & \sqrt{3} & \sqrt{3} & 1 \\ \sqrt{3} & 3 & 3 & \sqrt{3} \\ \sqrt{3} & 3 & 3 & \sqrt{3} \\ 1 & \sqrt{3} & \sqrt{3} & 1 \end{pmatrix}\begin{pmatrix} \langle\uparrow\uparrow| \\ \langle\uparrow\downarrow| \\ \langle\downarrow\uparrow| \\ \langle\downarrow\downarrow| \end{pmatrix}.$$

Частичный след по системе B дает:

$$\begin{aligned}\rho_A &= \mathrm{tr}_B\,|\Phi\rangle\langle\Phi| = \\ &= \langle\uparrow_B|\Phi\rangle\langle\Phi|\uparrow_B\rangle + \langle\downarrow_B|\Phi\rangle\langle\Phi|\downarrow_B\rangle = \\ &= \frac{1}{8}\begin{bmatrix} |\uparrow\rangle\langle\uparrow| + \sqrt{3}|\uparrow\rangle\langle\downarrow| + \sqrt{3}|\downarrow\rangle\langle\uparrow| + 3|\downarrow\rangle\langle\downarrow| + \\ +3|\uparrow\rangle\langle\uparrow| + \sqrt{3}|\uparrow\rangle\langle\downarrow| + \sqrt{3}|\downarrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| \end{bmatrix} = \\ &= \frac{1}{8}\Big(4|\uparrow\rangle\langle\uparrow| + 2\sqrt{3}|\uparrow\rangle\langle\downarrow| + 2\sqrt{3}|\downarrow\rangle\langle\uparrow| + 4|\downarrow\rangle\langle\downarrow|\Big) = \\ &= (|\uparrow\rangle, |\downarrow\rangle)\begin{pmatrix} 1/2 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/2 \end{pmatrix}\begin{pmatrix} \langle\uparrow| \\ \langle\downarrow| \end{pmatrix}.\end{aligned}$$

Поскольку состояние $|\Phi\rangle$ симметрично относительно обмена системами A и B, то оказывается, что частичный след по системе A дает тот же самый вид приведенной матрицы плотности ρ_B:

$$\begin{aligned}\rho_B &= \mathrm{tr}_A\,|\Phi\rangle\langle\Phi| = \\ &= \langle\uparrow_A|\Phi\rangle\langle\Phi|\uparrow_A\rangle + \langle\downarrow_A|\Phi\rangle\langle\Phi|\downarrow_A\rangle = \\ &= \frac{1}{8}\begin{bmatrix} |\uparrow\rangle\langle\uparrow| + \sqrt{3}|\uparrow\rangle\langle\downarrow| + \sqrt{3}|\downarrow\rangle\langle\uparrow| + 3|\downarrow\rangle\langle\downarrow| + \\ +3|\uparrow\rangle\langle\uparrow| + \sqrt{3}|\uparrow\rangle\langle\downarrow| + \sqrt{3}|\downarrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| \end{bmatrix} = \\ &= \frac{1}{8}\Big(4|\uparrow\rangle\langle\uparrow| + 2\sqrt{3}|\uparrow\rangle\langle\downarrow| + 2\sqrt{3}|\downarrow\rangle\langle\uparrow| + 4|\downarrow\rangle\langle\downarrow|\Big) = \\ &= (|\uparrow\rangle, |\downarrow\rangle)\begin{pmatrix} 1/2 & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/2 \end{pmatrix}\begin{pmatrix} \langle\uparrow| \\ \langle\downarrow| \end{pmatrix}.\end{aligned}$$

2.3.2 Разложение Шмидта. Решение этой части задачи тоже получается в результате простых манипуляций определениями, но с дополнительным преобразованием, с помощью которого мы предварительно поворачиваем базис системы A так, чтобы диагонализовать приведенную матрицу ρ_A.

Чтобы выполнить это, найдем сначала собственные состояния, диагонализующие ρ_A:

$$\begin{vmatrix} 1/2-\lambda & \sqrt{3}/4 \\ \sqrt{3}/4 & 1/2-\lambda \end{vmatrix} = 0,$$

$$1/4 - \lambda + \lambda^2 - 3/16 = 0,$$

$$\lambda^2 - \lambda + 1/16 = 0,$$

$$\lambda_\pm = 1/2 \pm \sqrt{3}/4,$$

$$|\psi^\pm\rangle_A = \frac{1}{\sqrt{2}}(|\uparrow\rangle_A \pm |\downarrow\rangle_A).$$

Чтобы выполнить (локальную) замену базиса, которая реализует эти собственные векторы в качестве базисных, используем обычную формулу перехода от одного представления к другому, основанную на очевидной тождественной вставке:

$$|\Phi\rangle = \sum_{i=\pm} |\psi^i\rangle\langle\psi^i|\Phi\rangle.$$

Коэффициентами этого разложения фактически являются состояния системы B, так как внутреннее произведение здесь вычисляется только по состояниям системы A. Действительно,

$$\begin{aligned}
\langle\psi^+|\Phi\rangle &= \frac{1}{2}({}_A\langle\uparrow| + {}_A\langle\downarrow|) \times \\
&\quad \times \left[|\uparrow\rangle_A \left(\frac{1}{2}|\uparrow\rangle_B + \frac{\sqrt{3}}{2}|\downarrow\rangle_B \right) + |\downarrow\rangle_A \left(\frac{\sqrt{3}}{2}|\uparrow\rangle_B + \frac{1}{2}|\downarrow\rangle_B \right) \right] = \\
&= \frac{1}{4}\left(|\uparrow\rangle_B + \sqrt{3}|\downarrow\rangle_B + \sqrt{3}|\uparrow\rangle_B + |\downarrow\rangle_B \right) = \\
&= \frac{1+\sqrt{3}}{4}\left(|\uparrow\rangle_B + |\downarrow\rangle_B \right) \equiv |\tilde{\varphi}_1\rangle_B,
\end{aligned}$$

$$\begin{aligned}
\langle\psi^-|\Phi\rangle &= \frac{1}{2}({}_A\langle\uparrow| - {}_A\langle\downarrow|) \times \\
&\quad \times \left[|\uparrow\rangle_A \left(\frac{1}{2}|\uparrow\rangle_B + \frac{\sqrt{3}}{2}|\downarrow\rangle_B \right) + |\downarrow\rangle_A \left(\frac{\sqrt{3}}{2}|\uparrow\rangle_B + \frac{1}{2}|\downarrow\rangle_B \right) \right] = \\
&= \frac{1}{4}\left(|\uparrow\rangle_B + \sqrt{3}|\downarrow\rangle_B - \sqrt{3}|\uparrow\rangle_B - |\downarrow\rangle_B \right) = \\
&= \frac{1-\sqrt{3}}{4}\left(|\uparrow\rangle_B - |\downarrow\rangle_B \right) \equiv |\tilde{\varphi}_2\rangle_B.
\end{aligned}$$

Мы почти у цели. Все, что нам осталось сделать — это нормировать полученные состояния:

$$\langle\tilde{\varphi}_1|\tilde{\varphi}_1\rangle = \frac{4+2\sqrt{3}}{16}({}_B\langle\uparrow| + {}_B\langle\downarrow|)(|\uparrow\rangle_B + |\downarrow\rangle_B) =$$

$$= \frac{1}{2}\left(1+\frac{\sqrt{3}}{2}\right) \equiv p_1,$$

$$\langle\tilde{\varphi}_2|\tilde{\varphi}_2\rangle = \frac{4-2\sqrt{3}}{16}({}_B\langle\uparrow| - {}_B\langle\downarrow|)(|\uparrow\rangle_B - |\downarrow\rangle_B) =$$

$$= \frac{1}{2}\left(1-\frac{\sqrt{3}}{2}\right) \equiv p_2.$$

Используя нормированные состояния $|\varphi_i\rangle_B \equiv \frac{1}{\sqrt{p_i}}|\tilde{\varphi}_i\rangle_B$, мы можем записать разложение Шмидта этого чистого состояния по ортонормированным состояниям систем A и B:

$$|\Phi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle_A|\varphi_i\rangle_B =$$

$$|\Phi\rangle = \sqrt{1+\frac{\sqrt{3}}{2}}\left[\frac{1}{\sqrt{2}}(|\uparrow\rangle_A+|\downarrow\rangle_A)\right]\left[\frac{1}{\sqrt{1+\frac{\sqrt{3}}{2}}}\,\frac{1+\sqrt{3}}{4}(|\uparrow\rangle_B+|\downarrow\rangle_B)\right] +$$

$$+\sqrt{1-\frac{\sqrt{3}}{2}}\left[\frac{1}{\sqrt{2}}(|\uparrow\rangle_A-|\downarrow\rangle_A)\right]\left[\frac{1}{\sqrt{1-\frac{\sqrt{3}}{2}}}\,\frac{1-\sqrt{3}}{4}(|\uparrow\rangle_B-|\downarrow\rangle_B)\right] =$$

$$= \left(\frac{1+\sqrt{3}}{2\sqrt{2}}\right)\left[\frac{1}{\sqrt{2}}(|\uparrow\rangle_A+|\downarrow\rangle_A)\right]\left[\frac{1}{\sqrt{2}}(|\uparrow\rangle_B+|\downarrow\rangle_B)\right] +$$

$$+\left(\frac{1-\sqrt{3}}{2\sqrt{2}}\right)\left[\frac{1}{\sqrt{2}}(|\uparrow\rangle_A-|\downarrow\rangle_A)\right]\left[\frac{1}{\sqrt{2}}(|\uparrow\rangle_B-|\downarrow\rangle_B)\right].$$

2.4. Трехкубитовое чистое состояние

Нет. Прежде чем объяснять, почему, я хотел бы отметить, что на самом деле неверно в этой задаче. Разложение Шмидта для трехкомпонентной системы должно выглядеть следующим образом:

$$\sum_i \sqrt{p_i}|i\rangle_A \otimes |i\rangle_B \otimes |i\rangle_C.$$

Надеюсь, это ни у кого не вызывает недоумения, так как множители $\sqrt{p_i}$, очевидно, должны присутствовать, чтобы нормировать состояние. В общем случае разложение Шмидта n-компонентной системы имело бы вид

$$\sum_i \sqrt{p_i} \bigotimes_j |i\rangle_j,$$

где каждый базис Шмидта $\{|i\rangle_j\}$ является ортонормированным в j-ой системе.

Теперь обратимся к «объяснению», которое я употребляю как эвфемизм «доказательства»[1].

Пусть $|\psi\rangle_{ABC}$ – чистое состояние трехкомпонентной системы; предположим, что $|\psi\rangle_{ABC}$ имеет разложение Шмидта. Коль скоро это так, то вычисление парциального следа по любым двум подсистемам даст диагональную в базисе Шмидта приведенную матрицу плотности оставшейся подсистемы. Более того, вычисляемые этом базисе приведенные матрицы плотности подсистем должны иметь *один и тот же* спектр значений p_i[2]. Любые локальные (действующие внутри одной подсистемы) унитарные преобразования базисов сохраняют собственные значения приведенных матриц плотности. Следовательно, спектры (ненулевые) приведенных матриц плотности всех этих подсистем должны быть идентичны независимо от того, в каком базисе они выражаются.

Это требование строгого «совпадения спектров» несправедливо для произвольных $|\psi\rangle_{ABC}$, и примеров этому множество. Я думаю, что простейшим контрпримером является:

$$|\psi\rangle_{ABC} = |0\rangle_A \left(\frac{1}{\sqrt{2}} \left(|00\rangle_{BC} + |11\rangle_{BC} \right) \right),$$

[1] Эвфемизм от греческого *euphemia* – воздержание от резких слов, смягченное выражение (перев.)

[2] Точнее, приведенные матрицы плотности должны иметь совпадающие спектры *ненулевых* собственных значений p_i (см. раздел 2.4). – *Прим. ред.*

$$\rho_{ABC} = |\psi\rangle_{ABC\,ABC}\langle\psi| =$$
$$= \frac{1}{2}\big(|000\rangle\langle 000| + |000\rangle\langle 011| + |011\rangle\langle 000| + |011\rangle\langle 011|\big),$$

$$\rho_A = \mathrm{tr}_B\,\mathrm{tr}_C\,\rho_{ABC}$$
$$= |0\rangle_{A\,A}\langle 0|$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\rho_B = \mathrm{tr}_A\,\mathrm{tr}_C\,\rho_{ABC}$$
$$= \frac{1}{2}\big(|0\rangle_{B\,B}\langle 0| + |1\rangle_{B\,B}\langle 1|\big)$$
$$= \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

$$\{1,0\} \neq \left\{\frac{1}{2}, \frac{1}{2}\right\}. \qquad \square$$

2.5. Квантовые корреляции в смешанном состоянии

Как мы видели в задаче 2.2, вероятность измерения собственного значения $\mathbf{P}_n$ равна $p_n = \mathrm{tr}\,\mathbf{P}_n\rho$. Взаимно однозначное соответствие между проекторами на кубиты и точками на поверхности сферы Блоха говорит о том, что вероятность результатов двух последовательных измерений — «спин вверх» вдоль оси $\hat{n}$ у первого кубита и «спин вверх» вдоль оси $\hat{m}$ у второго кубита — равна

$$p = \mathrm{tr}_B\left\{\mathbf{P}_{\hat{m}}\,\mathrm{tr}_A\left[\big(\mathbf{P}_{\hat{n}}\otimes\mathbf{1}_B\big)\rho\right]\right\},$$
$$p = \mathrm{tr}_B\left\{\frac{1}{2}\big(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B\big)\,\mathrm{tr}_A\left[\frac{1}{2}\big(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A\big)\otimes\mathbf{1}_B\big)\rho\right]\right\}.$$

По отношению к следу по системе A проектор $\mathbf{P}_{\hat{m}}$ является мультипликативной константой, поскольку его действие на систему A тривиально. В силу линейности следа такую константу можно внести под его знак:

$$p = \mathrm{tr}_B\,\mathrm{tr}_A\left[\left(\mathbf{1}_A\otimes\frac{1}{2}\big(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B\big)\right)\left(\frac{1}{2}\big(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A\big)\otimes\mathbf{1}_B\right)\rho\right] =$$
$$= \mathrm{tr}_B\,\mathrm{tr}_A\left[\left(\frac{1}{2}\big(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A\big)\otimes\frac{1}{2}\big(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B\big)\right)\rho\right].$$

С помощью данной в задаче $\boldsymbol{\rho}$, эти следы можно вычислить явно, учитывая линейность следа и равенство нулю следов матриц Паули:

$$= \mathrm{tr}_B\,\mathrm{tr}_A\left[\left(\frac{1}{2}(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\right)\otimes\left(\frac{1}{2}(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)\right)\left(\frac{1}{8}\mathbf{1}_{AB}+\frac{1}{2}|\psi^-\rangle\langle\psi^-|\right)\right] =$$

$$= \frac{1}{32}\,\mathrm{tr}_B\,\mathrm{tr}_A\left[(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)\right] +$$

$$+\frac{1}{8}\,\mathrm{tr}_B\,\mathrm{tr}_A\left[\Big((\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)\Big)|\psi^-\rangle\langle\psi^-|\right] =$$

$$= \frac{1}{32}\,\mathrm{tr}_B\,\mathrm{tr}_A\left[\mathbf{1}\otimes\mathbf{1}\right] + \frac{1}{8}\langle\psi^-|(\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes(\mathbf{1}+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)|\psi^-\rangle =$$

$$= \frac{1}{8} + \frac{1}{8}\langle\psi^-|\mathbf{1}+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A+\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B+\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A\otimes\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B|\psi^-\rangle =$$

$$= \frac{1}{4}+\frac{1}{8}\Big[\hat{n}\cdot\langle\psi^-|\vec{\boldsymbol{\sigma}}_A|\psi^-\rangle+\hat{m}\cdot\langle\psi^-|\vec{\boldsymbol{\sigma}}_B|\psi^-\rangle+\langle\psi^-|\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A\otimes\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B|\psi^-\rangle\Big].$$

В синглетном состоянии $|\psi^-\rangle$ математические ожидания $\boldsymbol{\sigma}_A$ и $\boldsymbol{\sigma}_B$ равны нулю, в чем можно убедиться или с помощью явных вычислений или заметив, что *синглет является скалярным* (спин-0) состоянием. Остается вычислить только одно слагаемое, самое правое из приведенных выше. Для этого имеется несколько способов. Возможно, проще всего показать, что благодаря спин-0 симметрии синглетное состояние имеет один и тот же вид в любом базисе, следовательно, мы можем выбрать систему координат, в которой $\hat{n}=\hat{z}$. Более того, симметрия состояния позволяет положить $\hat{m}=\hat{z}\cos\theta+\hat{x}\sin\theta$, так что мы находим

$$\langle\psi^-|\hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes\hat{m}\cdot\vec{\boldsymbol{\sigma}}_B|\psi^-\rangle = \langle\psi^-|\boldsymbol{\sigma}_z\otimes\boldsymbol{\sigma}_z|\psi^-\rangle\cos\theta +$$
$$+\langle\psi^-|\boldsymbol{\sigma}_z\otimes\boldsymbol{\sigma}_x|\psi^-\rangle\sin\theta = -\cos\theta,$$

что дает ответ

$$p=\frac{1}{4}-\frac{1}{8}\cos\theta.$$

Этот результат интуитивно понятен. С большей вероятностью мы обнаруживаем спины антипараллельными, так как матрица плотности имеет большую синглетную компоненту. По этой же причине менее вероятно обнаружить спины параллельными. Вариация между двумя возможностями, естественно, синусоидальная.

Решения упражнений к главе 3

3.1. Реализация ПОЗМ

Поскольку мы имеем $n = 4$ положительных оператора, действующих в $N = 2$-мерном гильбертовом пространстве $\mathcal{H}$, то согласно теореме Наймарка эту ПОЗМ можно расширить до «ортогонального измерения фон Неймана»[1] в $n = 4$-мерном гильбертовом пространстве $\mathcal{H} \oplus \mathcal{H}^{\perp}$. Сделаем это путем расширения $N = 2$ проекторов до четырех, требуя ортонормированность состояний, из которых формируются эти проекторы. На лекциях было показано, что с помощью следующего отображения

$$|\varphi_a\rangle \oplus |\varphi_a^{\perp}\rangle \rightarrow |\varphi_a\rangle_A|0\rangle_B + |0\rangle_A|\varphi_a^{\perp}\rangle_B, \qquad a = 1, \ldots, n,$$

вариант «прямой суммы» теоремы Наймарка можно преобразовать в вариант «тензорного произведения».

Если вспомогательная система приготовлена в состоянии $|0\rangle_B$, то это отображение гарантирует, что дальнейшая эволюция системы A будет ограничена подпространством $\mathcal{H}$. Размерность расширенного тензорным произведением пространства равна $N(n - N + 1)$, что в нашем случае равно шести. Однако это не самое эффективное из возможных отображений. Мы можем использовать следующее, более рациональное, отображение той же самой размерности:

$$|\varphi_a\rangle \oplus |\varphi_a^{\perp}\rangle \rightarrow |\varphi_a\rangle_A|0\rangle_B + |\varphi_a^{\perp}\rangle_A|1\rangle_B, \qquad a = 1, \ldots, n.$$

Очевидно, это отображение также ограничивает систему A подпространством $\mathcal{H}$, если вспомогательная система приготовлена в состоянии $|0\rangle_B$, но размерность тензорного произведения гильбертовых пространств теперь только $2N = 4$.

Чтобы найти ПЗИ, сначала найдем расширение в прямую сумму пространств, а затем применим приведенное выше отображение. Состояниями, включающими ПОЗМ, которую мы хотели бы расширить, являются

$$|\tilde{\psi}_1\rangle = \frac{1}{\sqrt{2}}|\uparrow_z\rangle, \qquad |\tilde{\psi}_3\rangle = \frac{1}{\sqrt{2}}|\uparrow_x\rangle,$$

$$|\tilde{\psi}_2\rangle = \frac{1}{\sqrt{2}}|\downarrow_z\rangle, \qquad |\tilde{\psi}_4\rangle = \frac{1}{\sqrt{2}}|\downarrow_x\rangle.$$

[1] Некоторые авторы называют этот тип измерения ПЗИ (проекторно-значное измерение). К их числу принадлежит и автор. ПЗИ гораздо более ясно, чем «ортогональное измерение фон Неймана».

Чтобы расширить базис, удобнее и понятнее переписать их в спинорной форме в $\hat{z}$-базисе, в котором два последних состояния записываются с помощью соотношений[1]

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle + |\downarrow_z\rangle),$$

$$|\downarrow_x\rangle = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle - |\downarrow_z\rangle),$$

$$|\tilde{\psi}_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\end{pmatrix}, \qquad |\tilde{\psi}_3\rangle = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix},$$

$$|\tilde{\psi}_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix}, \qquad |\tilde{\psi}_4\rangle = \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}.$$

Первым состоянием, которое будет расширено, является $|\tilde{\psi}_1\rangle$. Единственным ограничением является нормировка, поэтому мы можем расширить его с помощью любого 2-спинора в $\mathcal{H}^-$, имеющего норму 1/2. Существует множество выборов, но реально разумны только те из них, у которых или равна нулю одна компонента, или обе компоненты имеют одинаковые значения. Я продемонстрирую, что получится, если выбрать расширение спинора первого типа:

$$|u_1\rangle = \begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\\0\end{pmatrix}.$$

Следующий вектор должен быть ортогонален предыдущему и также нормирован. С точностью до произвольной фазы это фиксирует его расширение, которое мы можем применить. Вновь имеется только один разумный выбор, дающий тривиальную фазу:

$$|u_2\rangle = \begin{pmatrix}0\\1/\sqrt{2}\\0\\1/\sqrt{2}\end{pmatrix}.$$

[1] Записывая состояния таким образом, я явно использую соглашение о фазе спинора $|\psi(\theta,\varphi)\rangle = \begin{pmatrix}\cos\theta/2\\ e^{i\varphi}\sin\theta/2\end{pmatrix}$, как это обычно делается. Хотя соглашение о «распределенной фазе» $|\psi(\theta,\varphi)\rangle = \begin{pmatrix}e^{-i\varphi/2}\cos\theta/2\\ e^{i\varphi/2}\sin\theta/2\end{pmatrix}$ выглядит более симметрично, оно ведет к общим фазам в $|\uparrow_x\rangle$ и $|\downarrow_x\rangle$, которые труднее запомнить и уж во всяком случае нефизичны.

Два последних выбора полностью фиксируются требованиями ортогональности и нормировки и имеют вид

$$|u_3\rangle = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix}, \qquad |u_4\rangle = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}.$$

Теперь нужно применить наше отображение, чтобы преобразовать эти «прямые суммы» состояний в тензорные произведения состояний. То есть взять нашу исходную систему A и ввести вспомогательную систему B таким образом, чтобы базисные состояния в $\mathcal{H}^A \oplus \mathcal{H}^{A\perp}$ отобразились на базисные состояния в $\mathcal{H}^A \otimes \mathcal{H}^B$. Непосредственная проверка показывает, что это получается, если выполнить следующее отображение базисных векторов

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{pmatrix} \to \begin{pmatrix} |\uparrow_z\rangle_A |\uparrow_z\rangle_B \\ |\downarrow_z\rangle_A |\uparrow_z\rangle_B \\ |\uparrow_z\rangle_A |\downarrow_z\rangle_B \\ |\downarrow_z\rangle_A |\downarrow_z\rangle_B \end{pmatrix},$$

где введены обозначения $|0\rangle_B = |\uparrow_z\rangle_B$, $|1\rangle_B = |\downarrow_z\rangle_B$. Это позволяет записать проекторы в виде

$$\begin{aligned}
\Pi_1 &= |u_1\rangle\langle u_1| \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = |\uparrow_z\uparrow_x\rangle\langle\uparrow_z\uparrow_x|, \\
\Pi_2 &= |u_2\rangle\langle u_2| \\
&= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = |\downarrow_z\uparrow_x\rangle\langle\downarrow_z\uparrow_x|, \\
\Pi_3 &= |u_3\rangle\langle u_3| \\
&= \frac{1}{4}\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} = |\uparrow_x\downarrow_x\rangle\langle\uparrow_x\downarrow_x|,
\end{aligned}$$

$$\Pi_4 = |u_4\rangle\langle u_4|$$
$$= \frac{1}{4}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} = |\downarrow_x\downarrow_x\rangle\langle\downarrow_x\downarrow_x|.$$

Для того, чтобы в исходной системе была реализована ПОЗМ, вспомогательная система должна быть приготовлена в состоянии $|\uparrow_z\rangle_B$.

Как упоминалось в ходе вывода, существует свобода в выборе путей, следовательно, возможны другие решения.

3.2. Обратимость супероператоров

а) Предположим, что супероператор $\mathcal{M}$ имеет левый обратный супероператор $\mathcal{N}$ такой, что $\mathcal{N}\circ\mathcal{M} = \mathcal{I}$. Согласно теореме о представлении Крауса $\mathcal{M}$ и $\mathcal{N}$ имеют представления операторных сумм:

$$\mathcal{M}(\rho) = \sum_\mu \mathbf{M}_\mu \rho \mathbf{M}_\mu^\dagger,$$
$$\mathcal{N}(\rho) = \sum_a \mathbf{N}_a \rho \mathbf{N}_a^\dagger,$$
$$\sum_\mu \mathbf{M}_\mu^\dagger \mathbf{M}_\mu = \sum_a \mathbf{N}_a^\dagger \mathbf{N}_a = \mathbf{1}.$$

Более того, представление операторной суммы их композиции выражается на языке операторов $\mathbf{R}_{\{a\mu\}} = \mathbf{N}_a\mathbf{M}_\mu$:

$$\sum_{a\mu} \mathbf{R}_{\{a\mu\}} \rho \mathbf{R}_{\{a\mu\}}^\dagger = \sum_{a\mu} \mathbf{N}_a\mathbf{M}_\mu \rho (\mathbf{N}_a\mathbf{M}_\mu)^\dagger =$$
$$= \sum_{a\mu} \mathbf{N}_a\mathbf{M}_\mu \rho \mathbf{M}_\mu^\dagger\mathbf{N}_a^\dagger = \mathcal{N}\circ\mathcal{M}(\rho),$$

$$\sum_{a\mu} \mathbf{R}_{\{a\mu\}}^\dagger \mathbf{R}_{\{a\mu\}} = \sum_{a\mu} (\mathbf{N}_a\mathbf{M}_\mu)^\dagger \mathbf{N}_a\mathbf{M}_\mu =$$
$$= \sum_{a\mu} \mathbf{M}_\mu^\dagger\mathbf{N}_a^\dagger\mathbf{N}_a\mathbf{M}_\mu =$$
$$= \sum_\mu \mathbf{M}_\mu^\dagger \left(\sum_a \mathbf{N}_a^\dagger\mathbf{N}_a\right)\mathbf{M}_\mu = \mathbf{1}.$$

Но поскольку $\mathcal{N} \circ \mathcal{M} = \mathcal{I}$, то операторы $\mathbf{R}_{\{a\mu\}}$ одновременно должны быть операторами Крауса для тождественного супероператора, имеющего тривиальное представление $\mathcal{I}(\rho) = \mathbf{1}\rho\mathbf{1}^\dagger$. В наиболее общем случае операторы Крауса определены с точностью до унитарного поворота; отсюда следует, что $\mathbf{N}_a\mathbf{M}_\mu = \lambda_{a\mu}\mathbf{1}$, где $\lambda_{a\mu}$ — элемент унитарной матрицы, а $\sum_{a\mu} |\lambda_{a\mu}|^2 = \mathbf{1}$ согласно нормировке столбцов унитарной матрицы[1].

b) Используя тождественную вставку $\sum_a \mathbf{N}_a^\dagger \mathbf{N}_a = 1$ и соотношение $\mathbf{N}_a\mathbf{M}_\mu = \lambda_{a\mu}\mathbf{1}$ из части (a), получим требуемый результат:

$$\mathbf{M}_\nu^\dagger \mathbf{M}_\mu = \sum_a \mathbf{M}_\nu^\dagger \mathbf{N}_a^\dagger \mathbf{N}_a \mathbf{M}_\mu = \tag{6.281}$$

$$= \sum_a \left(\mathbf{N}_a\mathbf{M}_\nu\right)^\dagger \mathbf{N}_a\mathbf{M}_\mu = \tag{6.282}$$

$$= \left(\sum_a \lambda_{a\nu}^* \lambda_{a\mu}\right) \mathbf{1} = \gamma_{\nu\mu}\mathbf{1}. \tag{6.283}$$

c) Из части (b) мы знаем, что $\mathbf{M}_\nu^\dagger \mathbf{M}_\mu = \gamma_{\nu\mu}\mathbf{1}$. Этого достаточно, чтобы показать пропорциональность друг другу всех операторов $\mathbf{M}$, поскольку в этом случае разложение Крауса имеет всего одно слагаемое, которое в соответствии с нормировкой должно быть унитарным.

Так как мы рассматриваем только ненулевые операторы, нам известно, что $\gamma_{\nu\mu} \neq 0$. Таким образом[2],

$$\det \mathbf{M}_\mu^\dagger \mathbf{M}_\mu = \det \gamma_{\mu\mu}\mathbf{1},$$

$$\det \mathbf{M}_\mu^\dagger \det \mathbf{M}_\mu = \left(\gamma_{\mu\mu}\right)^n,$$

$$\left(\det \mathbf{M}_\mu\right)^* \det \mathbf{M}_\mu \neq 0,$$

$$\det \mathbf{M}_\mu \neq 0.$$

---

[1] На самом деле требование унитарности матрицы $\lambda_{a\mu}$ здесь излишне. Справедливость равенства $\mathcal{N} \circ \mathcal{M}(\rho) = \mathcal{I}(\rho) = \rho$ для любого оператора плотности ρ требует выполнения $\mathbf{N}_a\mathbf{M}_\mu = \lambda_{a\mu}\mathbf{1}$, где $\lambda_{a\mu}$ — элемент *произвольной* матрицы с единичной нормой Гильберта–Шмидта, определяемой уравнениями (6.104), (6.105). Впрочем уже в следующем пункте решения данной задачи унитарность матрицы $\lambda_{a\mu}$ не предполагается, в противном случае матрица $\gamma_{\nu\mu}$, определяемая как $\sum_a \lambda_{a\nu}^* \lambda_{a\mu} = \gamma_{\nu\mu}$ была бы равна единичной матрице $\gamma_{\nu\mu} = \delta_{\nu\mu}$. — *Прим. ред.*

[2] Здесь во втором равенстве n — размерность гильбертова пространства состояний рассматриваемой квантовой системы. — *Прим. ред.*

Следовательно, каждый из операторов $\mathbf{M}_\mu$ должен быть обратим и, в частности,

$$\mathbf{M}_\mu^\dagger \mathbf{M}_\mu = \gamma_{\mu\mu}\mathbf{1},$$
$$\mathbf{M}_\mu^\dagger = \gamma_{\mu\mu}\mathbf{M}_\mu^{-1}.$$

Отсюда следует, что все операторы $\mathbf{M}$ пропорциональны друг другу:

$$\mathbf{M}_\nu^\dagger \mathbf{M}_\mu = \gamma_{\nu\mu}\mathbf{1},$$
$$\mathbf{M}_\nu \mathbf{M}_\nu^\dagger \mathbf{M}_\mu = \gamma_{\nu\mu}\mathbf{M}_\nu,$$
$$\mathbf{M}_\nu(\gamma_{\nu\nu}\mathbf{M}_\nu^{-1})\mathbf{M}_\mu = \gamma_{\nu\mu}\mathbf{M}_\nu,$$
$$\mathbf{M}_\mu = \frac{\gamma_{\nu\mu}}{\gamma_{\nu\nu}}\mathbf{M}_\nu.$$

3.3. Как много супероператоров?

На лекции мы видели, что существует три эквивалентных способа установить, что $ является супероператором:

1. $ преобразует матрицы плотности в матрицы плотности.

2. $ является вполне положительным линейным отображением, сохраняющим эрмитовость и след своего аргумента.

3. $ имеет представление операторной суммы.

Необходимо найти количество степеней свободы $, используя любой из этих критериев. Здесь я опишу подходы, использующие только критерии (1) и (3). В каждом из этих подходов ρ рассматривается как $N \times N$ оператор плотности, который полностью описывает смешанное состояние (то есть ансамбль чистых состояний) в N-мерном гильбертовом пространстве.

Матрица плотности ρ является эрмитовой матрицей с единичным следом и, следовательно, зависит от $N^2 - 1$ свободных параметров. Однако было бы ошибкой думать, что действие $ сводится всего лишь к случайному перемешиванию этих параметров. Базис для ρ фактически является N^2-мерным, а ρ может быть записана как $\rho = \frac{1}{2}(\mathbf{1} + \vec{\alpha}\cdot\vec{\boldsymbol{\lambda}})$, где $\boldsymbol{\lambda}_i$ представляет собой $N^2 - 1$ линейно независимых базисных матриц. Как видно из этой записи, $ способен не только случайно перемешивать матрицы $\boldsymbol{\lambda}_i$, изменяя $\vec{\alpha}$, но также может отображать единицу на линейную комбинацию 1 и $\vec{\boldsymbol{\lambda}}$:

$$\$\left(\frac{1}{2}\mathbf{1}\right) = \frac{1}{2}(\mathbf{1} + \vec{\beta}\cdot\vec{\boldsymbol{\lambda}}) \quad \text{для некоторого } \vec{\beta}.$$

При таком подсчете количество свободных параметров равно $(N^2-1)^2$ для отображения $\vec{\lambda}$ и N^2-1 для аффинного сдвига $\mathbf{1}$, что в сумме дает $(N^2-1)^2+N^2-1=N^4-N^2$ вещественных параметров.

Если вы не убеждены в существовании аффинного сдвига, то посмотрите, как сдвигается центр сферы Блоха под действием канала затухания амплитуды в задаче 3.6 b.

Поскольку $ имеет представление операторной суммы, мы можем записать

$$\$(\rho)=\sum_{\mu}\mathbf{M}_{\mu}\rho\mathbf{M}_{\mu}^{\dagger},$$

где каждый оператор $\mathbf{M}_{\mu}\in \mathrm{GL}(N,C)$[1] зависит от $2N^2$ вещественных параметров. В $\mathrm{GL}(N,C)$ существует N^2 линейно независимых матриц, что означает, что *prima facie*[2] $ зависит самое большее от $2N^2(N^2)=2N^4$ вещественных параметров.

Матрицы $\mathbf{M}_{\mu}$ должны также удовлетворять условию нормировки

$$\sum_{\mu}\mathbf{M}_{\mu}^{\dagger}\mathbf{M}_{\mu}=\mathbf{1}.$$

Это дает только N^2 дополнительных связей, так как эрмитово сопряженное уравнение идентично записанному выше. Наконец, мы видели на лекции, что наиболее общей неоднозначностью в определении матриц $\mathbf{M}_{\mu}$ является унитарная перестановка операторов:

$$\mathbf{M}_{\mu}\rightarrow U_{\mu\nu}\mathbf{M}_{\nu}.$$

Так как существует самое большее N^2 матриц $\mathbf{M}_{\mu}$, то $U_{\mu\nu}\in \mathrm{U}(N^2)$ зависит от N^4 вещественных параметров. Таким образом, мы находим, что $ зависит самое большее от $2N^4-N^2-N^4=N^4-N^2$ вещественных параметров.

В обоих подсчетах мы нашли, что $ зависит самое большее от N^4-N^2 вещественных параметров.

3.4. Насколько быстра декогерентизация?

a) Уравнение движения простого затухающего гармонического осциллятора имеет вид

$$m\ddot{x}+b\dot{x}+m\omega^2x=0.$$

[1] $\mathrm{GL}(N,C)$ группа невырожденных матриц размерности N над полем комплексных чисел C. — *Прим. ред.*

[2] *Prima facie* (лат) — по первому виду, на первый взгляд. — *Прим. перев.*

Мы ожидаем, что при слабом затухании средняя энергия осциллятора убывает экспоненциально:

$$\langle E(t)\rangle = E_0 e^{-bt/m}.$$

Таким образом, *амплитуда* осцилляций должна затухать как $e^{-bt/2m}$. Из классической механики или откуда-нибудь еще мы помним, что при слабом затухании добротность определяется как

$$Q = 2\pi \left(\frac{\text{Полная энергия}}{\text{Потеря энергии за период}} \right) = \frac{\omega}{b/m}.$$

На лекции мы нашли, что декогерентизация хорошо моделируется каналом затухания фазы. Из основного уравнения для этого канала следует, что недиагональные в базисе когерентных состояний элементы матрицы плотности затухают как

$$\boldsymbol{\rho}_{nm}(t) = \boldsymbol{\rho}_{nm}(0) e^{-\Gamma |n-m|^2 t/2},$$

где Γ — темп рассеяния одного кванта осциллятора его окружением. Такой вид затухания наводит на мысль интерпретировать Γ как коэффициент эффективной радиационной силы затухания с добротностью

$$Q = \frac{\omega}{\Gamma}.$$

Время декогерентизации системы по порядку величины представляет собой время, за которое недиагональные элементы уменьшаются в e раз по сравнению с их начальными значениями:

$$t_{\text{decoh}} = \frac{2}{\Gamma |n-m|^2}.$$

Данное в задаче кот-состояние не выражается в базисе когерентных состояний. Однако для сильно локализованных гауссовских волновых пакетов мы ожидаем, что собственное состояние оператора уничтожения будет примерно пропорционально собственному состоянию $\hat{\mathbf{x}}$-оператора:

$$\hat{\mathbf{a}} = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{\mathbf{x}} + \frac{i}{m\omega} \hat{\mathbf{p}} \right),$$

$$\langle \hat{\mathbf{a}} \rangle = \sqrt{\frac{m\omega}{2\hbar}} \left(\langle \hat{\mathbf{x}} \rangle + \frac{i}{m\omega} \langle \hat{\mathbf{p}} \rangle \right) = \sqrt{\frac{m\omega}{2\hbar}} \langle \hat{\mathbf{x}} \rangle.$$

Следовательно, мы ожидаем, что показатель экспоненты недиагональных элементов матрицы плотности будет иметь порядок

$$|n-m|^2 = \frac{m\omega}{2\hbar}|x-(-x)|^2 = \frac{2m\omega x^2}{\hbar}.$$

Теперь у нас есть все необходимое, чтобы вычислить время декогерентизации маятника:

$$\begin{aligned} t_{\mathrm{decoh}} &= \frac{2}{\Gamma|n-m|^2} = \\ &= \frac{2Q\hbar}{\omega(2m\omega x^2)} = \\ &= \frac{Q\hbar}{m\omega^2 x^2} = \\ &= \frac{10^9 \cdot 10^{-34}\mathrm{J}\cdot\mathrm{s}}{10^{-3}\mathrm{kg}\cdot 1\,\mathrm{s}^{-2}\cdot 10^{-4}\mathrm{m}^2} = 10^{-18}\mathrm{s}. \end{aligned}$$

b) При нулевой температуре все уровни энергии осциллятора были связаны с основным состоянием окружения. При конечной температуре $n = \frac{kT}{\hbar\omega}$ состояний окружения доступны для взаимодействия[1]. Таким образом, по порядку величины темп затухания становится в n раз быстрее. Соответственно время декогерентизации должно уменьшиться на этот фактор:

$$\begin{aligned} t_{\mathrm{decoh}}(T) &= \frac{\hbar\omega}{kT} t_{\mathrm{decoh}}(0) = \\ &= \frac{10^{-34}\mathrm{J}\cdot\mathrm{s}\cdot 1\,\mathrm{s}^{-1}}{10^{-23}\mathrm{J}\cdot\mathrm{K}^{-1}\cdot 10^{2}\mathrm{K}} \cdot 10^{-18}\mathrm{s} = 10^{-31}\mathrm{s}. \end{aligned}$$

Мораль: декогерентизация — очень быстра. Это один из самых быстрых известных в настоящее время физических процессов.

3.5. Затухание фазы

a) Непосредственно видно, что $\mathbf{M}_0$, $\mathbf{M}_1$ и $\mathbf{M}_2$ выражаются только через две линейно независимые матрицы ($\mathbf{1}$ и σ_3). Это наводит на мысль, что возможно представление операторной суммы, использующее *только два*

[1] Это справедливо при $kT \gg \hbar\omega$. — *Прим. ред.*

оператора Крауса. (Фактически всегда, когда набор операторов Крауса зависит от n линейно независимых матриц, можно найти представление операторной суммы, использующее эти n операторов.)

Посмотрим явно, как операторы $\mathbf{M}$ действуют на матрицу плотности $\boldsymbol{\rho}$ общего вида

$$\begin{aligned}\boldsymbol{\rho} \rightarrow\ & \mathbf{M}_0\boldsymbol{\rho}\mathbf{M}_0^\dagger + \mathbf{M}_1\boldsymbol{\rho}\mathbf{M}_1^\dagger + \mathbf{M}_2\boldsymbol{\rho}\mathbf{M}_2^\dagger = \\ =\ & (1-p)\boldsymbol{\rho} + \frac{p}{4}(\mathbf{1}+\boldsymbol{\sigma}_3)\boldsymbol{\rho}(\mathbf{1}+\boldsymbol{\sigma}_3) + \frac{p}{4}(\mathbf{1}-\boldsymbol{\sigma}_3)\boldsymbol{\rho}(\mathbf{1}-\boldsymbol{\sigma}_3) = \\ =\ & \left(1-\frac{p}{2}\right)\boldsymbol{\rho} + \frac{p}{2}\boldsymbol{\sigma}_3\boldsymbol{\rho}\boldsymbol{\sigma}_3.\end{aligned}$$

Эта форма подсказывает выбор

$$\mathbf{N}_0 = \sqrt{1-\frac{p}{2}}\,\mathbf{1},$$

$$\mathbf{N}_1 = \sqrt{\frac{p}{2}}\,\boldsymbol{\sigma}_3$$

в качестве операторов Крауса канала затухания фазы. Действительно, $\mathbf{N}_0$ и $\mathbf{N}_1$ удовлетворяют условию $\mathbf{N}_0^\dagger\mathbf{N}_0 + \mathbf{N}_1^\dagger\mathbf{N}_1 = \mathbf{1}$ и, следовательно, должным образом нормированы.

b) Соотношение $\mathbf{M}_\mu = U_{\mu a}\mathbf{N}_a$ дает следующую систему уравнений для компонент $U_{\mu a}$:

$$\sqrt{1-p}\,\mathbf{1} = U_{00}\sqrt{1-\frac{p}{2}}\,\mathbf{1} + U_{01}\sqrt{\frac{p}{2}}\,\boldsymbol{\sigma}_3,$$

$$\sqrt{\frac{p}{4}}(\mathbf{1}+\boldsymbol{\sigma}_3) = U_{10}\sqrt{1-\frac{p}{2}}\,\mathbf{1} + U_{11}\sqrt{\frac{p}{2}}\,\boldsymbol{\sigma}_3,$$

$$\sqrt{\frac{p}{4}}(\mathbf{1}-\boldsymbol{\sigma}_3) = U_{20}\sqrt{1-\frac{p}{2}}\,\mathbf{1} + U_{21}\sqrt{\frac{p}{2}}\,\boldsymbol{\sigma}_3.$$

Сравнение коэффициентов при линейно независимых $\mathbf{1}$ и $\boldsymbol{\sigma}_3$ дает

$$U_{00} = \sqrt{\frac{2-2p}{2-p}}, \qquad U_{01} = 0,$$

$$U_{10} = \sqrt{\frac{p}{4-2p}}, \qquad U_{11} = \sqrt{\frac{1}{2}},$$

$$U_{20} = \sqrt{\frac{p}{4-2p}}, \qquad U_{21} = -\sqrt{\frac{1}{2}}.$$

Осталось лишь дополнить матрицу U до унитарной, потребовав, чтобы все ее строки и столбцы были взаимно ортогональны и нормированы:

$$|U_{00}|^2 + |U_{01}|^2 + |U_{02}|^2 = 1 \Rightarrow U_{02} = e^{i\theta}\sqrt{\frac{p}{2-p}},$$

$$|U_{10}|^2 + |U_{11}|^2 + |U_{12}|^2 = 1 \Rightarrow U_{12} = e^{i\varphi}\sqrt{\frac{1-p}{2-p}},$$

$$|U_{20}|^2 + |U_{21}|^2 + |U_{22}|^2 = 1 \Rightarrow U_{22} = e^{i\psi}\sqrt{\frac{1-p}{2-p}},$$

$$U_{01}^* U_{02} + U_{11}^* U_{12} + U_{21}^* U_{22} = 0 \Rightarrow e^{i(\varphi-\psi)} = 1 \Rightarrow \varphi = \psi,$$

$$U_{00}^* U_{02} + U_{10}^* U_{12} + U_{20}^* U_{22} = 0 \Rightarrow e^{i(\theta-\varphi)} = -1 \Rightarrow \theta = \varphi + \pi.$$

Больше связей нет, следовательно, с точностью до неопределенной общей фазы ($\mathbf{N}_2 = 0$ не может иметь хорошо определенной фазы)

$$U = \begin{pmatrix} \sqrt{\dfrac{2-2p}{2-p}} & 0 & -e^{i\varphi}\sqrt{\dfrac{p}{2-p}} \\ \sqrt{\dfrac{p}{4-2p}} & \sqrt{\dfrac{1}{2}} & e^{i\varphi}\sqrt{\dfrac{1-p}{2-p}} \\ \sqrt{\dfrac{p}{4-2p}} & -\sqrt{\dfrac{1}{2}} & e^{i\varphi}\sqrt{\dfrac{1-p}{2-p}} \end{pmatrix}.$$

c) Операторы Крауса для канала, имеющего унитарное представление U_{AE}, определяются как

$$\mathbf{M}_\mu \equiv \langle \mu_E | U_{AE} | 0_E \rangle,$$

где $|\mu\rangle_E$ — *ортогональные* состояния окружения. Мы можем обычным способом сформировать ортогональный базис окружения из $\{|0\rangle_E, |\gamma_0\rangle_E, |\gamma_1\rangle_E\}$. Одним из методов является применение процесса Грама — Шмидта, но вместо этого я выберу базис, отражающий симметрию между $|\gamma_0\rangle_E$ и $|\gamma_1\rangle_E$:

$$|\pm\rangle_E = \frac{\alpha_\pm}{\sqrt{2}}\left(|\gamma_0\rangle_E \pm |\gamma_1\rangle_E\right),$$

$$\langle \pm | \pm \rangle = 1,$$

$$\Rightarrow |\alpha_\pm|^2\left(1 \pm \langle \gamma_0 | \gamma_1 \rangle\right) = 1,$$

$$\Rightarrow \alpha_\pm = \sqrt{\frac{1}{1 \pm (1-\varepsilon)}}.$$

В этом базисе операторы Крауса имеют вид

$$\begin{aligned}
\mathbf{M}_0 &= \langle 0_E|U_{AE}|0_E\rangle = \sqrt{1-p}\,\mathbf{1},\\
\mathbf{M}_\pm &= \langle \pm_E|U_{AE}|0_E\rangle =\\
&= \sqrt{\frac{p}{2[1\pm(1-\varepsilon)]}}\begin{pmatrix} 1\pm(1-\varepsilon) & 0\\ 0 & \pm[1\pm(1-\varepsilon)]\end{pmatrix} =\\
&= \sqrt{\frac{p[1\pm(1-\varepsilon)]}{2}}\begin{pmatrix} 1 & 0\\ 0 & \pm 1\end{pmatrix}.
\end{aligned}$$

Они не похожи на операторы канала затухания фазы, но их можно преобразовать в три таких оператора. И даже более того, их можно преобразовать в *два* оператора, которые выглядят как операторы канала затухания фазы. Чтобы найти их, рассмотрим, как и в части (a), действие операторов $\mathbf{M}_{0,\pm}$ на произвольную матрицу плотности

$$\begin{aligned}
\boldsymbol{\rho} \;&\to\; \mathbf{M}_0\boldsymbol{\rho}\mathbf{M}_0^\dagger + \mathbf{M}_+\boldsymbol{\rho}\mathbf{M}_+^\dagger + \mathbf{M}_-\boldsymbol{\rho}\mathbf{M}_-^\dagger\\
&= (1-p)\boldsymbol{\rho} + \frac{p(2-\varepsilon)}{2}\boldsymbol{\rho} + \frac{p\varepsilon}{2}\boldsymbol{\sigma}_3\boldsymbol{\rho}\boldsymbol{\sigma}_3\\
&= \left(1-\frac{p\varepsilon}{2}\right)\boldsymbol{\rho} + \frac{p\varepsilon}{2}\boldsymbol{\sigma}_3\boldsymbol{\rho}\boldsymbol{\sigma}_3,
\end{aligned}$$

$$\begin{aligned}
\mathbf{N}_0 &= \sqrt{1-\frac{p\varepsilon}{2}}\,\mathbf{1},\\
\mathbf{N}_1 &= \sqrt{\frac{p\varepsilon}{2}}\,\boldsymbol{\sigma}_3.
\end{aligned}$$

В такой форме очевидно, что это операторы Крауса для канала затухания фазы, имеющего вероятность декогерентизации с его окружением, равную εp. Обратим внимание на то, что при $\varepsilon \to 1$ мы воспроизводим канал затухания фазы из части (a), а при $\varepsilon \to 0$ затухание фазы исчезает.

d) Если канал из (c) описывает рассеяние отдельного фотона, то мы имеем $\Gamma_{\text{scatt}} = p\Delta t$. Но декогерентизация возникает только тогда, когда окружение может различить результаты рассеяния, то есть $\Gamma_{\text{decoh}} = \varepsilon p\Delta t$. Следовательно,

$$\Gamma_{\text{decoh}} = \varepsilon\Gamma_{\text{scatt}}.$$

3.6. Декогерентизация на сфере Блоха

a) Под действием канала затухания фазы матрица плотности $\rho = \frac{1}{2}(\mathbf{1} + \vec{P}\cdot\vec{\sigma})$ эволюционирует как (используя операторы M_μ из задачи 3.5)

$$\begin{aligned}
\rho \to\ & \mathbf{M}_0\rho\mathbf{M}_0^\dagger + \mathbf{M}_1\rho\mathbf{M}_1^\dagger + \mathbf{M}_2\rho\mathbf{M}_2^\dagger = \\
=\ & (1-p)\rho + \frac{p}{4}(\mathbf{1}+\sigma_3)\rho(\mathbf{1}+\sigma_3) + \frac{p}{4}(\mathbf{1}-\sigma_3)\rho(\mathbf{1}-\sigma_3) = \\
=\ & \left(1-\frac{p}{2}\right)\rho + \frac{p}{2}\sigma_3\rho\sigma_3 = \\
=\ & \left(1-\frac{p}{2}\right)\rho + \frac{p}{2}\left[\frac{1}{2}\Big(\mathbf{1}+\sigma_3(\vec{P}\cdot\vec{\sigma})\sigma_3\Big)\right] = \\
=\ & \left(1-\frac{p}{2}\right)\rho + \frac{p}{2}\left[\frac{1}{2}\Big(\mathbf{1}-\vec{P}\cdot\vec{\sigma}+2P_3\sigma_3\Big)\right] = \\
=\ & \frac{1}{2}\left[\left(1-\frac{p}{2}+\frac{p}{2}\right)\mathbf{1} + \left(1-\frac{p}{2}-\frac{p}{2}\right)\vec{P}\cdot\vec{\sigma} + pP_3\sigma_3\right] = \\
=\ & \frac{1}{2}\Big[\mathbf{1} + (1-p)\vec{P}\cdot\vec{\sigma} + pP_3\sigma_3\Big] = \\
=\ & \frac{1}{2}\Big[\mathbf{1} + \Big((1-p)P_1, (1-p)P_2, P_3\Big)\cdot\vec{\sigma}\Big].
\end{aligned}$$

Таким образом, мы видим, что действие канала затухания фазы сжимает сферу Блоха, превращая ее в вытянутый вдоль оси z сфероид (эллипсоид вращения). Выделенное положение оси z означает, что канал затухания фазы действует в некотором предпочтительном базисе.

b) Под действием канала затухания амплитуды матрица плотности ρ эволюционирует как

$$\begin{aligned}
\rho \to\ & \mathbf{M}_0\rho\mathbf{M}_0^\dagger + \mathbf{M}_1\rho\mathbf{M}_1^\dagger \\
=\ & \frac{1}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}(\mathbf{1}+\vec{P}\cdot\vec{\sigma})\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}\right] + \\
& + \frac{1}{2}\left[\begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}(\mathbf{1}+\vec{P}\cdot\vec{\sigma})\begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}\right] = \\
=\ & \frac{1}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}\begin{pmatrix} 1+P_3 & P_1-iP_2 \\ P_1+iP_2 & 1-P_3 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}\right] +
\end{aligned}$$

$$+\frac{1}{2}\left[\begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1+P_3 & P_1-iP_2 \\ P_1+iP_2 & 1-P_3 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}\right] =$$

$$= \frac{1}{2}\begin{pmatrix} 1+P_3+p-pP_3 & (P_1-iP_2)\sqrt{1-p} \\ (P_1+iP_2)\sqrt{1-p} & 1-P_3-p+pP_3 \end{pmatrix} =$$

$$= \frac{1}{2}\left[\mathbf{1} + \left(\sqrt{1-p}\,P_1, \sqrt{1-p}\,P_2, P_3 + p(1-P_3)\right)\cdot\vec{\boldsymbol{\sigma}}\right].$$

Таким образом, мы видим, что действие канала затухания амплитуды сжимает сферу Блоха в сплюснутый вдоль оси z эллипсоид вращения и сдвигает ее вверх.

c) Под действием «двойного канала Паули» матрица плотности $\boldsymbol{\rho}$ эволюционирует как

$$\boldsymbol{\rho} \to \mathbf{M}_0\boldsymbol{\rho}\mathbf{M}_0^\dagger + \mathbf{M}_1\boldsymbol{\rho}\mathbf{M}_1^\dagger + \mathbf{M}_2\boldsymbol{\rho}\mathbf{M}_2^\dagger =$$

$$= (1-p)\boldsymbol{\rho} + \frac{p}{2}\boldsymbol{\sigma}_1\boldsymbol{\rho}\boldsymbol{\sigma}_1 + \frac{p}{2}\boldsymbol{\sigma}_3\boldsymbol{\rho}\boldsymbol{\sigma}_3 =$$

$$= (1-p)\boldsymbol{\rho} + \frac{p}{2}\left[\frac{1}{2}\left(\mathbf{1}+\boldsymbol{\sigma}_1(\vec{P}\cdot\vec{\boldsymbol{\sigma}})\boldsymbol{\sigma}_1\right) + \frac{1}{2}\left(\mathbf{1}+\boldsymbol{\sigma}_3(\vec{P}\cdot\vec{\boldsymbol{\sigma}})\boldsymbol{\sigma}_3\right)\right] =$$

$$= (1-p)\boldsymbol{\rho} + \frac{p}{2}\left[1 - \frac{1}{2}\vec{P}\cdot\vec{\boldsymbol{\sigma}} + P_1\boldsymbol{\sigma}_1 - \frac{1}{2}\vec{P}\cdot\vec{\boldsymbol{\sigma}} + P_3\boldsymbol{\sigma}_3\right] =$$

$$= \frac{1}{2}\left[(1-p+p)\mathbf{1} + (1-p-p)\vec{P}\cdot\vec{\boldsymbol{\sigma}} + p\vec{P}\cdot\vec{\boldsymbol{\sigma}} - pP_2\boldsymbol{\sigma}_2\right] =$$

$$= \frac{1}{2}\left[1 + \left((1-p)P_1, (1-2p)P_2, (1-p)P_3\right)\cdot\vec{\boldsymbol{\sigma}}\right].$$

Таким образом, действие двойного канала Паули при $p < \frac{1}{2}$ сжимает сферу Блоха в сплюснутый вдоль оси y эллипсоид вращения, а при $p > \frac{1}{2}$ — в вытянутый вдоль оси y инвертированный сфероид (однополостной гиперболоид вращения).

3.7. Декогерентизация затухающего осциллятора

a) Рассмотрим производную X по времени:

$$\dot{X} = \mathrm{tr}\left[\dot{\boldsymbol{\rho}}_I(t)e^{\lambda\mathbf{a}^\dagger}e^{-\lambda^*\mathbf{a}}\right] =$$

$$= \Gamma\,\mathrm{tr}\left[\left(\mathbf{a}\boldsymbol{\rho}_I\mathbf{a}^\dagger - \frac{1}{2}\mathbf{a}^\dagger\mathbf{a}\boldsymbol{\rho}_I - \frac{1}{2}\boldsymbol{\rho}_I\mathbf{a}^\dagger\mathbf{a}\right)e^{\lambda\mathbf{a}^\dagger}e^{-\lambda^*\mathbf{a}}\right].$$

Чтобы упростить это выражение, мы хотим преобразовать два вторых слагаемых под знаком следа к такому же виду, что и первое (с целью по возможности сократить их друг с другом). Это можно сделать, используя свойство инвариантности следа относительно циклических перестановок и коммутационные соотношения между операторами уничтожения и рождения:

$$[\mathbf{a}, \mathbf{a}^\dagger] = \mathbf{1},$$

$$[\mathbf{a}, e^{\lambda \mathbf{a}^\dagger}] = [\mathbf{a}, \mathbf{a}^\dagger] \frac{\partial}{\partial \mathbf{a}^\dagger} \left(e^{\lambda \mathbf{a}^\dagger} \right) = \lambda e^{\lambda \mathbf{a}^\dagger},$$

$$[e^{-\lambda^* \mathbf{a}}, \mathbf{a}^\dagger] = \frac{\partial}{\partial \mathbf{a}} \left(e^{-\lambda^* \mathbf{a}} \right) [\mathbf{a}, \mathbf{a}^\dagger] = -\lambda^* e^{-\lambda^* \mathbf{a}}.$$

Применяя эти манипуляции к $\dot{X}$, найдем

$$\dot{X} = \Gamma \operatorname{tr} \left[\left(\mathbf{a} \boldsymbol{\rho}_I \mathbf{a}^\dagger - \frac{1}{2} \mathbf{a} \boldsymbol{\rho}_I (\mathbf{a}^\dagger - \lambda^*) - \frac{1}{2} (\mathbf{a} + \lambda) \boldsymbol{\rho}_I \mathbf{a}^\dagger \right) e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}} \right] =$$

$$= \frac{\Gamma}{2} \operatorname{tr} \left[\lambda^* \boldsymbol{\rho}_I e^{\lambda \mathbf{a}^\dagger} \mathbf{a} e^{-\lambda^* \mathbf{a}} - \lambda \boldsymbol{\rho}_I \mathbf{a}^\dagger e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}} \right].$$

От лишних операторов рождения и уничтожения можно избавиться, используя правила дифференцирования экспонент:

$$\frac{\partial}{\partial \lambda^*} e^{-\lambda^* \mathbf{a}} = -\mathbf{a} e^{-\lambda^* \mathbf{a}},$$

$$\frac{\partial}{\partial \lambda} e^{\lambda \mathbf{a}^\dagger} = \mathbf{a}^\dagger e^{\lambda \mathbf{a}^\dagger};$$

таким образом, мы получаем для $\dot{X}$ дифференциальное уравнение в частных производных:

$$\dot{X} = -\frac{\Gamma}{2} \operatorname{tr} \left[\lambda^* \boldsymbol{\rho}_I e^{\lambda \mathbf{a}^\dagger} \frac{\partial}{\partial \lambda^*} \left(e^{-\lambda^* \mathbf{a}} \right) + \lambda \boldsymbol{\rho}_I \frac{\partial}{\partial \lambda} \left(e^{\lambda \mathbf{a}^\dagger} \right) e^{-\lambda^* \mathbf{a}} \right] =$$

$$= -\frac{\Gamma}{2} \lambda^* \frac{\partial}{\partial \lambda^*} \operatorname{tr} \left[\boldsymbol{\rho}_I e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}} \right] - \frac{\Gamma}{2} \lambda \frac{\partial}{\partial \lambda} \operatorname{tr} \left[\boldsymbol{\rho}_I e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}} \right] =$$

$$= -\frac{\Gamma}{2} \left(\lambda^* \frac{\partial X}{\partial \lambda^*} + \lambda \frac{\partial X}{\partial \lambda} \right).$$

Здесь я довольно бесцеремонно переставил порядок операций дифференцирования и вычисления следа. Для данных целей я буду предполагать, что

здесь все относящиеся к делу функции равномерно непрерывны, так что эта коммутация разрешена.

Используя правило дифференцирования сложных функций (цепное правило), мы можем записать для X линейное уравнение в частных производных с постоянными коэффициентами:

$$\dot{X} = -\frac{\Gamma}{2}\left(\frac{\partial X}{\partial \ln \lambda^*} + \frac{\partial X}{\partial \ln \lambda}\right).$$

Естественно предположить, что решение этого равнения является функцией от линейной комбинации его аргументов:

$$X = X(\alpha \ln \lambda^* + \beta \ln \lambda + \gamma t).$$

Подставляя этот *анзац* в уравнение, мы находим соотношение между коэффициентами

$$\gamma = -\frac{\Gamma}{2}(\alpha + \beta),$$

что дает искомый скейлинговый закон:

$$\begin{aligned} X(\vec{\lambda}, t) &= X\left(\alpha \ln \lambda^* + \beta \ln \lambda - \frac{\Gamma}{2}(\alpha + \beta)t\right) \\ &= X\left(\alpha \ln\left(\lambda^* e^{-\Gamma t/2}\right) + \beta \ln\left(\lambda e^{-\Gamma t/2}\right)\right) \\ &= X(\vec{\lambda}', 0), \end{aligned}$$

$$\vec{\lambda}' = \vec{\lambda} e^{-\Gamma t/2}.$$

b) Прежде чем начинать решение, следует заметить, что этот кот ненормален не только в житейском смысле, но и в смысле борновской интерпретации. Чтобы должным образом нормировать этого кота, нам нужно положить равным единице «бра-кот кет-кот»:

$$|\mathrm{cat}\rangle = \frac{N}{\sqrt{2}}\left(|\alpha_1\rangle + |\alpha_2\rangle\right),$$

$$\langle \mathrm{cat}|\mathrm{cat}\rangle = \frac{|N|^2}{2}\left(\langle\alpha_1|\alpha_1\rangle + \langle\alpha_1|\alpha_2\rangle + \langle\alpha_2|\alpha_1\rangle + \langle\alpha_2|\alpha_2\rangle\right) = 1.$$

Но вместо того чтобы отвлекаться на дальнейшие детали нормировки этого кота, я просто замечу, что перед ним должен быть нормирующий множитель.

Результаты части (а) показывают, как связаны между собой след кота и оператора в момент времени t с их следом в момент $t = 0$. Однако оператором под знаком следа является оператор «смещения» $\mathbf{D}_\lambda = e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}}$, который переводит одно когерентное состояние в другое. Поскольку оператор любой наблюдаемой осциллятора можно разложить по этим операторам сдвига[1], то временная эволюция кота полностью определяется тем, как изменяется во времени действие на него этих операторов сдвига.

Конкретнее, согласно части (а):

$$\mathrm{tr}\left[|\mathrm{cat}(0)\rangle\langle\mathrm{cat}(0)| e^{\lambda' \mathbf{a}^\dagger} e^{-\lambda'^* \mathbf{a}}\right] = \mathrm{tr}\left[\rho_{\mathrm{cat}}(t) e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}}\right].$$

В общем случае $\rho_{\mathrm{cat}}(t)$ не обязано быть чистым состоянием (фактически мы увидим, что оно им не является), но до поры до времени будем предполагать его чистым. Это даст возможность преобразовать следы в математические ожидания. С помощью внутреннего произведения когерентных состояний

$$\begin{aligned}\langle\alpha|\beta\rangle &= e^{\alpha^*\beta - \left(|\alpha|^2 + |\beta|^2\right)/2} = \\ &= e^{-\frac{1}{2}\left(|\alpha|^2 + |\beta|^2 - 2\mathrm{Re}(\alpha^*\beta)\right)} e^{i\mathrm{Im}(\alpha^*\beta)} = \\ &= e^{-\frac{1}{2}|\alpha-\beta|^2} e^{i\mathrm{Im}(\alpha^*\beta)}\end{aligned}$$

мы находим, что

$$\langle\mathrm{cat}(t)| e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}} |\mathrm{cat}(t)\rangle = \langle\mathrm{cat}(0)| e^{\lambda' \mathbf{a}^\dagger} e^{-\lambda'^* \mathbf{a}} |\mathrm{cat}(0)\rangle,$$

$$\begin{pmatrix} e^{\lambda\alpha_1^*(t) - \lambda^*\alpha_1(t)} + \\ e^{\lambda\alpha_2^*(t) - \lambda^*\alpha_2(t)} + \\ \langle\alpha_1(t)|\alpha_2(t)\rangle e^{\lambda\alpha_1^*(t) - \lambda^*\alpha_2(t)} + \\ \langle\alpha_2(t)|\alpha_1(t)\rangle e^{\lambda\alpha_2^*(t) - \lambda^*\alpha_1(t)} \end{pmatrix} \cdot \begin{pmatrix} e^{(\lambda\alpha_1^* - \lambda^*\alpha_1) e^{-\Gamma t/2}} + \\ e^{(\lambda\alpha_2^* - \lambda^*\alpha_2) e^{-\Gamma t/2}} + \\ \langle\alpha_1|\alpha_2\rangle e^{(\lambda\alpha_1^* - \lambda^*\alpha_2) e^{-\Gamma t/2}} + \\ \langle\alpha_2|\alpha_1\rangle e^{(\lambda\alpha_2^* - \lambda^*\alpha_1) e^{-\Gamma t/2}} \end{pmatrix}.$$

[1] См., например, A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer Academic Publishers, New York et al, 2002. [Оператор сдвига $\mathbf{D}(\lambda)$, действуя на $|0\rangle$ — основное состояние гармонического осциллятора, преобразует в $|\lambda\rangle = \mathbf{D}(\lambda)|0\rangle$ — собственное состояние оператора уничтожения $\mathbf{a}|\lambda\rangle = \lambda|\lambda\rangle$ (когерентное состояние). Преобразование одного когерентного состояния в другое обеспечивает закон умножения операторов сдвига $\mathbf{D}(\lambda)\mathbf{D}(\mu) = \mathbf{D}(\lambda + \mu)\exp[(\lambda\mu^* - \lambda^*\mu)/2]$. На русском языке с теорией когерентных состояний можно познакомиться по книгам И. А. Малкин, В. И. Манько, *Динамические симметрии и когерентные состояния квантовых систем*, М.: Наука (1979); А. М. Переломов, *Обобщенные когерентные состояния и их применения*, (1987). — *Прим. ред.*]

Эти слагаемые можно было бы почти согласовать друг с другом, предполагая, что эволюция во времени преобразует чистое состояние в другое чистое, но это ведет к тому, что недиагональные элементы не подходят друг к другу:

$$
\begin{aligned}
|\alpha_{1,2}\rangle &\to \left|\alpha_{1,2}e^{-\Gamma t/2}\right\rangle,\\
e^{\lambda \mathbf{a}^\dagger} e^{-\lambda^* \mathbf{a}} \left|\alpha_{1,2}e^{-\Gamma t/2}\right\rangle &= e^{(\lambda\alpha_{1,2}^* - \lambda^*\alpha_{1,2})e^{-\Gamma t/2}} \left|\alpha_{1,2}e^{-\Gamma t/2}\right\rangle,\\
\left\langle \alpha_{1,2}e^{-\Gamma t/2}\right| \left.\alpha_{2,1}e^{-\Gamma t/2}\right\rangle &= \langle\alpha_2|\alpha_1\rangle^{e^{-\Gamma t}} \neq \langle\alpha_2|\alpha_1\rangle.
\end{aligned}
$$

Чтобы исправить это несоответствие недиагональных элементов, мы должны потребовать, чтобы недиагональные компоненты кота затухали быстрее диагональных, как раз базис для когерентных состояний является затухающим. Это ведет к полностью перемешивающей состояния эволюции:

$$
\begin{aligned}
|\alpha_1\rangle\langle\alpha_1| &\to \left|\alpha_1 e^{-\Gamma t/2}\right\rangle \left\langle \alpha_1 e^{-\Gamma t/2}\right|,\\
|\alpha_2\rangle\langle\alpha_2| &\to \left|\alpha_2 e^{-\Gamma t/2}\right\rangle \left\langle \alpha_2 e^{-\Gamma t/2}\right|,\\
|\alpha_1\rangle\langle\alpha_2| &\to \langle\alpha_1|\alpha_2\rangle^{(1-e^{-\Gamma t})} \left|\alpha_1 e^{-\Gamma t/2}\right\rangle \left\langle \alpha_2 e^{-\Gamma t/2}\right|,\\
|\alpha_2\rangle\langle\alpha_1| &\to \langle\alpha_2|\alpha_1\rangle^{(1-e^{-\Gamma t})} \left|\alpha_2 e^{-\Gamma t/2}\right\rangle \left\langle \alpha_1 e^{-\Gamma t/2}\right|.
\end{aligned}
$$

Таким образом, наш кот эволюционирует в нечто более диагональное:

$$
\begin{aligned}
|\mathrm{cat}(0)\rangle\langle\mathrm{cat}(0)| &\to \frac{|N|^2}{2} \begin{pmatrix} 1 & \langle\alpha_1|\alpha_2\rangle^{(1-e^{-\Gamma t})} \\ \langle\alpha_2|\alpha_1\rangle^{(1-e^{-\Gamma t})} & 1 \end{pmatrix} =\\
&= \frac{|N|^2}{2}\Big[\mathbf{1} + e^{-\frac{1}{2}|\alpha_1-\alpha_2|^2(1-e^{-\Gamma t})}\big(\sigma_x \cos\theta_{21}(t) +\\
&\qquad + \sigma_y \sin\theta_{21}(t)\big)\Big],
\end{aligned}
$$

где матрица плотности выше выражается в зависящем от времени базисе $\begin{pmatrix} |\alpha_1 e^{-\Gamma t/2}\rangle \\ |\alpha_2 e^{-\Gamma t/2}\rangle \end{pmatrix}$, а углы поворота $\theta_{21}(t)$ определяются как $\theta_{21}(t) = \mathrm{Im}\,(\alpha_2^*\alpha_1)(1-e^{-\Gamma t})$.

Если мы рассматриваем затухание только недиагональных элементов, то можно игнорировать фазу $\theta_{21}(t)$. Для времен $t \ll 1/\Gamma$ базисные состоя-

ния остаются приблизительно такими же:

$$|\alpha_1(t)\rangle \simeq \left|\alpha_1\left(1-\frac{\Gamma t}{2}\right)\right\rangle \simeq |\alpha_1\rangle,$$

а амплитуды недиагональных элементов экспоненциально затухают со временем:

$$e^{-\frac{1}{2}|\alpha_1-\alpha_2|^2(1-e^{-\Gamma t})} \simeq e^{-\frac{1}{2}|\alpha_1-\alpha_2|^2(1-(1-\Gamma t))} \simeq e^{-\frac{\Gamma t}{2}|\alpha_1-\alpha_2|^2}.$$

Решения упражнений к главе 4

4.1. Теорема Харди

Боб и Клер делят множество идентично приготовленных копий состояния

$$|\psi\rangle_{BC} = \sqrt{1-2x}|0\rangle_B \otimes |0\rangle_C + \sqrt{x}|0\rangle_B \otimes |1\rangle_C + \sqrt{x}|1\rangle_B \otimes |0\rangle_C,$$

где x — вещественное число из интервала $[0, 1/2]$.

а) Если Боб выполняет измерение в базисе $\{|0\rangle, |1\rangle\}$, а Клер — в базисе $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, то всякий раз, когда результатом Боба является $|0\rangle$, Клер получает $|\varphi\rangle$. Вследствие симметрии подсистем в $|\psi\rangle_{BC}$, такая же картина будет наблюдаться, если Боб и Клер обменяются базисами.

Тогда мы можем спроецировать разделенное состояние на $|0\rangle$ в одной подсистеме, чтобы найти $|\varphi\rangle$ в другой подсистеме:

$$\begin{aligned}(|0\rangle_B\,{}_B\langle 0| \otimes \mathbf{1}_C)\,|\psi\rangle_{BC} &= \sqrt{1-2x}|0\rangle_B \otimes |0\rangle_C + \sqrt{x}|0\rangle_B \otimes |1\rangle_C = \\ &= |0\rangle_B \otimes \left(\sqrt{1-2x}|0\rangle_C + \sqrt{x}|1\rangle_C\right) \\ (|0\rangle_B\,{}_B\langle 0| \otimes \mathbf{1}_C)\,|\psi\rangle_{BC} &= |0\rangle_B \otimes (\mathcal{N}|\varphi\rangle_C).\end{aligned}$$

Нормируя проекцию, мы находим

$$|\varphi\rangle = \sqrt{\frac{1-2x}{1-x}}|0\rangle + \sqrt{\frac{x}{1-x}}|1\rangle.$$

Рассмотрим некоторое нормированное состояние $|\chi\rangle = a|0\rangle + b|1\rangle$, где $a, b \in \mathbb{C}$. Поскольку

$$\langle\chi^\perp|\chi\rangle = (b\langle 0| - a\langle 1|)\,(a|0\rangle + b|1\rangle) = 0,$$

то ортогональное ему нормированное состояние равно $|\chi^{\perp}\rangle = b^*|0\rangle - a^*|1\rangle$. Следовательно, учитывая, что при $x \in [0, 1/2]$ коэффициенты вещественны, мы можем определить

$$|\varphi^{\perp}\rangle = \sqrt{\frac{x}{1-x}}|0\rangle - \sqrt{\frac{1-2x}{1-x}}|1\rangle,$$

b) Боб и Клер оба выбрали $\{|\varphi\rangle, |\varphi^{\perp}\rangle\}$ в качестве базиса измерения. Вероятность $P(x)$ того, что результатом обоих измерений является $|\varphi^{\perp}\rangle$, вычисляется как квадрат соответствующей амплитуды состояния $|\psi\rangle_{BC}$:

$$P(x) = \left|\left({}_B\langle\varphi^{\perp}| \otimes {}_C\langle\varphi^{\perp}|\right)|\psi\rangle_{BC}\right|^2.$$

Мы можем вычислить

$$\begin{aligned}
{}_B\langle\varphi^{\perp}| \otimes {}_C\langle\varphi^{\perp}| &= \left(\sqrt{\frac{x}{1-x}}\,{}_B\langle 0| - \sqrt{\frac{1-2x}{1-x}}\,{}_B\langle 1|\right) \otimes \\
&\quad \left(\sqrt{\frac{x}{1-x}}\,{}_C\langle 0| - \sqrt{\frac{1-2x}{1-x}}\,{}_C\langle 1|\right), \\
{}_B\langle\varphi^{\perp}| \otimes {}_C\langle\varphi^{\perp}| &= \frac{x}{1-x}\,{}_B\langle 0| \otimes {}_C\langle 0| + \frac{1-2x}{1-x}\,{}_B\langle 1| \otimes {}_C\langle 1| - \\
&\quad - \frac{\sqrt{x(1-2x)}}{1-x}\left({}_B\langle 0| \otimes {}_C\langle 1| + {}_B\langle 1| \otimes {}_C\langle 0|\right).
\end{aligned}$$

Подстановка разложений ${}_B\langle\varphi^{\perp}| \otimes {}_C\langle\varphi^{\perp}|$ и $|\psi\rangle_{BC}$ в базисе $\{|0\rangle, |1\rangle\}$ в определение $P(x)$ дает следующий результат:

$$\begin{aligned}
P(x) &= \left|\frac{x}{1-x}\sqrt{1-2x} + \frac{1-2x}{1-x}\cdot(0) - \frac{\sqrt{x(1-2x)}}{1-x}(\sqrt{x}+\sqrt{x})\right|^2 = \\
&= \left|\frac{x\sqrt{1-2x}}{1-x} - \frac{2x\sqrt{1-2x}}{1-x}\right|^2 = \\
&= \left|-\frac{x\sqrt{1-2x}}{1-x}\right|^2, \\
P(x) &= \frac{x^2(1-2x)}{(1-x)^2}.
\end{aligned}$$

c) Заметим, что $P(x) \geqslant 0$ при $x \in [0, 1/2]$ и $P(0) = P(1/2) = 0$. Так как $P(x)$ непрерывна в этом интервале, ее максимум достигается в некоторой внутренней критической точке, удовлетворяющей условию $\dfrac{dP(x)}{dx} = 0$ (или ∞):

$$\frac{dP(x)}{dx} = \frac{d}{dx}\left[\frac{x^2(1-2x}{(1-x)^2}\right] = \frac{2x}{(1-x)^3}\Big[1 - 4x + 3x^2 + x - 2x^2\Big],$$

$$\frac{dP(x)}{dx} = \frac{2x}{(1-x)^3}\big(x^2 - 3x + 1\big).$$

Корнями приведенного выше квадратного трехчлена являются $x = \frac{1}{2}(3 \pm \sqrt{5})$. Внутри $0 < x < 1/2$ лежит только одна критическая точка $P(x)$: $x = \frac{1}{2}(3 - \sqrt{5})$. Подставляя ее значение в выражение для $P(x)$, находим $P_{\max} = \frac{1}{2}(5\sqrt{5} - 11)$.

d) Если $P(x)$ не равна нулю (что соответствует $0 < x < 1/2$), то существует измеримое нарушение предсказания Альберта (и теоремы Харди). Рассуждения Альберта некорректны на этапе комбинирования результатов двух взаимно исключающих экспериментов. Наблюдаемые, соответствующие измерениям в различных базисах, не коммутируют между собой, то есть не имеет смысла рассматривать систему общих собственных состояний обеих наблюдаемых. Коль скоро Боб (или Клер) выбирает измерение в базисе $\{|0\rangle, |1\rangle\}$, мы никогда не сможем (с определенностью) узнать, каким был бы результат измерения в базисе $\{|\varphi\rangle, |\varphi^{\perp}\rangle\}$.

Заметим, что одного запутывания недостаточно, чтобы опровергнуть рассуждения Альберта. До тех пор, пока Альберт рассматривает только коммутирующие наблюдаемые, он может построить теорию скрытых переменных, чтобы объяснить корреляции результатов измерений. Например, в случаях $x = 0$ и $x = 1/2$ оба базиса становятся идентичными и, как предсказывал Альберт, $P(x) = 0$.[1]

4.2. Закрытие лазейки детектирования

a) Выберем переменные $x, x', y, y' \in \{0, 1\}$. Мы хотим показать, что

$$xy + xy' + x'y - x'y' \leqslant x + y, \quad \forall\, x, x', y, y'.$$

[1] Заметим, что состояние $|\psi\rangle_{BC}$, факторизуемое при $x = 0$, остается запутанным при $x = 1/2$. – *Прим. ред.*

Конечно, можно перебрать все 16 возможностей. С другой стороны, мы могли бы воспользоваться неравенством КГШХ (доказанным на лекциях), определив переменные $\alpha = 2x - 1$, $\alpha' = 2x' - 1$, $\beta = 2y - 1$, $\beta' = 2y' - 1$. Заметим, что $\alpha, \alpha', \beta, \beta' \in \{-1, 1\}$. Применим неравенство КГШХ:

$$
\begin{aligned}
2 &\geqslant \alpha\beta + \alpha\beta' + \alpha'\beta - \alpha'\beta', \\
2 &\geqslant (2x-1)(2y-1) + (2x-1)(2y'-1) + \\
&\quad + (2x'-1)(2y-1) - (2x'-1)(2y'-1), \\
2 &\geqslant 4xy - 2x - 2y + 1 + 4xy' - 2x - 2y' + 1 + \\
&\quad + 4x'y - 2x' - 2y + 1 - 4x'y' + 2x' + 2y' - 1, \\
2 &\geqslant 4xy + 4xy' + 4x'y - 4x'y' - 4x - 4y + 2, \\
2 &\geqslant 4\Big[(xy + xy' + x'y - x'y') - (x + y)\Big] + 2, \\
0 &\geqslant (xy + xy' + x'y - x'y') - (x + y).
\end{aligned}
$$

Следовательно, $xy + xy' + x'y - x'y' \leqslant x + y \ \ \forall x, x', y, y' \in \{0, 1\}$.

b) Предположим, что существует локальная теория скрытых переменных, описывающая результаты выполняемых Алисой и Бобом измерений N фотонных пар. Пусть переменные $x_i, x_i', y_i, y_i' \in \{0, 1\}$ обозначают результаты регистрации фотонов i-й пары. А именно, $x_i, x_i' \in \{0, 1\}$ обозначают, сработал или нет детектор Алисы, ориентированный вдоль оси $\hat{a}$ или $\hat{a}'$ соответственно. Аналогично переменные $y_i, y_i' \in \{0, 1\}$ обозначают, сработал или нет детектор Боба, ориентированный вдоль оси $\hat{b}$ или $\hat{b}'$ соответственно. Каждый набор переменных x, x', y, y' должен удовлетворять доказанному в части (а) соотношению

$$x_i y_i + x_i y_i' + x_i' y_i - x_i' y_i' \leqslant x_i + y_i.$$

Сложим N неравенств, чтобы получить

$$
\begin{gathered}
\sum_{i=1}^{N} (x_i y_i + x_i y_i' + x_i' y_i - x_i' y_i') \leqslant \sum_{i=1}^{N} (x_i + y_i), \\
\sum_{i=1}^{N} x_i y_i + \sum_{i=1}^{N} x_i y_i' + \sum_{i=1}^{N} x_i' y_i - \sum_{i=1}^{N} x_i' y_i' \leqslant \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} y_i.
\end{gathered}
$$

Деление обеих частей на положительное целое число не меняет неравенство, поэтому

$$\frac{1}{N}\sum_{i=1}^{N} x_i y_i + \frac{1}{N}\sum_{i=1}^{N} x_i y_i' + \frac{1}{N}\sum_{i=1}^{N} x_i' y_i - \frac{1}{N}\sum_{i=1}^{N} x_i' y_i' \leqslant \frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N}\sum_{i=1}^{N} y_i.$$

Выражения в левой части этого неравенства дают оценки вероятностей того, что Алиса и Боб одновременно детектируют отдельную фотонную пару (при определенном расположении их детекторов). Выражения в правой части дают оценки вероятностей того, что Алиса или Боб независимо детектируют фотон детекторами, ориентированными вдоль осей $\hat{a}$ и $\hat{b}$ соответственно. Пусть N настолько велико, что эти оценки достаточно близки к соответствующим вероятностям скрытых переменных. Тогда мы приходим к выводу, что

$$P_{++}(ab) + P_{++}(ab') + P_{++}(a'b) - P_{++}(a'b') \leqslant P_{+\cdot}(a) + P_{\cdot+}(b).$$

c) Заметим, что базис состояний Белла $\{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\}$ удовлетворяет

$$\begin{aligned}
\boldsymbol{\sigma}_X \otimes \mathbf{1}|\Phi^+\rangle &= |\Psi^+\rangle = \mathbf{1} \otimes \boldsymbol{\sigma}_X|\Phi^+\rangle, \\
\boldsymbol{\sigma}_Y \otimes \mathbf{1}|\Phi^+\rangle &= i|\Psi^-\rangle = -\mathbf{1} \otimes \boldsymbol{\sigma}_Y|\Phi^+\rangle, \\
\boldsymbol{\sigma}_Z \otimes \mathbf{1}|\Phi^+\rangle &= |\Phi^-\rangle = \mathbf{1} \otimes \boldsymbol{\sigma}_Z|\Phi^+\rangle.
\end{aligned}$$

Таким образом, $\langle\Phi^+|\hat{a}\cdot\vec{\boldsymbol{\sigma}} \otimes \mathbf{1}|\Phi^+\rangle = 0 = \langle\Phi^+|\mathbf{1} \otimes \hat{b}\cdot\vec{\boldsymbol{\sigma}}|\Phi^+\rangle$ при любых $\hat{a}$, $\hat{b}$.

При данном единичном 3-векторе $\hat{a}$ оператор $\hat{a}\cdot\vec{\boldsymbol{\sigma}}$ имеет собственные значения $\{-1, +1\}$, соответствующие собственным состояниям кубита, ориентированным или антипараллельно, или параллельно оси $\hat{a}$. Оператор $\frac{1}{2}(\mathbf{1} + \hat{a}\cdot\vec{\boldsymbol{\sigma}})$ имеет те же собственные состояния, но с собственными значениями $\{0, +1\}$. Математическое ожидание этого последнего оператора в точности совпадает с вероятностью срабатывания ориентированного вдоль оси $\hat{a}$ детектора при условии его идеальной эффективности.

Тогда мы можем выразить вероятность одновременного детектирования Алисой и Бобом фотонов из разделенного состояния $|\Phi^+\rangle$ при эффективностях детекторов η_A, η_B:

$$\begin{aligned}
P_{++}(ab) &= \left\langle \Phi^+ \left| \frac{\eta_A}{2}\left(\mathbf{1} + \hat{a}\cdot\vec{\boldsymbol{\sigma}}\right) \otimes \frac{\eta_B}{2}\left(\mathbf{1} + \hat{b}\cdot\vec{\boldsymbol{\sigma}}\right) \right| \Phi^+ \right\rangle = \\
&= \frac{\eta_A\eta_B}{4}\left\langle \Phi^+ \left| \mathbf{1}\otimes\mathbf{1} + \hat{a}\cdot\vec{\boldsymbol{\sigma}}\otimes\mathbf{1} + \mathbf{1}\otimes\hat{b}\cdot\vec{\boldsymbol{\sigma}} + (\hat{a}\cdot\vec{\boldsymbol{\sigma}})\otimes(\hat{b}\cdot\vec{\boldsymbol{\sigma}}) \right| \Phi^+ \right\rangle = \\
&= \frac{\eta_A\eta_B}{4}\left[1 + 0 + 0 + \hat{a}\cdot\hat{b}\right], \\
P_{++}(ab) &= \frac{1}{4}\eta_A\eta_B(1 + \hat{a}\cdot\hat{b}).
\end{aligned}$$

Аналогичным образом находим

$$P_{++}(ab') = \frac{1}{4}\eta_A\eta_B(1 + \hat{a}\cdot\hat{b}'),$$

$$P_{++}(a'b) = \frac{1}{4}\eta_A\eta_B(1 + \hat{a}'\cdot\hat{b}),$$

$$P_{++}(a'b') = \frac{1}{4}\eta_A\eta_B(1 + \hat{a}'\cdot\hat{b}').$$

Также можно вычислить вероятности независимого детектирования:

$$P_{+\cdot}(a) = \left\langle \Phi^+ \left| \frac{\eta_A}{2}\left(\mathbf{1} + \hat{a}\cdot\vec{\boldsymbol{\sigma}}\right)\otimes\mathbf{1} \right| \Phi^+ \right\rangle = \frac{\eta_A}{2},$$

$$P_{\cdot+}(b) = \left\langle \Phi^+ \left| \mathbf{1}\otimes\frac{\eta_B}{2}\left(\mathbf{1} + \hat{b}\cdot\vec{\boldsymbol{\sigma}}\right) \right| \Phi^+ \right\rangle = \frac{\eta_B}{2}.$$

Чтобы максимально нарушить неравенство КГШХ, следует выбрать $\hat{a} = \hat{x}$, $\hat{a}' = \hat{z}$, $\hat{b} = \frac{1}{\sqrt{2}}(\hat{x} + \hat{z})$ и $\hat{b}' = \frac{1}{\sqrt{2}}(\hat{x} - \hat{z})$. Подставим их в найденные выше выражения:

$$P_{++}(ab) = \frac{1}{4}\eta_A\eta_B\left(1 + \hat{x}\cdot\frac{1}{\sqrt{2}}(\hat{x} + \hat{z})\right) = \frac{1}{4}\eta_A\eta_B\left(1 + \frac{1}{\sqrt{2}}\right),$$

$$P_{++}(ab') = \frac{1}{4}\eta_A\eta_B\left(1 + \hat{x}\cdot\frac{1}{\sqrt{2}}(\hat{x} - \hat{z})\right) = \frac{1}{4}\eta_A\eta_B\left(1 + \frac{1}{\sqrt{2}}\right),$$

$$P_{++}(a'b) = \frac{1}{4}\eta_A\eta_B\left(1 + \hat{z}\cdot\frac{1}{\sqrt{2}}(\hat{x} + \hat{z})\right) = \frac{1}{4}\eta_A\eta_B\left(1 + \frac{1}{\sqrt{2}}\right),$$

$$P_{++}(a'b') = \frac{1}{4}\eta_A\eta_B\left(1 + \hat{z}\cdot\frac{1}{\sqrt{2}}(\hat{x} - \hat{z})\right) = \frac{1}{4}\eta_A\eta_B\left(1 - \frac{1}{\sqrt{2}}\right).$$

Комбинируя эти вероятности, получим

$$P_{++}(ab) + P_{++}(ab') + P_{++}(a'b) - P_{++}(a'b') = \frac{1}{4}\eta_A\eta_B\left(2 + \frac{4}{\sqrt{2}}\right),$$

$$P_{++}(ab) + P_{++}(ab') + P_{++}(a'b) - P_{++}(a'b') = \frac{1}{2}\eta_A\eta_B(1 + \sqrt{2}).$$

Максимально запутанное состояние $|\Phi^+\rangle$ может нарушить выведенное в части (b) неравенство для локальных скрытых переменных, если

$$P_{++}(ab) + P_{++}(ab') + P_{++}(a'b) - P_{++}(a'b') \geqslant P_{+\cdot}(a) + P_{\cdot+}(b),$$

$$\frac{1}{2}\eta_A\eta_B(1+\sqrt{2}) > \frac{\eta_A}{2}+\frac{\eta_B}{2},$$

$$\frac{\eta_A\eta_B}{\eta_A+\eta_B} > \frac{1}{1+\sqrt{2}}.$$

4.3. Телепортация с помощью непрерывных переменных

a) Проверим сформулированное утверждение, выражая запутанные состояния в базисе $\{|q_1\rangle \otimes |q_2\rangle\}$:

$$\begin{aligned}\langle Q', P'|Q, P\rangle &= \frac{1}{2\pi}\int dq\, dq' e^{i(Pq-P'q')}\langle q'|q\rangle\langle q'+Q'|q+Q\rangle = \\ &= \frac{1}{2\pi}\int dq e^{i(P-P')q}\delta(Q-Q') = \\ &= \delta(Q-Q')\delta(P-P').\end{aligned}$$

b) Чтобы найти коэффициенты, вновь разложим в базисе $\{|q_1\rangle \otimes |q_2\rangle\}$:

$$\begin{aligned}\langle Q', P'|q_1, q_2\rangle &= \frac{1}{\sqrt{2\pi}}\int dq e^{-iPq}\langle q|q_1\rangle\langle q+Q|q_2\rangle = \\ &= \frac{1}{\sqrt{2\pi}}\int dq e^{-iPq}\delta(q-q_1)\delta(q+Q-q_2) = \\ &= \frac{1}{\sqrt{2\pi}}e^{-iPq_1}\,\delta[Q-(q_2-q_1)].\end{aligned}$$

c) В этой части нам нужно оставить неизменными переменные p и q. Как и в уравнении (4.84) в лекциях, мы хотим выразить состояние системы AC в запутанном базисе. В этом базисе Алиса будет выполнять измерения, посылая их результаты Бобу. Тогда, используя эти результаты, Боб сможет реконструировать в своей лаборатории исходное состояние системы C. Это – грубое описание того, как должна работать телепортация; после выполнения некоторых предварительных вычислений я представлю полную процедуру, которой должны следовать Алиса и Боб.

В качестве первого шага представим систему AC в запутанном базисе. Предварительно записав $|\psi\rangle_{ABC}$ в базисе

$$\{|q_1\rangle_A \otimes |q_2\rangle_B \otimes |q_3\rangle_C\},$$

сделаем это, используя тождественную вставку

$$1 = \int dQ'\, dP' |Q', P'\rangle_{CA\,CA}\langle Q', P'|.$$

Получающиеся при этом коэффициенты ${}_{CA}\langle Q', P'|q_1, q_2\rangle_{CA}$, которые уже были вычислены в части (b), позволяют перевыразить состояние в запутанном базисе системы AC:

$$
\begin{aligned}
|\psi\rangle_C|Q, P\rangle_{AB} &= \int dq\, {}_C\langle q|\psi\rangle_C|q\rangle_C \times \frac{1}{\sqrt{2\pi}} \int dq'\, e^{iPq'}|q'\rangle_A|q'+Q\rangle_B = \\
&= \frac{1}{\sqrt{2\pi}} \int dq\, dq'\, dQ'\, dP'\, {}_C\langle q|\psi\rangle_C e^{iPq'} \times \\
&\quad \times |Q', P'\rangle_{CA}\, {}_{CA}\langle Q', P'|q, q'\rangle_{CA} \otimes |q'+Q\rangle_B = \\
&= \frac{1}{2\pi} \int dq\, dq'\, dQ'\, dP'\, {}_C\langle q|\psi\rangle_C e^{iPq'} \times \\
&\quad \times e^{-iP'q}\delta[Q' - (q'-q)]|Q', P'\rangle_{CA} \otimes |q'+Q\rangle_B = \\
&= \frac{1}{2\pi} \int dq\, dQ'\, dP'\, {}_C\langle q|\psi\rangle_C e^{iP(Q'+q)-iP'q} \times \\
&\quad \times |Q', P'\rangle_{CA} \otimes |q+Q'+Q\rangle_B.
\end{aligned}
$$

С этого момента Алиса выполняет измерение в запутанном базисе AC, получая некоторое состояние $|Q', P'\rangle_{CA}$. Результирующим состоянием Боба является

$$
|\mathrm{Bob}\rangle = \int dq\, {}_C\langle q|\psi\rangle_C e^{iPQ'-i(P-P')q}|q+Q'+Q\rangle_B.
$$

Боб имеет почти все, что ему необходимо. Если Алиса посылает ему результаты своего измерения (Q', P'), то он может применить параллельные переносы координаты и импульса:

$$
\begin{aligned}
\mathbf{D}(q) &= e^{iq\mathbf{P}} = \int dq'\, |q'+q\rangle\langle q'|, \\
\mathbf{D}(p) &= e^{-ip\mathbf{Q}} = \int dq'\, e^{-ipq'}|q'\rangle\langle q'|,
\end{aligned}
$$

чтобы преобразовать свое состояние $|\mathrm{Bob}\rangle$ в то, в какое ему нужно. Проверяя состояние Боба, мы видим, что он должен применить к нему $\mathbf{D}(-Q'-Q)$ и $\mathbf{D}(P-P')$. Однако выполнение этих сдвигов[1] оставляет состояние Боба с общей фазой $e^{iPQ'}$. Конечно, она не имеет физического значения,

[1] Именно в этом порядке, сначала $\mathbf{D}(-Q'-Q)$, а затем $\mathbf{D}(P-P')$. *Прим. ред.*

но если угодно, то можно избавиться и от нее, применяя операторы сдвига в специальном порядке. Сначала заметим, что

$$\begin{aligned}\mathbf{D}(p)\mathbf{D}(q) &= \int dq''\,dq'\,e^{-ipq''}|q''\rangle\langle q''|q'+q\rangle\langle q'| = \\ &= \int dq'\,e^{-ipq'}e^{-ipq}|q'+q\rangle\langle q'| = \\ &= e^{-ipq}\int dq'\,dq''\,e^{-ipq''}|q'+q\rangle\langle q'|q''\rangle\langle q''| = \\ &= e^{-ipq}\mathbf{D}(q)\mathbf{D}(p).\end{aligned}$$

Используя этот результат, мы видим, что применение к состоянию Боба $|\mathrm{Bob}\rangle$ оператора

$$\begin{aligned}U &= \mathbf{D}(-P')\mathbf{D}(-Q')\mathbf{D}(P)\mathbf{D}(-Q) = \\ &= e^{-iPQ'}\mathbf{D}(-P')\mathbf{D}(P)\mathbf{D}(-Q')\mathbf{D}(-Q) = \\ &= e^{-iPQ'}\mathbf{D}(P-P')\mathbf{D}(-Q'-Q)\end{aligned}$$

восстанавливает состояние $|\psi\rangle$:

$$\begin{aligned}U|\mathrm{Bob}\rangle &= \int dq\,{}_C\langle q|\psi\rangle_C\,e^{iPQ'+i(P-P')q}e^{-iPQ'}\times \\ &\quad\times\mathbf{D}(P-P')\mathbf{D}(-Q'-Q)|q+Q'+Q\rangle_B = \\ &= \int dq\,{}_C\langle q|\psi\rangle_C\,e^{i(P-P')q}\mathbf{D}(P-P')|q\rangle_B = \\ &= \int dq\,{}_C\langle q|\psi\rangle_C\,e^{i(P-P')q}e^{-i(P-P')q}|q\rangle_B = \\ &= \int dq\,{}_C\langle q|\psi\rangle_C|q\rangle_B = \\ &= |\psi\rangle_B.\end{aligned}$$

Итак, протокол, которому должны следовать Алиса и Боб для телепортации с помощью непрерывных переменных, выглядит следующим образом:

1) Готовится запутанное состояние $|Q,P\rangle_{AC}$.

2) Алиса измеряет (Q',P') в запутанном базисе системы AC.

3) Алиса посылает Бобу результаты своего измерения (Q', P').

4) Боб применяет оператор $\mathbf{D}(-P')\mathbf{D}(-Q')\mathbf{D}(P)\mathbf{D}(-Q)$ к своему состоянию. В итоге он имеет состояние $|\psi\rangle_B$.

4.4. Телепортация со смешанными состояниями

a) Мы знаем, что если Алиса и Боб поделили синглет, то они могут осуществить телепортацию с идеальной точностью воспроизведения. Если вместо этого Алиса и Боб нечаянно разделили смешанное состояние, то выполняемое Алисой измерение Белла ничего не говорит об ее состоянии (следовательно, у нее нет классической информации, которую необходимо послать Бобу), а состояние Боба никак не коррелирует с состоянием Алисы. В этом случае лучшая стратегиея Боба состоит в угадывании, которое, как мы показали, имеет точность воспроизведения $1/2$. Так как данная в задаче матрица плотности может рассматриваться как *ансамбль* этих альтернатив, имеющих вероятности $(1-\lambda)$ и λ соответственно, то полная точность воспроизведения телепортации с помощью этого состояния равна

$$F = 1 \cdot (1-\lambda) + \frac{1}{2} \cdot \lambda = 1 - \frac{\lambda}{2}.$$

b) Эта точность воспроизведения больше, чем $2/3$, при $\lambda < 2/3$.

c) Очень похожие спин-спиновые корреляции рассматривались в задаче 2.5. Вырезая и склеивая ее фрагменты, я воспроизведу здесь (с небольшим изменением) решение. (Более детальное изложение смотрите в решении задачи 2.5).

$$\begin{aligned}
p &= \mathrm{tr}_B\,\mathrm{tr}_A\left[\left(\frac{1}{2}(\mathbf{1}_A + \hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes\frac{1}{2}(\mathbf{1}_B + \hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)\right)\times\right.\\
&\qquad\left.\times\left(\frac{\lambda}{4}\mathbf{1}_{AB} + (1-\lambda)|\psi^-\rangle\langle\psi^-|\right)\right] =\\
&= \frac{\lambda}{16}\,\mathrm{tr}_B\,\mathrm{tr}_A\left[(\mathbf{1}_A + \hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes(\mathbf{1}_B + \hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)\right] +\\
&\qquad + \frac{1-\lambda}{4}\,\mathrm{tr}_B\,\mathrm{tr}_A\left[(\mathbf{1}_A + \hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes(\mathbf{1}_B + \hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)|\psi^-\rangle\langle\psi^-|\right] =\\
&= \frac{\lambda}{16}\,\mathrm{tr}_B\,\mathrm{tr}_A\left[\mathbf{1}_A\otimes\mathbf{1}_B\right] + \frac{1-\lambda}{4}\langle\psi^-|(\mathbf{1}_A + \hat{n}\cdot\vec{\boldsymbol{\sigma}}_A)\otimes(\mathbf{1}_B + \hat{m}\cdot\vec{\boldsymbol{\sigma}}_B)|\psi^-\rangle =
\end{aligned}$$

$$= \frac{\lambda}{4} + \frac{1-\lambda}{4} + \frac{1-\lambda}{4}\langle\psi^-|\hat{n}\cdot\vec{\sigma}_A + \hat{m}\cdot\vec{\sigma}_B + \hat{n}\cdot\vec{\sigma}_A \otimes \hat{m}\cdot\vec{\sigma}_B|\psi^-\rangle =$$

$$= \frac{1}{4} + \frac{1-\lambda}{4}\Big[\hat{n}\cdot\langle\psi^-|\vec{\sigma}_A|\psi^-\rangle + \hat{m}\cdot\langle\psi^-|\vec{\sigma}_B|\psi^-\rangle +$$

$$+\langle\psi^-|\hat{n}\cdot\vec{\sigma}_A \otimes \hat{m}\cdot\vec{\sigma}_B|\psi^-\rangle\Big] =$$

$$= \frac{1}{4} + \frac{1-\lambda}{4}\langle\psi^-|\hat{n}\cdot\vec{\sigma}_A \otimes \hat{m}\cdot\vec{\sigma}_B|\psi^-\rangle =$$

$$= \frac{1}{4} - \frac{1-\lambda}{4}\cos\theta.$$

d) При $\lambda = 1/2$ вероятность того, что спины Алисы и Боба коррелированы, равна $p = \frac{1}{4} - \frac{1}{8}\hat{n}\cdot\hat{m}$. Очень естественным предположением относительно порождающего эту корреляцию распределения вероятностей скрытых переменных выглядят

$$f_A(\hat{\alpha}\cdot\hat{n}) = \frac{1}{2} + a(\hat{\alpha}\cdot\hat{n}),$$

$$f_B(\hat{\alpha}\cdot\hat{n}) = \frac{1}{2} + b(\hat{\alpha}\cdot\hat{m}).$$

Этот вид подсказывается взаимно-однозначным соответствием между проекторами на сфере Блоха и единичными векторами на S^2. Он автоматически порождает индивидуальные распределения наблюдаемых Алисы и Боба. Для того чтобы воспроизводить квантово-механические спин-спиновые корреляции между Алисой и Бобом, a и b должны удовлетворять условию

$$\int\limits_{\hat{\alpha}} f_A(\hat{\alpha}\cdot\hat{n})f_B(\hat{\alpha}\cdot\hat{m}) = \frac{1}{4} + \frac{ab}{4\pi}\int(\hat{\alpha}\cdot\hat{n})(\hat{\alpha}\cdot\hat{m})d\Omega =$$

$$= \frac{1}{4} + \frac{ab}{4\pi}\left(\frac{4\pi}{3}\right)\hat{n}\cdot\hat{m} =$$

$$= \frac{1}{4} + \frac{1}{3}ab\cos\theta \Rightarrow$$

$$\Rightarrow ab = -\frac{3}{8}.$$

Для того чтобы f_A и f_B действительно были распределениями вероятностей (то есть принимали значения в $[0, 1]$), должны выполняться неравенства $|a| \leqslant 1/2$ и $|b| \leqslant 1/2$. Но, согласно неравенству Шварца, это означает,

что $|ab| \leqslant |a| \cdot |b| \leqslant 1/4 < 3/8$ (!). Таким образом, этой простой модели не достаточно — квантовые корреляции слишком сильны, чтобы моделироваться наивной теорией скрытых переменных.

Чтобы добиться сильных корреляций, рассмотрим *разрывные* функции распределения вероятностей

$$f_A(\hat{\alpha}\cdot\hat{n}) = \frac{1}{2} + a\,\mathrm{sign}(\hat{\alpha}\cdot\hat{n}),$$

$$f_B(\hat{\alpha}\cdot\hat{m}) = \frac{1}{2} + b(\hat{\alpha}\cdot\hat{n}).$$

Очевидно, что эта новая теория по-прежнему воспроизводит индивидуальные распределения Алисы и Боба с $\langle f_A\rangle = \langle f_B\rangle = \frac{1}{2}$. Чтобы вычислить их интеграл спин-спиновых корреляций, запишем $\hat{\alpha}$, $\hat{n}$ и $\hat{m}$ в конкретном базисе:

$$\hat{\alpha} = \hat{x}\cos\varphi\sin\theta + \hat{y}\sin\varphi\sin\theta + \hat{z}\cos\theta,$$
$$\hat{n} = \hat{z},$$
$$\hat{m} = \hat{x}\sin\psi + \hat{z}\cos\psi.$$

В этом базисе корреляционный интеграл имеет вид

$$\int\limits_{\hat{\alpha}} f_A(\hat{\alpha}\cdot\hat{n})f_B(\hat{\alpha}\cdot\hat{m}) = \frac{1}{4} + \frac{ab}{4\pi}\int(\hat{\alpha}\cdot\hat{m})\,\mathrm{sign}(\hat{\alpha}\cdot\hat{n})d\Omega =$$

$$= \frac{1}{4} + \frac{ab}{4\pi}\int\limits_0^{2\pi} d\varphi \int\limits_1^1 d(\cos\theta)(\cos\varphi\sin\theta\sin\psi + \cos\theta\cos\psi)\mathrm{sign}(\cos\theta) =$$

$$= \frac{1}{4} + \frac{ab}{2}\cos\psi\left[\int\limits_0^1 d(\cos\theta)\cos\theta - \int\limits_{-1}^0 d(\cos\theta)\cos\theta\right] =$$

$$= \frac{1}{4} + \frac{ab}{2}\cos\psi\left[\frac{1}{2} + \frac{1}{2}\right] = \frac{1}{4} + \frac{ab}{2}\cos\psi.$$

Это соответствует предсказанию квантовой механики при $ab = -1/4$, что выполняется, например, при $a = 1/2$, $b = -1/2$:

$$f_A(\hat{\alpha}\cdot\hat{n}) = \begin{cases} 1, & \hat{\alpha}\cdot\hat{n} \geqslant 0, \\ 0, & \hat{\alpha}\cdot\hat{n} < 0, \end{cases}$$

$$f_B(\hat{\alpha}\cdot\hat{m}) = \frac{1}{2}(1 - \hat{\alpha}\cdot\hat{m}).$$

4.5. Распределение квантовых ключей

а) Решение первой части этой задачи в целом совпадает с решением задачи следующей главы 5.2(b). Как там показано, ограничение собственных чисел оператора $\mathbf{F}_{DK}$ неотрицательными значениями накладывает верхнюю границу на возможные значения A. Я приведу здесь доказательство:

$$\begin{aligned}\mathbf{F}_{DK} &= \begin{pmatrix} 1-2A & 0 \\ 0 & 1-2A \end{pmatrix} \\ &+ A\begin{pmatrix} \cos^2\alpha & \cos\alpha\sin\alpha + \\ \cos\alpha\sin\alpha & \sin^2\alpha \end{pmatrix} + \\ &+ A\begin{pmatrix} \sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & \cos^2\alpha \end{pmatrix} = \\ &= \begin{pmatrix} 1-A & 2A\cos\alpha\sin\alpha \\ 2A\cos\alpha\sin\alpha & 1-A \end{pmatrix}.\end{aligned}$$

Характеристическое уравнение имеет вид:

$$\begin{aligned}0 &= \lambda^2 - \lambda\,\mathrm{tr}\,\mathbf{F}_{DK} + \det\mathbf{F}_{DK} = \\ &= \lambda^2 - 2(1-A)\lambda + (1-A)^2 - 4A^2\cos^2\alpha\sin^2\alpha \\ \lambda &= 1 - A \pm \sqrt{(1-A)^2 - (1-A)^2 + A^2\sin^2 2\alpha} = \\ &= 1 - A \pm A\sin 2\alpha.\end{aligned}$$

Из условия неотрицательности собственных чисел $\lambda \geqslant 0$ следует неравенство:

$$A \leqslant \frac{1}{1 \pm \sin 2\alpha}.$$

Поскольку $0 < \alpha < \pi/4$, ограничение положительности может быть переписано как

$$0 \leqslant A \leqslant \frac{1}{1 + \sin 2\alpha}.$$

Если Алиса делает равновероятный выбор из $\{|u\rangle, |v\rangle\}$, то матрица плотности Боба выглядит как $\boldsymbol{\rho} = \frac{1}{2}(|u\rangle\langle u| + |v\rangle\langle v|)$. Следовательно, вероятность получения результата DK:

$$\begin{aligned}p_{DK} &= \mathrm{tr}(\boldsymbol{\rho}\mathbf{F}_{DK}) = \\ &= \mathrm{tr}\left[\frac{1}{2}\begin{pmatrix} 1 & \sin 2\alpha \\ \sin 2\alpha & 1 \end{pmatrix}\begin{pmatrix} 1-A & A\sin 2\alpha \\ A\sin 2\alpha & 1-A \end{pmatrix}\right] =\end{aligned}$$

$$= 1 - A + A\sin^2 2\alpha =$$
$$= 1 + A(\sin^2 2\alpha - 1).$$

Для того чтобы минимизировать p_{DK}, мы должны выбрать *максимально возможное* значение A, а именно: $A = 1/(1 + \sin 2\alpha)$. Тогда вероятность того, что Боб не знает, что послала Алиса:

$$p_{DK} = 1 + \frac{\sin^2 2\alpha - 1}{1 + \sin 2\alpha} = \sin 2\alpha.$$

b) Наиболее естественный способ построения распределения квантовых ключей вокруг источника Алисы и ПОЗМ Боба представляет собой небольшую модификацию схемы BB84 с целью адаптировать ее к ПОЗМ. (См. раздел 4.2.2 в лекциях.) Алиса случайным образом готовит состояния, а Боб измеряет их с помощью своей ПОЗМ. Затем он открыто объявляет, как только узнает, что послала Алиса. Конечно, он не распространяется о том, *что* он открыл, а только о том, что это ему известно. При идентификации $|u\rangle \equiv 0$, $|v\rangle \equiv 1$, Алиса и Боб теперь имеют безопасно разделенную строку случайных битов, с помощью которой они могут выполнять шифрование (используя одноразовый протокол). Конечно, прежде чем ее использовать, им также будет необходимо провести коррекцию ошибок и секретное увеличение их строки, чтобы свести вероятность подслушивания к тому уровню, при котором они будут чувствовать себя комфортно. Однако эта «пост-обработка» их строки представляет именно то, что они должны были бы сделать, осуществляя стандартный протокол ортогональных состояний BB84.

c) Вмешательство Евы вызовет лишь ошибку, когда Ева перехватывает посланное Алисой $|u\rangle$ ($|v\rangle$), Бобу передается неправильное состояние $|v\rangle$ ($|u\rangle$). Для любого посланного Алисой сигнала это происходит с вероятностью $\sin^2\alpha$. [Эта симметрия имеет место благодаря тому, что для выполнения своего измерения Ева выбрала в качестве базиса векторы $\binom{1}{0}, \binom{0}{1}$, относительно которых $|u\rangle$ и $|v\rangle$ повернуты на один и тот же угол α по и против часовой стрелки соответственно.]

Имеется два способа описания влияния Евы на скорость появления ошибки, которые лишь слегка различаются в семантике, но дают различные численные результаты. Первым способом описания ее воздействия является:

> *пусть Боб имеет «убедительный» результат, вероятность того, что он отличается от посланного Алисой, равна*
>
> $$\sin^2\alpha.$$

Но поскольку нам известно, что Боб получает «убедительный» результат только с вероятностью $1 - \sin 2\alpha$, можно также сказать:

отнесенная к измеряемому Бобом состоянию вероятность того, что оно не может быть использовано в качестве правильного ключевого бита, равна

$$(1 - \sin 2\alpha) \sin^2 \alpha.$$

Оба этих ответа могут быть полезными. Первый описывает, что Ева испортила часть ключа. Второй описывает, сколько еще необходимо Бобу выполнить измерений, чтобы получить ключ той же длины, что и раньше. Если Алиса и Боб готовы пожертвовать некоторой частью своего ключа, чтобы обнаруживать любые гнусные делишки Евы, им следует выбрать α, максимизирующее последнее выражение, чтобы было как можно легче обнаруживать ее вмешательство. А именно им следует выбрать $\alpha = \pi/8$. Всякий раз, когда частота скорость ошибки их протокола будет превышать $(1 - \sin 2\alpha) \sin^2 \alpha$, они будут подозревать неладное.

Точный смысл поставленного в этой задаче вопроса состоит в рассмотрении влияний на убедительные результаты Боба. Следовательно, фактически первое выражение из приведенных выше следует представить как влияние Евы на протокол.

4.6. Минимальное возмущение

Алиса случайным образом (с равной вероятностью) готовит одно из двух возможных состояний: или $|\psi\rangle = \cos\alpha|0\rangle + \sin\alpha|1\rangle$, или $|\tilde{\psi}\rangle = \sin\alpha|0\rangle + \cos\alpha|1\rangle$. В части (d) упражнения 2.1 мы нашли, что оптимальная ПОЗМ, различающая эти два состояния, состоит из проекторов $|0\rangle\langle 0|$ и $|1\rangle\langle 1|$.

Пусть $\mathbf{M}_0 = |\phi_0\rangle\langle 0|$, а $\mathbf{M}_1 = |\phi_1\rangle\langle 1|$ с произвольными нормированными векторами $|\phi_0\rangle$ и $|\phi_1\rangle$. Это операторы измерения для реализации оптимальной ПОЗМ, различающей приготовленные Алисой состояния:

$$\mathbf{M}_0^\dagger \mathbf{M}_0 = |0\rangle\langle\phi_0|\phi_0\rangle\langle 0| = |0\rangle\langle 0|,$$
$$\mathbf{M}_1^\dagger \mathbf{M}_1 = |1\rangle\langle\phi_1|\phi_1\rangle\langle 1| = |1\rangle\langle 1|.$$

Если Ева выполняет это измерение до того как состояние достигнет Боба, то, в зависимости от того, посылала Алиса $|\psi\rangle$ или $|\tilde{\psi}\rangle$, он получит одно из состояний:

$$\boldsymbol{\rho}' = \sum_i \mathbf{M}_i|\psi\rangle\langle\psi|\mathbf{M}_i^\dagger = \cos^2\alpha|\phi_0\rangle\langle\phi_0| + \sin^2\alpha|\phi_1\rangle\langle\phi_1|,$$

$$\tilde{\boldsymbol{\rho}}' = \sum_i \mathbf{M}_i|\tilde{\psi}\rangle\langle\tilde{\psi}|\mathbf{M}_i^\dagger = \sin^2\alpha|\phi_0\rangle\langle\phi_0| + \cos^2\alpha|\phi_1\rangle\langle\phi_1|.$$

Ева хочет минимизировать «возмущение» $D=1-\frac{1}{2}(F+\tilde{F})$, чтобы максимизировать среднюю точность воспроизведения получаемого Бобом состояния.

а) Мы можем вычислить эти точности воспроизведения:

$$
\begin{aligned}
F &= \langle\psi|\boldsymbol{\rho}'|\psi\rangle = \\
&= \cos^2\alpha\langle\psi|\phi_0\rangle\langle\phi_0|\psi\rangle + \sin^2\alpha\langle\psi|\phi_1\rangle\langle\phi_1|\psi\rangle, \\
F &= \cos^2\alpha\langle\phi_0|\psi\rangle\langle\psi|\phi_0\rangle + \sin^2\alpha\langle\phi_1|\psi\rangle\langle\psi|\phi_1\rangle, \\
\tilde{F} &= \langle\psi|\tilde{\boldsymbol{\rho}}'|\psi\rangle = \\
&= \sin^2\alpha\langle\tilde{\psi}|\phi_0\rangle\langle\phi_0|\tilde{\psi}\rangle + \cos^2\alpha\langle\tilde{\psi}|\phi_1\rangle\langle\phi_1|\tilde{\psi}\rangle, \\
\tilde{F} &= \sin^2\alpha\langle\phi_0|\tilde{\psi}\rangle\langle\tilde{\psi}|\phi_0\rangle + \cos^2\alpha\langle\phi_1|\tilde{\psi}\rangle\langle\tilde{\psi}|\phi_1\rangle.
\end{aligned}
$$

Тогда, складывая их, получим

$$
\begin{aligned}
F+\tilde{F} &= \cos^2\alpha\langle\phi_0|\psi\rangle\langle\psi|\phi_0\rangle + \sin^2\alpha\langle\phi_1|\psi\rangle\langle\psi|\phi_1\rangle + \\
&\quad + \sin^2\alpha\langle\phi_0|\tilde{\psi}\rangle\langle\tilde{\psi}|\phi_0\rangle + \cos^2\alpha\langle\phi_1|\tilde{\psi}\rangle\langle\tilde{\psi}|\phi_1\rangle = \\
&= \left\langle\phi_0\left|\left[\cos^2\alpha|\psi\rangle\langle\psi| + \sin^2\alpha|\tilde{\psi}\rangle\langle\tilde{\psi}|\right]\right|\phi_0\right\rangle + \\
&\quad + \left\langle\phi_1\left|\left[\sin^2\alpha|\psi\rangle\langle\psi| + \cos^2\alpha|\tilde{\psi}\rangle\langle\tilde{\psi}|\right]\right|\phi_1\right\rangle \\
F+\tilde{F} &= \langle\phi_0|A|\phi_0\rangle + \langle\phi_1|B|\phi_1\rangle,
\end{aligned}
$$

где

$$
\begin{aligned}
A &= \cos^2\alpha\begin{pmatrix}\cos^2\alpha & \sin\alpha\cos\alpha\\ \sin\alpha\cos\alpha & \sin^2\alpha\end{pmatrix} + \sin^2\alpha\begin{pmatrix}\sin^2\alpha & \sin\alpha\cos\alpha\\ \sin\alpha\cos\alpha & \cos^2\alpha\end{pmatrix} = \\
&= \begin{pmatrix}(1-\sin^2\alpha)\cos^2\alpha & (1-\sin^2\alpha)\sin\alpha\cos\alpha\\ (1-\sin^2\alpha)\sin\alpha\cos\alpha & \sin^2\alpha\cos^2\alpha\end{pmatrix} + \\
&\quad + \begin{pmatrix}(1-\cos^2\alpha)\sin^2\alpha & \sin^3\alpha\cos\alpha\\ \sin^3\alpha\cos\alpha & \sin^2\alpha\cos^2\alpha\end{pmatrix}, \\
A &= \begin{pmatrix}1-2\sin^2\alpha\cos^2\alpha & \sin\alpha\cos\alpha\\ \sin\alpha\cos\alpha & 2\sin^2\alpha\cos^2\alpha\end{pmatrix}, \\
B &= \sin^2\alpha\begin{pmatrix}\cos^2\alpha & \sin\alpha\cos\alpha\\ \sin\alpha\cos\alpha & \sin^2\alpha\end{pmatrix} + \cos^2\alpha\begin{pmatrix}\sin^2\alpha & \sin\alpha\cos\alpha\\ \sin\alpha\cos\alpha & \cos^2\alpha\end{pmatrix} =
\end{aligned}
$$

$$= \begin{pmatrix} \sin^2\alpha\cos^2\alpha & (1-\cos^2\alpha)\sin\alpha\cos\alpha \\ (1-\cos^2\alpha)\sin\alpha\cos\alpha & (1-\cos^2\alpha)\sin^2\alpha \end{pmatrix} +$$

$$+ \begin{pmatrix} \sin^2\alpha\cos^2\alpha & \sin\alpha\cos^3\alpha \\ \sin\alpha\cos^3\alpha & (1-\cos^2\alpha)\cos^2\alpha \end{pmatrix},$$

$$B = \begin{pmatrix} 2\sin^2\alpha\cos^2\alpha & \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha & 1-2\sin^2\alpha\cos^2\alpha \end{pmatrix}.$$

b) Заметим, что $A^\dagger = A$, а $B^\dagger = B$ – эрмитовы матрицы, следовательно они могут быть диагонализованы. В части (d) упражнения 2.3 мы нашли, что собственные числа 2×2-матрицы можно выразить через их след и детерминант:

$$\lambda_i = \frac{1}{2}\left(\operatorname{tr} M \pm \sqrt{(\operatorname{tr} M)^2 - 4\det M}\right).$$

Заметим также, что $\operatorname{tr} A = \operatorname{tr} B = 1$, а $\det A = \det B$, так что эти матрицы имеют одинаковые (вещественные) собственные значения.

Пусть $\{\lambda_i\}$ и $\{|a_i\rangle\}$ – собственные значения и соответствующие им собственные векторы матрицы A. До тех пор пока A несингулярна (то есть $\det A \neq 0$), мы можем разлагать $|\phi_0\rangle = c_0|a_0\rangle + c_1|a_1\rangle$ в собственном базисе. Тогда мы можем вычислить

$$\langle\phi_0|A|\phi_0\rangle = \left(c_0^*\langle a_0| + c_1^*\langle a_1|\right)\sum_i |a_i\rangle\lambda_i\langle a_i|\left(c_0|a_0\rangle + c_1|a_1\rangle\right),$$

$$\langle\phi_0|A|\phi_0\rangle = \lambda_0|c_0|^2 + \lambda_1|c_1|^2.$$

Это просто взвешенная сумма собственных значений матрицы A. Без потери общности можно предположить, что $\lambda_0 > \lambda_1$. Тогда мы максимизируем выражение $\langle\phi_0|A|\phi_0\rangle$, выбирая $|\phi_0\rangle$ так, чтобы $|c_0| = 1$, а $|c_1| = 0$. Максимальное значение $\langle\phi_0|A|\phi_0\rangle$ просто равно максимальному собственному значению A.

Аналогично оптимальный выбор $|\phi_1\rangle$ сделает выражение $\langle\phi_1|B|\phi_1\rangle$ равным максимальному собственному значению B (которое совпадает с максимальным собственным значением A).

Из предыдущего следует, что $\lambda_{\max} = \frac{1}{2}\left(1 + \sqrt{1 - 4\det A}\right)$. Минимально возможное возмущение равно:

$$D_{\min} = 1 - \frac{1}{2}\left({}_{\text{opt}}\langle\phi_0|A|\phi_0\rangle_{\text{opt}} + {}_{\text{opt}}\langle\phi_1|B|\phi_1\rangle_{\text{opt}}\right) =$$

$$= 1 - \frac{1}{2}\left(\lambda_{\max} + \lambda_{\max}\right) = 1 - \lambda_{\max}$$

$$D_{\min} = \frac{1}{2}\left(1 - \sqrt{1 - 4\det A}\right).$$

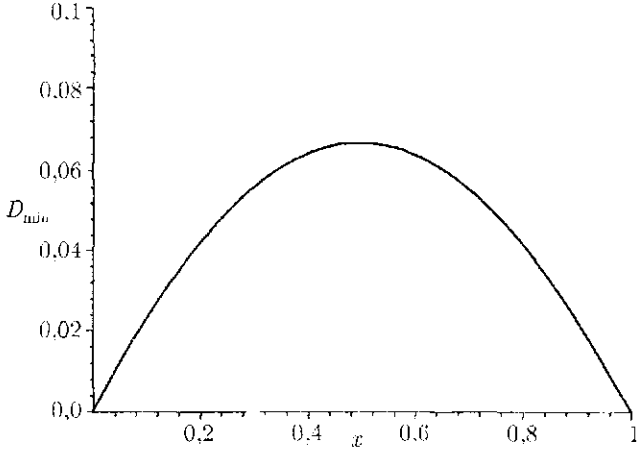
Мы можем переписать A через θ, определяемое соотношением $\cos\theta = \sin 2\alpha$:

$$A = \begin{pmatrix} 1 - \frac{1}{2}\sin^2 2\alpha & \frac{1}{2}\sin 2\alpha \\ \frac{1}{2}\sin 2\alpha & \frac{1}{2}\sin^2 2\alpha \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{2}\cos^2\theta & \frac{1}{2}\cos\theta \\ \frac{1}{2}\cos\theta & \frac{1}{2}\cos^2\theta \end{pmatrix}.$$

Тогда $\det A = \frac{1}{2}\cos^2\theta - \frac{1}{4}\cos^4\theta = \frac{1}{4}\cos^2\theta = \frac{1}{4}\left(\cos^2\theta - \cos^4\theta\right)$ и, следовательно,

$$D_{min} = \frac{1}{2}\left(1 - \sqrt{1 - \cos^2\theta + \cos^4\theta}\right).$$

c) Построим график функции $D_{\min}(x) = \frac{1}{2}\left(1 - \sqrt{1 - x + x^2}\right)$, где $x = \cos^2\theta$.



Если $\cos\theta = 1$, то начальные состояния $|\psi\rangle$ и $|\tilde{\psi}\rangle$ неразличимы. Тогда, независимо от того, выполняла Ева измерение или нет, Боб не может получить никакой информации о том, какое состояние приготовлено Алисой. Это означает, что он не может обнаружить вмешательство Евы, то есть вносимое ей возмущение неизмеримо.

Когда $\cos\theta = 0$, начальные состояния ортогональны, следовательно, Ева может выполнить оптимальную ПОЗМ в базисе $\{|0\rangle, |1\rangle\}$ и, не возмущая состояние, узнать, что приготовила Алиса. Если она пересылает его Бобу, то в этом случае измеримое возмущение отсутствует. Именно это препятствует использованию классических (ортогональных) состояний для безопасной связи — как Алисе и Бобу узнать, что их подслушивают?

Наибольшее значение $D_{\min}$ достигается при $x = \frac{1}{2}$, что соответствует $\theta = \arccos\frac{1}{\sqrt{2}} = \frac{\pi}{4}$. Это является фактическим выбором состояний схемы распределения квантовых ключей (такой как BB84), в которой Алиса и Боб хотят создать разделенную секретную строку битов, одновременно максимизируя величину возмущения, которое в среднем будет вносить подслушивающий, пытаясь узнать значения передаваемых битов. Позднее в этом курсе мы еще больше узнаем о квантовой криптографии.

Для этого выбора приготовлений мы можем вычислить

$$
\begin{aligned}
D_{\min}(x) &= \frac{1}{2} - \left(1 - \sqrt{1 - x + x^2}\right), \\
D_{\min}\left(x = \frac{1}{2}\right) &= \frac{1}{2} - \left(1 - \sqrt{1 - \frac{1}{2} + \frac{1}{4}}\right), \\
D_{\min}\left(x = \frac{1}{2}\right) &= \frac{1}{2} - \left(1 - \sqrt{\frac{3}{4}}\right), \\
D_{\min}\left(x = \frac{1}{2}\right) &\approx 0{,}067, \\
(p_{\text{error}})_{\text{optimal}} &= \frac{1}{2}(1 - \sin\theta), \\
(p_{\text{error}})_{\text{optimal}} &= \frac{1}{2}\left(1 - \frac{1}{\sqrt{2}}\right), \\
(p_{\text{error}})_{\text{optimal}} &\approx 0{,}146.
\end{aligned}
$$

4.7. Приближенное клонирование

а) Эта машина физически реализуема в том и только в том случае, когда она корректно сохраняет вероятности. Это требует, чтобы отображение было унитарным или антиунитарным, а из теоремы Вигнера следует, что на са-

мом деле оно должно быть унитарным. Унитарные отображения сохраняют не только вероятности, но и внутренние произведения. Следовательно, достаточно показать, что машина сохраняет внутреннее произведение, чтобы показать, что она физически реализуема.
До:

$$\begin{aligned}\langle 100|000\rangle &= 0,\\ \langle 000|000\rangle &= 1,\\ \langle 100|100\rangle &= 1.\end{aligned}$$

После:

$$\begin{aligned}\langle 100|\boldsymbol{U}^\dagger\boldsymbol{U}|000\rangle &= \left(\sqrt{\frac{2}{3}}\langle 111| + \sqrt{\frac{1}{3}}\langle\psi^+|\langle 0|\right)\\ &\quad\times\left(\sqrt{\frac{2}{3}}|000\rangle + \sqrt{\frac{1}{3}}|\psi^+\rangle|1\rangle\right)\\ &= 0,\\ \langle 000|\boldsymbol{U}^\dagger\boldsymbol{U}|000\rangle &= \frac{2}{3} + \frac{1}{3} = 1,\\ \langle 100|\boldsymbol{U}^\dagger\boldsymbol{U}|100\rangle &= \frac{2}{3} + \frac{1}{3} = 1.\end{aligned}$$

b) Переход от унитарного представления к представлению операторной суммы \$ выполняется путем отождествления

$$\mathbf{M}_\mu = {}_{BC}\langle\mu|\boldsymbol{U}|\bar{0}\rangle_{BC},$$

где $|0\rangle_{BC}$ — некоторое фиксированное состояние. К данному в условии задачи описанию отображения естественно подходит выбор $|\bar{0}\rangle_{BC} = |00\rangle_{BC}$. В этом базисе необходимые нам существенные слагаемые $\boldsymbol{U}$ имеют вид

$$\begin{aligned}\boldsymbol{U} = &\sqrt{\frac{2}{3}}|000\rangle\langle 000| + \sqrt{\frac{1}{6}}|011\rangle\langle 000| + \sqrt{\frac{1}{6}}|101\rangle\langle 000| +\\ &+ \sqrt{\frac{2}{3}}|111\rangle\langle 100| + \sqrt{\frac{1}{6}}|010\rangle\langle 100| + \sqrt{\frac{1}{6}}|100\rangle\langle 100|.\end{aligned}$$

Следовательно, операторами представления операторной суммы являются:

$$\mathbf{M}_{00} = {}_{BC}\langle 00|\boldsymbol{U}|00\rangle_{BC} = \sqrt{\frac{2}{3}}|0\rangle\langle 0| + \sqrt{\frac{1}{6}}|1\rangle\langle 1|,$$

$$\mathbf{M}_{01} = {}_{BC}\langle 01|\boldsymbol{U}|00\rangle_{BC} = \sqrt{\frac{1}{6}}|1\rangle\langle 0|,$$

$$\mathbf{M}_{10} = {}_{BC}\langle 10|\boldsymbol{U}|00\rangle_{BC} = \sqrt{\frac{1}{6}}|0\rangle\langle 1|,$$

$$\mathbf{M}_{11} = {}_{BC}\langle 11|\boldsymbol{U}|00\rangle_{BC} = \sqrt{\frac{1}{6}}|0\rangle\langle 0| + \sqrt{\frac{2}{3}}|1\rangle\langle 1|.$$

c) Запись в виде суммы проекторов для матриц операторной суммы $\mathbf{M}_\mu$ позволяет быстро вычислить точность воспроизведения F даже без явного вычисления $\boldsymbol{\rho}'_A$:

$$F = \langle\psi|\boldsymbol{\rho}'_A|\psi\rangle = \left\langle\psi\left|\left(\sum_\mu \mathbf{M}_\mu|\psi\rangle\langle\psi|\mathbf{M}_\mu^\dagger\right)\right|\psi\right\rangle =$$

$$= \sum_\mu \langle\psi|\mathbf{M}_\mu|\psi\rangle\langle\psi|\mathbf{M}_\mu^\dagger|\psi\rangle = \sum_\mu \left|\langle\psi|\mathbf{M}_\mu|\psi\rangle\right|^2 =$$

$$= \left|\sqrt{\frac{2}{3}}|a|^2 + \sqrt{\frac{1}{6}}|b|^2\right|^2 + \frac{1}{3}|a|^2|b|^2 + \left|\sqrt{\frac{1}{6}}|a|^2 + \sqrt{\frac{2}{3}}|b|^2\right|^2 =$$

$$= \frac{5}{6}|a|^4 + \frac{5}{6}|b|^4 + \frac{5}{3}|a|^2|b|^2 = \frac{5}{6}\left(|a|^2 + |b|^2\right)^2 = \frac{5}{6}.$$

Этот результат означает, что, имея ресурс двух дополнительных вспомогательных кубитов, мы можем копировать неизвестный кубит с точностью воспроизведения $5/6$.

Действие этого «квантового ксерокса» на входящее состояние отображает его на состояние

$$\boldsymbol{\rho}'_A = \sum_\mu \mathbf{M}_\mu|\psi\rangle\langle\psi|\mathbf{M}_\mu^\dagger =$$

$$= \frac{1}{6}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} +$$

$$+\frac{1}{6}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} +$$

$$
\begin{aligned}
&+\frac{1}{6}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}+ \\
&+\frac{1}{6}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}= \\
&=\frac{1}{6}\begin{pmatrix} 4|a|^2 & 2ab^* \\ 2a^*b & |b|^2 \end{pmatrix}+\frac{1}{6}\begin{pmatrix} 0 & 0 \\ 0 & |a|^2 \end{pmatrix}= \\
&=\frac{1}{6}\begin{pmatrix} |b|^2 & 0 \\ 0 & 0 \end{pmatrix}+\frac{1}{6}\begin{pmatrix} |a|^2 & 2ab^* \\ 2a^*b & 4|b|^2 \end{pmatrix}= \\
&=\frac{1}{6}\begin{pmatrix} 5|a|^2+|b|^2 & 4ab^* \\ 4a^*b & |a|^2+5|b|^2 \end{pmatrix}= \\
&=\frac{1}{6}\begin{pmatrix} 4|a|^2 & 4ab^* \\ 4a^*b & 4|b|^2 \end{pmatrix}+\frac{1}{6}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=\frac{2}{3}|\psi\rangle\langle\psi|+\frac{1}{6}\mathbf{1}.
\end{aligned}
$$

Этот квантовый ксерокс действует так же, как и деполяризующий канал с $p = 1/4$. Следовательно, в части (b) мы с полным основанием могли использовать операторы из раздела 3.4.1.

4.8. Прости нас, дядюшка Альберт

a) Пусть $\boldsymbol{\Sigma}_n$ обозначает оператор $(\boldsymbol{\sigma}_1 + i\boldsymbol{\sigma}_2)^{\otimes n} + (\boldsymbol{\sigma}_1 - i\boldsymbol{\sigma}_2)^{\otimes n}$. Из элементарной квантовой механики известно, что

$$\boldsymbol{\sigma}_\pm \equiv \frac{1}{2}(\boldsymbol{\sigma}_1 \pm i\boldsymbol{\sigma}_2)$$

являются повышающими и понижающими операторами для спина. Следовательно, действие $\boldsymbol{\Sigma}_n$ на $|\psi\rangle_n$ имеет вид:

$$
\begin{aligned}
\boldsymbol{\Sigma}_n|\psi\rangle_n &= \boldsymbol{\Sigma}_n\sqrt{\frac{1}{2}}\big(|000\ldots0\rangle + |111\ldots1\rangle\big) = \\
&= \sqrt{\frac{1}{2}}\Big[\big(2\boldsymbol{\sigma}_+\big)^{\otimes n}|000\ldots0\rangle + \big(2\boldsymbol{\sigma}_-\big)^{\otimes n}|111\ldots1\rangle\big)\Big] = \\
&= 2^n\cdot\sqrt{\frac{1}{2}}\big(|111\ldots1\rangle + |000\ldots0\rangle\big) = 2^n|\psi\rangle_n.
\end{aligned}
$$

b) Теория скрытых переменных утверждает, что $\boldsymbol{\sigma}_1$ и $\boldsymbol{\sigma}_2$ в любой момент времени являются функциями набора недоступных для нас «скрытых переменных». Она утверждает, что наша неспособность узнать эти переменные

заставляет все измерения σ_1 и σ_2 давать только *усредненные по ансамблю* этих скрытых переменных результаты.

Заметим, однако, что *модуль* оператора $(\sigma_1 \pm i\sigma_2)^{\otimes n}$ имеет только одно значение для любого возможного распределения значений наблюдаемых σ_1 и σ_2:

$$\left|(\sigma_1 \pm i\sigma_2)^{\otimes n}\right| = |(\pm 1) \pm i(\pm 1)|^n = 2^{n/2}.$$

c) Оператор $(\sigma_1 + i\sigma_2)^{\otimes n} + (\sigma_1 - i\sigma_2)^{\otimes n}$ эрмитов и, следовательно, его модуль является наблюдаемой величиной. Теория скрытых переменных предсказывает, что ее измеряемое значение дается усредненным по ансамблю определенных значений, принимаемых σ_1 и σ_2. Используя неравенство треугольника для нормы, мы можем ограничить этот модуль суммой выражений, вычисленных в части (b). Важность этого ограничения в том, что вычисленные в части (b) слагаемые *независимы* от любого такого распределения:

$$\begin{aligned} |\Sigma_n| &= \left|(\sigma_1 + i\sigma_2)^{\otimes n} + (\sigma_1 - i\sigma_2)^{\otimes n}\right| \leqslant \\ &\leqslant \left|(\sigma_1 + i\sigma_2)^{\otimes n}\right| + \left|(\sigma_1 - i\sigma_2)^{\otimes n}\right| = \\ &= |\sigma_1 + i\sigma_2|^n + |\sigma_1 - i\sigma_2|^n = |\sqrt{2}|^n + |\sqrt{2}|^n = 2^{n/2+1}. \end{aligned}$$

d) Эйнштейн сказал бы:

> *Sie haben demonstriert (wie auf der Hand gelegen haben sollte), daß mein Argument gegen Quanten Mechanik wissenschaftlich stichhaltig ist, weil es durch Experiment falsifizierbar ist. Für die Systeme, die $n > 2$ haben, sind die Voraussagungen der lokalen versteckten variablen Theorie und Quanten Mechanik offenbar inkompatibel. Lassen Sie uns ein Experiment machen, um zu überprüfen, daß ich Recht habe*[1]...

Под впечатлением экспериментального свидетельства, которое поддерживает квантовую механику и опровергает теорию скрытых переменных, Эйнштейн бы сказал:

> *Ach! Dieses ist wirklich mein größter Fehlgriff. Es scheint, daß Gott tatsächlich Würfel spielt*[2].

[1] Вы продемонстрировали (как это и должно было быть очевидно), что мои аргументы против квантовой механики научно обоснованы, пока они не опровергнуты экспериментально. Для системы из $n > 2$ частей предсказания локальной теории скрытых переменных и квантовой механики, очевидно, несовместимы. Давайте поставим эксперимент, чтобы убедиться в том, что я прав... .

[2] Ах! Это поистине моя самая большая ошибка. Похоже, что Бог действительно играет в кости.

4.9. Манипуляция запутыванием

а) Алиса может связать команды с помощью обмена запутыванием. Я опишу этот процесс на языке состояний и языке стабилизатора. Судите сами, какой язык покажется вам наиболее подходящим для этой задачи.

Язык состояний. Начальным состоянием системы является

$$
\begin{aligned}
|A_1Y, A_2P\rangle = \frac{1}{2}\big(&|00\rangle_A|\bar{0}\rangle_Y|\bar{0}\rangle_P + |11\rangle_A|\bar{1}\rangle_Y|1\rangle_P + \\
&+|01\rangle_A|\bar{0}\rangle_Y|\bar{1}\rangle_P + |10\rangle_A|\bar{1}\rangle_Y|\bar{0}\rangle_P\big) = \\
= \frac{1}{2\sqrt{2}}\Big[&|\Phi^+\rangle_A\big(|\bar{0}\rangle_Y|\bar{0}\rangle_P + |\bar{1}\rangle_Y|\bar{1}\rangle_P\big) + \\
&+|\Phi^-\rangle_A\big(|\bar{0}\rangle_Y|\bar{0}\rangle_P - |\bar{1}\rangle_Y|\bar{1}\rangle_P\big) + \\
&+|\Psi^+\rangle_A\big(|\bar{0}\rangle_Y|\bar{1}\rangle_P + |\bar{1}\rangle_Y|\bar{0}\rangle_P\big) + \\
&+|\Psi^-\rangle_A\big(|\bar{0}\rangle_Y|\bar{1}\rangle_P - |\bar{1}\rangle_Y|\bar{0}\rangle_P\big)\Big].
\end{aligned}
$$

Алиса измеряет два ее состояния в базисе Белла. Затем она посылает одной из команд два полученных ей классических бита. Тогда эта команда выполняет одну из следующих операций, гарантирующих, что получающееся в результате 50-кубитовое состояние является кот-состоянием

Состояние		Действие
$\|\Phi^+\rangle_A$	$\longrightarrow$	Ничего не делает.
$\|\Phi^-\rangle_A$	$\longrightarrow$	Один участник применяет σ_z.
$\|\Psi^+\rangle_A$	$\longrightarrow$	Все участники применяют σ_x.
$\|\Psi^-\rangle_A$	$\longrightarrow$	Все участники применяют σ_x Один участник применяет σ_z

Язык стабилизатора. Исходным стабилизатором системы является[1]

[1] Стабилизаторы — симплектические коды, корректирующие ошибки, рассматриваются в седьмой главе лекций, вошедшей во вторую часть этой книги. — *Прим. ред.*

Янки (25 кубитов)	Алиса	Святые отцы (25 кубитов)	Собственное значение
Z Z 1 ⋯ 1	1 1		+1
1 Z Z ⋯ 1	1 1		+1
⋮ ⋮ ⋮ ⋮	⋮ ⋮	1	⋮
1 1 1 ⋯ Z	Z 1		+1
X X X ⋯ X	X 1		+1
	1 Z	Z 1 1 ⋯ 1	+1
	1 1	Z Z 1 ⋯ 1	+1
1	⋮ ⋮	⋮ ⋮ ⋮ ⋮	⋮
	1 1	1 1 1 ⋯ Z	+1
	1 X	X X X ⋯ X	+1

Алиса измеряет **ZZ**, затем **XX**, получая собственные значения α_{par} и α_{ph}. После каждого измерения генераторы стабилизатора заменяются только на те генераторы, которые коммутируют с измерением. (Заметим, что произведение двух антикоммутирующих с измерением генераторов коммутирует с этим измерением.) Результирующим стабилизатором является (для ясности две колонки Алисы сдвинуты влево)

Алиса	Янки — Святые отцы	Собственное значение
Z Z	1	α_{par}
X X		α_{ph}
	Z Z 1 ⋯ ⋯ 1	+1
	1 Z Z ⋯ ⋯ 1	+1
	⋮ ⋮	⋮
1	1 ⋯ Z Z ⋯ 1	α_{par}
	⋮ ⋮	⋮
	1 1 1 ⋯ ⋯ Z	+1
	X X X ⋯ ⋯ X	α_{ph}

Для того чтобы команды разделили кот-состояние, Алиса должна сообщить одной из них собственные значения α_{par} и α_{ph}. После чего эта команда выполняет одну из следующих операций, гарантирующих, что получающийся в результате стабилизатор Янки и Святых Отцов имеет все собственные значения равные $+1$ (вспомним, что операция A переводит генератор стабилизатора $\mathbf{M}$ в $AMA^{\dagger}$):

$(\alpha_{\text{par}}, \alpha_{\text{ph}})$		Действие
$(+1,+1)$	$\longrightarrow$	Ничего не делает.
$(+1,-1)$	$\longrightarrow$	Один участник применяет $\mathbf{Z}$.
$(-1,+1)$	$\longrightarrow$	Все участники применяют $\mathbf{X}$.
$(-1,-1)$	$\longrightarrow$	Все участники применяют $\mathbf{X}$ Один участник применяет $\mathbf{Z}$

b) (I) Если Алиса имеет вспомогательный кубит...

Имея вспомогательный кубит, Алиса может оставить команду в некотором смысле в том же положении, что и в части (a). Снова я опишу ее действия на языке состояний и языке стабилизатора.

Язык состояний. Алиса готовит свой вспомогательный кубит в состоянии

$$|A_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

Следовательно, начальным состоянием системы является

$$\begin{aligned}|A_1A_2Y\rangle = \frac{1}{2}\big(&|00\rangle_A|\bar{0}\rangle_Y + |11\rangle_A|\bar{1}\rangle_Y + \\ &+|01\rangle_A|\bar{0}\rangle_Y + |10\rangle_A|\bar{1}\rangle_Y\big) = \\ = \frac{1}{2\sqrt{2}}\Big[&|\Phi^+\rangle_A\big(|\bar{0}\rangle_Y + |\bar{1}\rangle_Y\big) + \\ &+|\Phi^-\rangle_A\big(|\bar{0}\rangle_Y - |\bar{1}\rangle_Y\big) + \\ &+|\Psi^+\rangle_A\big(|\bar{0}\rangle_Y + |\bar{1}\rangle_Y\big) - \\ &-|\Psi^-\rangle_A\big(|\bar{0}\rangle_Y - |\bar{1}\rangle_Y\big)\Big].\end{aligned}$$

Алиса измеряет два ее состояния в базисе Белла. (На самом деле ей нужно измерить только бит фазы.) Затем она посылает результат измерения оставшейся команде. Тогда эта команда выполняет одну из следующих

операций, гарантирующих, что получающееся в результате 24-кубитовое состояние является кот-состоянием:

Состояние		Действие
$\|\Phi^+\rangle_A, \|\Psi^+\rangle_A$	$\longrightarrow$	Ничего не делает.
$\|\Phi^-\rangle_A, \|\Psi^-\rangle_A$	$\longrightarrow$	Один участник применяет σ_z

Язык стабилизатора. Алиса готовит свой вспомогательный кубит в собственном состоянии $\mathbf{X} = +1$.[1] Начальным стабилизатором системы является

Алиса		Янки (24 кубита)			Собственное значение
X	**1**	**1**	$\cdots$	**1**	+1
1	**X**	**X**	$\cdots$	**X**	+1
1	**Z**	**Z**	$\cdots$	**1**	+1
1	**1**	**Z**	$\cdots$	1	+1
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
1	**1**	**1**	$\cdots$	**Z**	+1

Алиса измеряет **XX** на ее двух кубитах, получая собственное значение α. Новым стабилизатором является

Алиса		Янки (24)					Собственное значение
X	**X**						α
X	**1**			**1**			+1
		Z	**Z**	**1**	$\cdots$	**1**	+1
		1	**Z**	**Z**	$\cdots$	**1**	+1
1		$\vdots$	$\cdots$	$\vdots$		$\vdots$	−1
		1	**1**	**1**	$\cdots$	**Z**	+1
		X	**X**	**X**	$\cdots$	**X**	α

Алиса сообщает свое собственное значение α оставшейся команде, которая выполняет одну из следующих операций, гарантирующих, что полу-

[1] То есть в собственном состоянии оператора **X** с собственным значением +1. — *Прим. ред.*

чающийся в результате их стабилизатор имеет все собственные значения, равные $+1$:

α		Действие
$+1$	$\longrightarrow$	Ничего не делает.
-1	$\longrightarrow$	Один участник применяет $\mathbf{Z}$

(II) ..., а если она не имеет вспомогательного кубита.

Даже если Алиса не имеет вспомогательного кубита, прицеплявшегося выше, она по-прежнему может покинуть команду. Снова я опишу ее действия (на обоих языках).

Язык состояний. Начальным состоянием системы является

$$|AY\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A|\bar{0}\rangle_Y + |1\rangle_A|\bar{1}\rangle_Y =$$

$$= \frac{1}{2}\Big[(|0\rangle_A + |1\rangle_A)(|\bar{0}\rangle_Y + |\bar{1}\rangle_Y) + (|0\rangle_A - |1\rangle_A)(|\bar{0}\rangle_Y - |\bar{1}\rangle_Y)\Big].$$

Алиса измеряет σ_x на своем спине и посылает результат оставшейся команде. Эта команда выполняет одну из следующих операций, гарантирующих, что получающееся в результате 24-кубитовое состояние является кот-состоянием:

Состояние		Действие
$\lvert\uparrow_x\rangle_A$	$\longrightarrow$	Ничего не делает.
$\lvert\downarrow_x\rangle_A$	$\longrightarrow$	Один участник применяет σ_z

Язык стабилизатора. Начальным стабилизатором системы является

Алиса	Янки (24 кубита)				Собственное значение
$\mathbf{X}$	$\mathbf{X}$	$\mathbf{X}$	$\cdots$	$\mathbf{X}$	$+1$
$\mathbf{Z}$	$\mathbf{Z}$	$\mathbf{1}$	$\cdots$	$\mathbf{1}$	$+1$
$\mathbf{1}$	$\mathbf{Z}$	$\mathbf{Z}$	$\cdots$	$\mathbf{1}$	$+1$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	$\cdots$	$\mathbf{Z}$	$+1$

Алиса измеряет $\mathbf{X}$ на своем кубите, получая собственное значение α. Новым стабилизатором является

Алиса	Янки (24)					Собственное значение
$\mathbf{X}$			$\mathbf{1}$			α
	$\mathbf{Z}$	$\mathbf{Z}$	$\mathbf{1}$	$\cdots$	$\mathbf{1}$	$+1$
	$\mathbf{1}$	$\mathbf{Z}$	$\mathbf{Z}$	$\cdots$	$\mathbf{1}$	$+1$
$\mathbf{1}$	$\vdots$	$\cdots$	$\vdots$		$\vdots$	$\vdots$
	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{1}$	$\cdots$	$\mathbf{Z}$	$+1$
	$\mathbf{X}$	$\mathbf{X}$	$\mathbf{X}$	$\cdots$	$\mathbf{X}$	α

Алиса сообщает свое собственное значение α оставшейся команде, которая выполняет одну из следующих операций, гарантирующих, что получающийся в результате их стабилизатор имеет все собственные значения, равные $+1$:

α		Действие
$+1$	$\longrightarrow$	Ничего не делает
-1	$\longrightarrow$	Один участник применяет $\mathbf{Z}$

Решения упражнений к главе 5

5.1. Различимость неортогональных состояний

В отсутствие какой-либо предварительной информации, мы (как правоверные бейесиане) должны предположить, что с равной вероятностью Алиса готовит одно из состояний $|u\rangle$ и $|v\rangle$.

Пусть во всех частях этой задачи определены следующие случайные переменные:

$$A = \text{состояние, которое готовит Алиса,}$$

$$B = \text{результат, который получает Боб.}$$

Пусть для краткости B принимает значения $1, 2$ или 3, соответствующие применяемому Бобом измерительному оператору, а A принимает значения u и v, соответствующие приготавливаемым Алисой состояниям. В каждой части этой задачи мы должны вычислить следующие величины[1]:

$$p(i|w) = \begin{cases} |\langle w|\mathbf{E}_i|w\rangle|^2 & \text{(ортогональное измерение)} \\ |\langle w|\mathbf{F}_i|w\rangle|^2 & \text{(ПОЗМ)} \end{cases},$$

[1] Конечно, учитывая связь $I(B;A) = I(A;B) = H(A) - H(A|B)$, вместо этого для вычислений можно выбрать $H(A|B)$ и $H(A)$.

$$H(B|A) = -\sum_{a,b} p(a,b)\log p(b|a) =$$

$$= -\sum_{a,b} p(b|a)p(a)\log p(b|a),$$

$$H(B) = -\sum_{a,b} p(a,b)\log p(b) =$$

$$= -\sum_{a,b} p(b|a)p(a)\log\left(\sum_{c} p(b|c)p(c)\right),$$

$$I(B;A) = H(B) - H(B|A).$$

а) Результатом исходной «фон неймановской» стратегии Боба является приобретение следующей информации.

Вероятности:

$$p(1|u) = 1, \quad p(2|u) = 0, \quad p(u) = \frac{1}{2},$$

$$p(1|v) = \cos^2\frac{\theta}{2}, \quad p(2|v) = \sin^2\frac{\theta}{2}, \quad p(v) = \frac{1}{2}.$$

Условная энтропия:

$$H(B|A) = -\frac{1}{2}\cos^2\frac{\theta}{2}\log\left(\cos^2\frac{\theta}{2}\right) - \frac{1}{2}\sin^2\frac{\theta}{2}\log\left(\sin^2\frac{\theta}{2}\right).$$

Энтропия Шеннона:

$$H(B) = -\frac{1}{2}\left(1+\cos^2\frac{\theta}{2}\right)\log\left[\frac{1}{2}\left(1+\cos^2\frac{\theta}{2}\right)\right] - \frac{1}{2}\sin^2\frac{\theta}{2}\log\left(\frac{1}{2}\sin^2\frac{\theta}{2}\right).$$

Взаимная информация:

$$I(B;A) = 1 - \frac{1}{2}\left(1+\cos^2\frac{\theta}{2}\right)\log\left(1+\cos^2\frac{\theta}{2}\right) + \frac{1}{2}\cos^2\frac{\theta}{2}\log\left(\cos^2\frac{\theta}{2}\right) =$$

$$= 1 - \frac{1}{2}\left(1+\cos^2\frac{\theta}{2}\right) H_2\left(\frac{1}{1+\cos^2\frac{\theta}{2}}\right),$$

где $H_2(x) = -x\log x - (1-x)\log(1-x)$ — бинарная функция энтропии.

b) Осуществляя более симметричное измерение ПОЗМ, Боб рассчитывает увеличить приобретение информации. Фактически мы находим, что, поступая так, он на самом деле *уменьшает* приобретение информации.

Вероятности:

$$p(1|u) = 0, \qquad p(2|v) = 0,$$

$$p(2|u) = A\sin^2\frac{\theta}{2}, \qquad p(1|v) = A\sin^2\frac{\theta}{2},$$

$$p(3|u) = 1 - A + A\cos^2\frac{\theta}{2}, \qquad p(3|v) = 1 - A + A\cos^2\frac{\theta}{2},$$

$$p(u) = \frac{1}{2}, \qquad p(v) = \frac{1}{2}.$$

Условная энтропия:

$$H(B|A) = -\left[A\sin^2\frac{\theta}{2}\log\left(A\sin^2\frac{\theta}{2}\right) + \right.$$
$$\left. + \left(1 - A + A\cos^2\frac{\theta}{2}\right)\log\left(1 - A + A\cos^2\frac{\theta}{2}\right)\right].$$

Энтропия Шеннона:

$$H(B) = -\left[A\sin^2\frac{\theta}{2}\log\left(\frac{1}{2}A\sin^2\frac{\theta}{2}\right) + \right.$$
$$\left. + \left(1 - A + A\cos^2\frac{\theta}{2}\right)\log\left(1 - A + A\cos^2\frac{\theta}{2}\right)\right].$$

Взаимная информация:

$$I(B;A) = -\left[A\sin^2\frac{\theta}{2}\log\frac{1}{2} + \left(1 - A + A\cos^2\frac{\theta}{2}\right)\log 1\right] -$$
$$- A\sin^2\frac{\theta}{2}.$$

Чтобы найти A, мы используем требование положительности $\mathbf{F}_3$, а именно $\mathbf{F}_3$ имеет положительные собственные значения:

$$\lambda^2 - \lambda\,\mathrm{tr}\,\mathbf{F}_3 + \det\mathbf{F}_3 = 0, \implies \lambda = 1 - A \pm A\cos\frac{\theta}{2},$$

$$\lambda \geqslant 0, \implies A \leqslant \frac{1}{1 \pm \cos\theta/2}.$$

При $\theta \in [0, \pi]$ таким наибольшим A, при котором оба собственных значения остаются положительными, является $A = \frac{1}{1+\cos\theta/2}$. Следовательно, максимальное приобретение информации равно

$$I(B;A) = 2\sin^2\frac{\theta}{4} = 1 - \cos\frac{\theta}{2}.$$

c) В последней отчаянной попытке Боб возвращается к измерению фон Неймана, которое «выявляет различие» между $|u\rangle$ и $|v\rangle$. Эта схема действительно оказывается наилучшей.

Вероятности:

$$p(1|u) = \cos^2\left(\frac{\theta+\pi}{4}\right), \qquad p(1|v) = \sin^2\left(\frac{\theta+\pi}{4}\right),$$

$$p(2|u) = \sin^2\left(\frac{\theta+\pi}{4}\right), \qquad p(2|v) = \cos^2\left(\frac{\theta+\pi}{4}\right),$$

$$p(u) = \frac{1}{2}, \qquad p(v) = \frac{1}{2}.$$

Условная энтропия:

$$H(B|A) = -\left[\sin^2\left(\frac{\theta+\pi}{4}\right)\log\left(\sin^2\left(\frac{\theta+\pi}{4}\right)\right) + \right.$$
$$\left. + \cos^2\left(\frac{\theta+\pi}{4}\right)\log\left(\cos^2\left(\frac{\theta+\pi}{4}\right)\right)\right].$$

Энтропия Шеннона:

$$H(B) = -\left[\frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}\right] = 1.$$

Взаимная информация:

$$I(B;A) = 1 + \sin^2\left(\frac{\theta+\pi}{4}\right)\log\left(\sin^2\left(\frac{\theta+\pi}{4}\right)\right) +$$
$$+ \cos^2\left(\frac{\theta+\pi}{4}\right)\log\left(\cos^2\left(\frac{\theta+\pi}{4}\right)\right) =$$
$$= 1 - H_2\left(\cos^2\left(\frac{\theta+\pi}{4}\right)\right).$$

d) Несмотря на то что в условии задачи не было раздела **(d)**, в ее контексте полезно рассмотреть границу Холево, если, конечно, мы уверены в разумности результатов, полученных в предыдущих частях этой задачи.

Граница Холево утверждает, что приобретаемая Бобом информация ограничена сверху доступной информацией источника Алисы $\mathrm{Acc}(\mathcal{E}) = S(\boldsymbol{\rho}) - \sum_x p_x S(\boldsymbol{\rho}_x)$. Поскольку оба сигнальных состояния Алисы являются чистыми состояниями, доступная информация сводится к энтропии фон Неймана, которую мы можем вычислить путем диагонализации:

$$\boldsymbol{\rho} = \begin{pmatrix} \frac{1}{2}\left(1+\cos^2\frac{\theta}{2}\right) & \frac{1}{2}\cos\frac{\theta}{2}\sin\frac{\theta}{2} \\ \frac{1}{2}\cos\frac{\theta}{2}\sin\frac{\theta}{2} & \frac{1}{2}\sin^2\frac{\theta}{2} \end{pmatrix},$$

$$\lambda^2 - \lambda + \frac{1}{4}\left[\left(1+\cos^2\frac{\theta}{2}\right)\sin^2\frac{\theta}{2} - \cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2}\right] = 0,$$

$$\lambda = \frac{1}{2} \pm \frac{1}{2}\sqrt{1-\sin^2\frac{\theta}{2}} = \frac{1}{2} \pm \frac{1}{2}\cos\frac{\theta}{2} = \cos^2\frac{\theta}{4} \quad \text{или} \quad \sin^2\frac{\theta}{4}.$$

Таким образом, приобретаемая Бобом информация ограничена условием

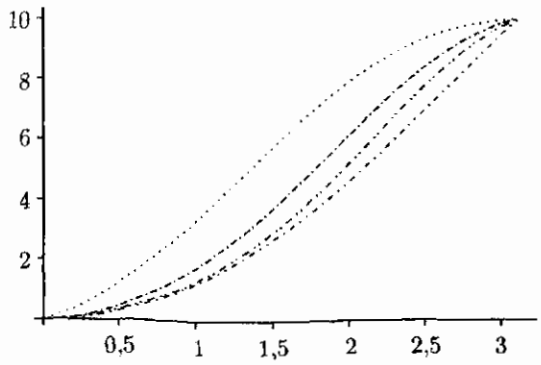
$$I(B;A) \leqslant \sin^2\frac{\theta}{4}\log\left(\sin^2\frac{\theta}{4}\right) - \cos^2\frac{\theta}{4}\log\left(\cos^2\frac{\theta}{4}\right) = H_2\left(\cos^2\frac{\theta}{4}\right).$$

Построив график приобретаемой Бобом информации, как функции для каждой из трех схем вместе с границей Холево, мы видим, что наилучшим выбором Боба является стратегия **(c)**, несмотря на то, что граница Холево не насыщается. На диаграмме ниже стратегии a, b, c и h (для Холево) помечены соответствующими кодами азбуки Морзе[1]

5.2. Относительная энтропия

a) Эта задача содежит небольшую трудность, поскольку A и B не обязательно коммутируют между собой. Однако, работая с некоммутирующими объектами, мы можем применить обычный трюк и разлагать все выражения по компонентам обычных коммутирующих чисел. Пусть $\{|i\rangle\}$ является базисом, диагонализующим A, и пусть $\{|j\rangle\}$ – базис, диагонализующий B.

[1] Коды азбуки Морзе: a ↔ ·−; b ↔ −···; c ↔ −·−·; h ↔ ····. – *Прим. перев.*



Разлагая по этим базисам, мы имеем

$$\begin{aligned}\operatorname{tr}\left[f(\mathbf{B})-f(\mathbf{A})\right] &= \sum_i \langle i|f(\mathbf{B})-f(a_i)|i\rangle = \\ &= \sum_{i,j} \langle i|f(\mathbf{B})-f(a_i)|j\rangle\langle j|i\rangle = \\ &= \sum_{i,j} \langle i|f(b_j)-f(a_i)|j\rangle\langle j|i\rangle \leqslant \sum_{i,j} \langle i|(b_j-a_i)f'(a_i)|j\rangle\langle j|i\rangle = \\ &= \sum_i \langle i|(\mathbf{B}-a_i)f'(a_i)|i\rangle = \operatorname{tr}\left[(\mathbf{B}-\mathbf{A})f'(\mathbf{A})\right].\end{aligned}$$

b) Этот результат, подобно многим неравенствам в теории информации, следует из неравенства $\ln x \leqslant x-1$. Достаточно показать, что $g(x) = -x\ln x$ является вогнутой функцией, следовательно, вогнутой является и функция $f(x) = g(x)/\ln 2$. Доказательство для $g(x)$:

$$\begin{aligned}g(y)-g(x) &= -y\ln y + x\ln x = \\ &= y(\ln x - \ln y) + (x-y)\ln x = \\ &= y\ln\frac{x}{y} - (y-x)\ln x \leqslant \\ &\leqslant y\left(\frac{x}{y}-1\right) - (y-x)\ln x = \\ &= (y-x)(-1-\ln x) = \\ &= (y-x)g'(x);\end{aligned}$$

$g(x)$ — вогнутая функция $\Longrightarrow$ $f(x)$ — вогнутая функция.

c) Применяя результаты **(a)** и **(b)**, находим, что относительная энтропия неотрицательна[1]:

$$\mathrm{tr}[-\rho\log\rho+\sigma\log\sigma]\leqslant\mathrm{tr}\left[(\rho-\sigma)\left(-\log\sigma-\frac{1}{\ln 2}\right)\right],$$

согласно частям **(a)** и **(b)**. Далее,

$$\begin{aligned}-\mathrm{tr}\,\rho\log\rho+\mathrm{tr}\,\sigma\log\sigma&\leqslant-\mathrm{tr}\,\rho\log\sigma+\mathrm{tr}\,\sigma\log\sigma,\\ -\mathrm{tr}\,\rho\log\rho&\leqslant-\mathrm{tr}\,\rho\log\sigma,\\ 0&\leqslant\mathrm{tr}\,\rho\log\rho-\mathrm{tr}\,\rho\log\sigma,\\ 0&\leqslant S(\rho|\sigma).\end{aligned}$$

d) Пусть σ — матрица плотности, совпадающая с единицей в подпространстве, являющемся носителем ρ. К искомому результату ведет выражение неотрицательности относительной энтропии между ρ и σ в базисе, в котором они обе диагональны:

$$\begin{aligned}0&\leqslant S(\rho|\sigma)-S(\rho)-\mathrm{tr}\,\rho\log\sigma,\\ S(\rho)&\leqslant-\mathrm{tr}\left[\rho_1\log\sigma_1+\ldots+\rho_D\log\sigma_D\right]=\\ &=-\left(\log\frac{1}{D}\right)\mathrm{tr}(\rho_1+\ldots+\rho_D)=\log D.\end{aligned}$$

e) Используя неотрицательность относительной энтропии между ρ_{AB} и $\rho_A\otimes\rho_B$, находим

$$\begin{aligned}S(\rho_{AB}|\rho_A\otimes\rho_B)&=-S(\rho_{AB})-\mathrm{tr}\left[\rho_{AB}\log(\rho_A\otimes\rho_B)\right]\geqslant 0,\\ S(\rho_{AB})&\leqslant-\mathrm{tr}\left[\rho_{AB}\left(\log\rho_A\otimes\mathbf{1}_B+\mathbf{1}_A\otimes\log\rho_B\right)\right]=\\ &=-\mathrm{tr}\,\rho_A\log\rho_A-\mathrm{tr}\,\rho_B\log\rho_B=S(\rho_A)+S(\rho_B).\end{aligned}$$

f) Рассмотрим матрицу плотности ρ_{AB} и ее частичные следы, данные соотношениями

$$\begin{aligned}\rho_{AB}&=\sum_i\lambda_i(\rho_i)_A\otimes(|e_i\rangle\langle e_i|)_B,\\ \rho_A&=\sum_i\lambda_i\rho_i,\\ \rho_B&=\sum_i\lambda_i|e_i\rangle\langle e_i|.\end{aligned}$$

[1]Формально, для того чтобы это доказательство было верным, нам нужно показать, что $f(x)$ является вогнутой при $x=0$. (ρ и σ могут иметь некоторые обращающиеся в нуль собственные значения.) Вы можете проверить самостоятельно, что при этом $f(x)$ остается вогнутой функцией.

Субаддитивность энтропии этой системы доказывает общую вогнутость $S(\boldsymbol{\rho})$:

$$S(\boldsymbol{\rho}_{AB}) \leqslant S(\boldsymbol{\rho}_A) + S(\boldsymbol{\rho}_B),$$

$$-\operatorname{tr}\left[\left(\sum_i \lambda_i (\boldsymbol{\rho}_i)_A \otimes (|e_i\rangle\langle e_i|)_B\right) \log\left(\sum_j \lambda_j (\boldsymbol{\rho}_j)_A \otimes (|e_j\rangle\langle e_j|)_B\right)\right] \leqslant$$

$$\leqslant S\left(\sum_i \lambda_i \boldsymbol{\rho}_i\right) - \operatorname{tr}\left[\sum_i \lambda_i |e_i\rangle\langle e_i| \log \sum_i \lambda_i |e_i\rangle\langle e_i|\right].$$

Если мы берем след по системе B в базисе $|e_i\rangle$, то сумма по j сводится к одному ее слагаемому с $i = j$. Упростим левую часть (LHS) этого неравенства:[1]

$$\begin{aligned}
\text{LHS} &= -\operatorname{tr}\left[\left(\sum_i \lambda_i \boldsymbol{\rho}_i \otimes |c_i\rangle\langle e_i|\right) \log\left(\lambda_i \boldsymbol{\rho}_i \otimes |e_i\rangle\langle e_i|\right)\right] = \\
&= -\operatorname{tr}\left[\left(\sum_i \lambda_i \boldsymbol{\rho}_i \otimes |e_i\rangle\langle e_i|\right)\left(\log \boldsymbol{\rho}_i \otimes \mathbf{1} + \mathbf{1} \otimes \log \lambda_i |e_i\rangle\langle e_i|\right)\right] = \\
&= -\operatorname{tr}\left[\sum_i \lambda_i \boldsymbol{\rho}_i \log \boldsymbol{\rho}_i \otimes |e_i\rangle\langle e_i|\right] - \\
&\quad -\operatorname{tr}\left[\sum_i \lambda_i \boldsymbol{\rho}_i \otimes |e_i\rangle\langle e_i| \log\left(\lambda_i |e_i\rangle\langle e_i|\right)\right] = \\
&= \sum_i \lambda_i S(\boldsymbol{\rho}_i) - \operatorname{tr} \boldsymbol{\rho}_i \left[\sum_i \lambda_i |e_i\rangle\langle e_i| \log\left(\lambda_i |e_i\rangle\langle e_i|\right)\right] = \\
&= \sum_i \lambda_i S(\boldsymbol{\rho}_i) - \sum_i \lambda_i \log \lambda_i.
\end{aligned}$$

[1] Здесь в ходе преобразований используется равенство

$$\begin{aligned}
\mathbf{P} \log \mathbf{P} &= \mathbf{P} \log(\mathbf{1} - \mathbf{Q}) = -\mathbf{P}\left(\mathbf{Q} + \tfrac{1}{2}\mathbf{Q}^2 + \tfrac{1}{3}\mathbf{Q}^3 + \ldots\right) \\
&= -\mathbf{P}\mathbf{Q}\left(1 + \tfrac{1}{2} + \tfrac{1}{3} + \ldots\right) = 0,
\end{aligned}$$

где $\mathbf{P} = |e\rangle\langle e|$, $\mathbf{Q} = \mathbf{1} - \mathbf{P}$ – проекторы на взаимно ортогональные подпространства. – *Прим. ред.*

Подставляя этот результат обратно в неравенство субаддитивности, мы получим условие вогнутости:

$$\sum_i \lambda_i S(\rho_i) - \sum_i \lambda_i \log \lambda_i \leqslant S\left(\sum_i \lambda_i \rho_i\right) - \sum_i \lambda_i \log \lambda_i,$$

$$\sum_i \lambda_i S(\rho_i) \leqslant S\left(\sum_i \lambda_i \rho_i\right).$$

5.3. Монотонность Линдблада — Ульмана

Используя некоторые разработанные в задаче 5.2 приемы, мы находим, что свойство монотонности позволяет вывести некоторые очень полезные результаты.

a) Применяя свойство монотонности к состоящей из трех частей системе, получим свойство строгой субаддитивности:

$$\begin{aligned}
S(\rho_{AB}|\rho_A \otimes \rho_B) &\leqslant S(\rho_{ABC}|\rho_A \otimes \rho_{BC}),\\
-S(\rho_{AB}) - \operatorname{tr}\left[\rho_{AB} \log(\rho_A \otimes \rho_B)\right] &\leqslant -S(\rho_{ABC})\\
&\quad - \operatorname{tr}\left[\rho_{ABC} \log(\rho_A \otimes \rho_{BC})\right],\\
-S(\rho_{AB}) + S(\rho_A) + S(\rho_B) &\leqslant -S(\rho_{ABC}) + S(\rho_A) + S(\rho_{BC}),\\
-S(\rho_{AB}) + S(\rho_B) &\leqslant -S(\rho_{ABC}) + S(\rho_{BC}),\\
S(\rho_{ABC}) + S(\rho_B) &\leqslant S(\rho_{AB}) + S(\rho_{BC}).
\end{aligned}$$

b) Действие супероператора \$ на матрицу плотности ρ_A (σ_A) можно представить, как вычисление следа по окружению после приведения его в контакт с системой A и совместной с ρ_A (σ_A) унитарной эволюции:

$$\begin{aligned}
\rho_{AB} &= U\left(\rho_A \otimes (|e\rangle\langle e|)_B\right)U^{-1},\\
\sigma_{AB} &= U\left(\sigma_A \otimes (|e\rangle\langle e|)_B\right)U^{-1},\\
\$\rho_A &= \operatorname{tr}_B \rho_{AB},\\
\$\sigma_A &= \operatorname{tr}_B \sigma_{AB}.
\end{aligned}$$

Энтропия фон Неймана матрицы плотности инвариантна относительно унитарной эволюции или присоединения чистого состояния. Опуская для простоты индексы чистого состояния $|e\rangle$ и унитарной матрицы U, мы видим:

$$\begin{aligned}
S(\boldsymbol{\rho}_{AB}|\boldsymbol{\sigma}_{AB}) &= \operatorname{tr}\left[U(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)U^{-1} \log\left(U(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)U^{-1}\right)\right] - \\
&- \operatorname{tr}\left[U(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)U^{-1} \log\left(U(\boldsymbol{\sigma}_A \otimes |e\rangle\langle e|)U^{-1}\right)\right] = \\
&= \operatorname{tr}\left[U(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)U^{-1} U \log(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)U^{-1}\right] - \\
&- \operatorname{tr}\left[U(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)U^{-1} U \log(\boldsymbol{\sigma}_A \otimes |e\rangle\langle e|)U^{-1}\right] = \\
&= \operatorname{tr}\left[(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|) \log(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|)\right] - \\
&- \operatorname{tr}\left[(\boldsymbol{\rho}_A \otimes |e\rangle\langle e|) \log(\boldsymbol{\sigma}_A \otimes |e\rangle\langle e|)\right] = \\
&= \operatorname{tr}(\boldsymbol{\rho}_A \log \boldsymbol{\rho}_A) - \operatorname{tr}(\boldsymbol{\rho}_A \log \boldsymbol{\sigma}_A) = \\
&= S(\boldsymbol{\rho}_A|\boldsymbol{\sigma}_A).
\end{aligned}$$

Вместе с монотонностью Линдблада–Ульмана это дает искомый результат

$$S(\$\boldsymbol{\rho}_A|\$\boldsymbol{\sigma}_A \leqslant S(\boldsymbol{\rho}_{AB}|\boldsymbol{\sigma}_{AB}) \leqslant S(\boldsymbol{\rho}_A|\boldsymbol{\sigma}_A).$$

c) Рассмотрим матрицы плотности, определенные соотношениями

$$\begin{aligned}
\boldsymbol{\rho}_{AB} &= \sum_x p_x(\boldsymbol{\rho}_x)_A \otimes (|e_x\rangle\langle e_x|)_B, \\
\boldsymbol{\rho}_A &= \sum_x p_x \boldsymbol{\rho}_x, \\
\boldsymbol{\rho}_B &= \sum_x p_x |e_x\rangle\langle e_x|.
\end{aligned}$$

Относительная энтропия $S(\boldsymbol{\rho}_{AB}|\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B)$ в точности совпадает с информацией Холево $\chi(\mathcal{E})$ (для краткости опущены индексы подсистем A и B):

$$\begin{aligned}
S(\boldsymbol{\rho}_{AB}|\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B) = \operatorname{tr}&\left[\left(\sum_x p_x \boldsymbol{\rho}_x \otimes |e_x\rangle\langle e_x|\right) \log\left(\sum_y p_y \boldsymbol{\rho}_y \otimes |e_y\rangle\langle e_y|\right)\right] - \\
&- \operatorname{tr}\left[\left(\sum_x p_x \boldsymbol{\rho}_x \otimes |e_x\rangle\langle e_x|\right) \log\left(\sum_w p_w \boldsymbol{\rho}_w \otimes \sum_z p_z |e_z\rangle\langle e_z|\right)\right].
\end{aligned}$$

Так же как и в задаче 5.2 (**f**), мы можем взять след по состояниям системы B в базисе $|e_x\rangle$, сводя суммы по y и z к одному слагаемому каждую:

$$\begin{aligned}
S(\boldsymbol{\rho}_{AB}|\boldsymbol{\rho}_A \otimes \boldsymbol{\rho}_B) = &- \operatorname{tr}\left[\sum_x p_x \boldsymbol{\rho}_x \log \boldsymbol{\rho}_x \otimes |e_x\rangle\langle e_x|\right] + \\
&+ \operatorname{tr}\left[\sum_x p_x \boldsymbol{\rho}_x \otimes |e_x\rangle\langle e_x| \log(p_x|e_x\rangle\langle e_x|)\right] - \\
&- \operatorname{tr}\left[\sum_x p_x \boldsymbol{\rho}_x \log\left(\sum_w p_w \boldsymbol{\rho}_w \otimes |e_x\rangle\langle e_x|\right)\right] - \\
&- \operatorname{tr}\left[\sum_x p_x \boldsymbol{\rho}_x \otimes \log\left(p_x|e_x\rangle\langle e_x|\right)\right] = \\
= &- \sum_x p_x S(\boldsymbol{\rho}_x) - \sum_x p_x \log p_x + \\
&+ S\left(\sum_x p_x \boldsymbol{\rho}_x\right) + \sum_x p_x \log p_x = \\
= &- \sum_x p_x S(\boldsymbol{\rho}_x) + S\left(\sum_x p_x \boldsymbol{\rho}_x\right) = \\
= &\, \chi(\mathcal{E}).
\end{aligned}$$

5.4. ПОЗМ Переса — Вутерса

а) Записанные в обозначениях Дирака сигнальные состояния Алисы имеют вид

$$\begin{aligned}
|\varphi_1\rangle &= |\uparrow\rangle, \\
|\varphi_2\rangle &= -\frac{1}{2}\left(|\uparrow\rangle - \sqrt{3}|\downarrow\rangle\right), \\
|\varphi_3\rangle &= -\frac{1}{2}\left(|\uparrow\rangle + \sqrt{3}|\downarrow\rangle\right).
\end{aligned}$$

Дираковские обозначения позволяют очень быстро выразить состояния $|\Phi_i\rangle = |\varphi_i\rangle|\varphi_i\rangle$ $(i = 1, 2, 3)$ в базисе Белла:

$$
\begin{aligned}
|\Phi_1\rangle &= |\uparrow\uparrow\rangle = \\
&= \frac{1}{\sqrt{2}}\Big(|\Phi^+\rangle + |\Phi^-\rangle\Big), \\
|\Phi_2\rangle &= \frac{1}{4}\Big(|\uparrow\rangle - \sqrt{3}|\downarrow\rangle\Big)\Big(|\uparrow\rangle - \sqrt{3}|\downarrow\rangle\Big) = \\
&= \frac{1}{4}\Big(|\uparrow\uparrow\rangle + 3|\downarrow\downarrow\rangle - \sqrt{3}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\Big) = \\
&= \frac{1}{4}\Big(2(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle) - (|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle) - \sqrt{3}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\Big) = \\
&= \frac{1}{2\sqrt{2}}\Big(2|\Phi^+\rangle - |\Phi^-\rangle - \sqrt{3}|\Psi^+\rangle\Big), \\
|\Phi_3\rangle &= \frac{1}{4}\Big(|\uparrow\rangle + \sqrt{3}|\downarrow\rangle\Big)\Big(|\uparrow\rangle + \sqrt{3}|\downarrow\rangle\Big) = \\
&- \frac{1}{4}\Big(|\uparrow\uparrow\rangle + 3|\downarrow\downarrow\rangle + \sqrt{3}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)\Big) = \\
&= \frac{1}{2\sqrt{2}}\Big(2|\Phi^+\rangle - |\Phi^-\rangle + \sqrt{3}|\Psi^+\rangle\Big),
\end{aligned}
$$

Матрица плотности, соответствующая приготовлению Алисы, представляет собой одну треть суммы проекторов

$$
|\Phi_1\rangle\langle\Phi_1| = \frac{1}{8}\begin{pmatrix} 4 & 4 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$

$$
|\Phi_2\rangle\langle\Phi_2| = \frac{1}{8}\begin{pmatrix} 4 & -2 & -2\sqrt{3} & 0 \\ -2 & 1 & \sqrt{3} & 0 \\ -2\sqrt{3} & \sqrt{3} & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$

$$
|\Phi_3\rangle\langle\Phi_3| = \frac{1}{8}\begin{pmatrix} 4 & -2 & 2\sqrt{3} & 0 \\ -2 & 1 & -\sqrt{3} & 0 \\ 2\sqrt{3} & -\sqrt{3} & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$

$$
\boldsymbol{\rho} = \operatorname{diag}\Big(1/2, 1/4, 1/4, 0\Big).
$$

Возможно это удивительно, что матрица плотности диагональна в базисе Белла. Следовательно, энтропией фон Неймана этого источника является

$$S(\boldsymbol{\rho}) = -\frac{1}{2}\log\frac{1}{2} - 2\left(\frac{1}{4}\log\frac{1}{4}\right) = \frac{3}{2}.$$

b) «Достаточно хорошее измерение» (ДХИ), которым следует пользоваться Бобу, чтобы декодировать сигнал Алисы, представляет собой ПОЗМ, определяемую операторами $\mathbf{F}_i = \mathbf{G}^{-1/2}|\Phi_i\rangle\langle\Phi_i|\mathbf{G}^{-1/2}$, где $\mathbf{G} = 3\boldsymbol{\rho}$:

$$\mathbf{F}_1 = \frac{1}{3}\begin{pmatrix} 1 & \sqrt{2} & 0 & 0 \\ \sqrt{2} & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{F}_2 = \frac{1}{6}\begin{pmatrix} 2 & -\sqrt{2} & -\sqrt{6} & 0 \\ -\sqrt{2} & 1 & \sqrt{3} & 0 \\ -\sqrt{6} & \sqrt{3} & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{F}_3 = \frac{1}{6}\begin{pmatrix} 2 & -\sqrt{2} & \sqrt{6} & 0 \\ -\sqrt{2} & 1 & -\sqrt{3} & 0 \\ \sqrt{6} & -\sqrt{3} & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Оказывается, что эта ПОЗМ в действительности является *ортогональным* измерением, определяющими состояниями которого служат

$$|\Psi_1\rangle = \frac{1}{\sqrt{3}}\Big(|\Phi^+\rangle + \sqrt{2}|\Phi^-\rangle\Big),$$

$$|\Psi_2\rangle = \frac{1}{\sqrt{6}}\Big(\sqrt{2}|\Phi^+\rangle - |\Phi^-\rangle - \sqrt{3}|\Psi^+\rangle\Big),$$

$$|\Psi_3\rangle = \frac{1}{\sqrt{6}}\Big(\sqrt{2}|\Phi^+\rangle - |\Phi^-\rangle + \sqrt{3}|\Psi^+\rangle\Big),$$

$$|\Psi_4\rangle = |\Psi^-\rangle, \quad \text{(для полноты базиса)}.$$

c) Как и в задаче 5.1, чтобы вычислить взаимную информацию между приготовлением Алисы и измерением Боба, мы должны начать с вычисления

вероятностей. В общем случае они представляют собой

$$\begin{aligned} p\,(\text{Боб измеряет } |\Psi_b\rangle|, \text{Алиса готовит } |\Phi_a\rangle) &= \langle\Phi_a|\mathbf{F}_b|\Phi_a\rangle \\ &= \langle\Phi_a|\Psi_b\rangle\langle\Psi_b|\Phi_a\rangle \\ &= |\langle\Psi_b|\Phi_a\rangle|^2. \end{aligned}$$

Последовательно вычисляя их для каждых a и b, мы действительно найдем цитированный в лекциях результат:

$$\begin{aligned} p(a|a) &= \frac{1}{3}\left(1+\frac{1}{\sqrt{2}}\right)^2, \\ p(b|a) &= \frac{1}{6}\left(1-\frac{1}{\sqrt{2}}\right)^2, \qquad b \neq a. \end{aligned}$$

Поскольку каждая вероятность здесь сводится к одному из двух значений, то нетрудно вычислить приобретаемую Бобом информацию:

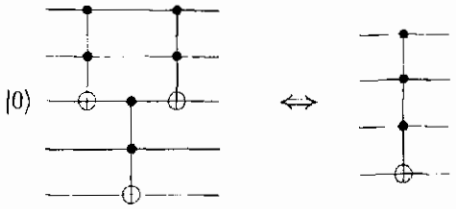
$$\begin{aligned} H(B) &= -[p(1|1)+2p(1|2)]\log\left[\frac{1}{3}\big(p(1|1)+2p(1|2)\big)\right] \\ &= -1\log\frac{1}{3} \approx 1{,}585 \\ H(B|A) &= -p(1|1)\log p(1|1) - 2p(1|2)\log p(1|2) \\ &= -\frac{1}{3}\left(1+\frac{1}{\sqrt{2}}\right)^2 \log\left[\frac{1}{3}\left(1+\frac{1}{\sqrt{2}}\right)^2\right] - \\ &\quad -\frac{1}{3}\left(1-\frac{1}{\sqrt{2}}\right)^2 \log\left[\frac{1}{3}\left(1-\frac{1}{\sqrt{2}}\right)^2\right] \\ &\approx 0{,}215893, \\ I(B;A) &\approx 1{,}36907. \end{aligned}$$

Решения упражнений к главе 6

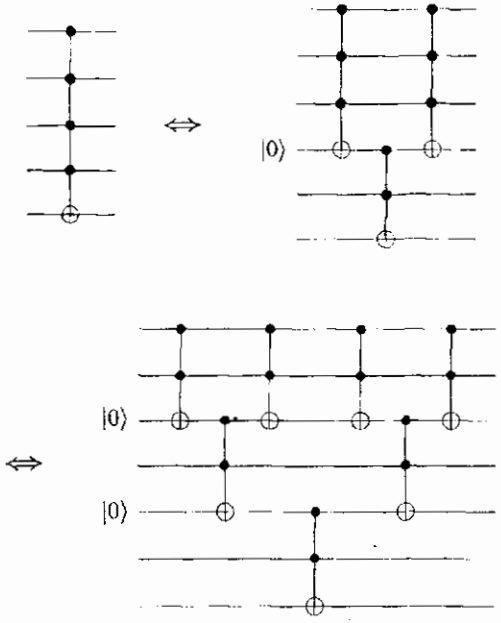
6.1. Линейное моделирование вентиля Тоффоли

а) Рассмотрим описанную в лекциях схему, которая выполняет $\theta^{(4)}$, используя только компоненты $\theta^{(3)}$ и один бит вспомогательного пространства,

первоначально полагаемый равным нулю

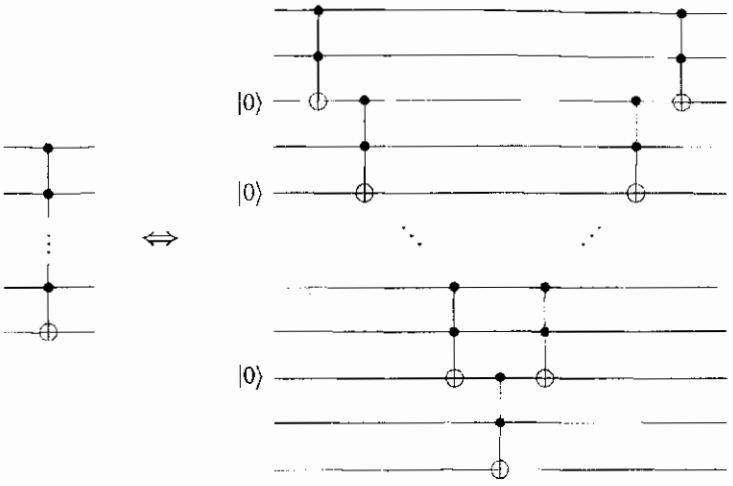


Используя конструкцию для вентиля $\theta^{(4)}$, мы можем рекурсивно построить вентиль $\theta^{(5)}$:



Заметим, что, благодаря тождеству $\left(\theta^{(3)}\right)^2 = \mathbf{1}$, два вентиля во внутренней части последней диаграммы необязательны. Продолжая рекурсивным образом эту процедуру, мы, очевидно, подобным образом можем исключить все внутренние вентили Тоффоли, получая в результате

конструкцию:



Поскольку каждой контрольной линии вентиля $\theta^{(n)}$, исключая линии 1, 2 и $2n-3$, сопоставлен вспомогательный бит, то для реализации этой схемы необходимо $n-3$ вспомогательных бита. Далее, поскольку каждый вспомогательный бит является целью двух вентилей $\theta^{(3)}$ (после вычисления необходимо восстановить исходное значение вспомогательного бита) и так как целевой бит вентиля $\theta^{(n)}$ в свою очередь также является целью вентиля $\theta^{(3)}$, то всего для этой конструкции необходимо $2(n-3)+1 = 2n-5$ вентилей $\theta^{(3)}$.

b) Описанный в **(a)** каскад вентилей Тоффоли отображает целевой бит y на

$$
\begin{aligned}
y \;\to\; & ((((s_1 \oplus x_1x_2)x_3 \oplus s_2)x_4 \oplus s_3 \ldots)x_{n-2} \oplus s_{n-3})x_{n-1} \oplus y \\
= \; & [s_1x_3x_4 \ldots x_{n-1} \oplus s_2x_4x_5 \ldots x_{n-1} \oplus \ldots \oplus s_{n-3}x_{n-1}] \\
& \oplus x_1 \ldots x_{n-1} \oplus y \\
= \; & [((((s_2 \oplus s_1x_3)x_4 \oplus s_3)x_5 \oplus s_4 \ldots)x_{n-2} \oplus s_{n-3})x_{n-1}] \\
& \oplus x_1 \ldots x_{n-1} \oplus y,
\end{aligned}
$$

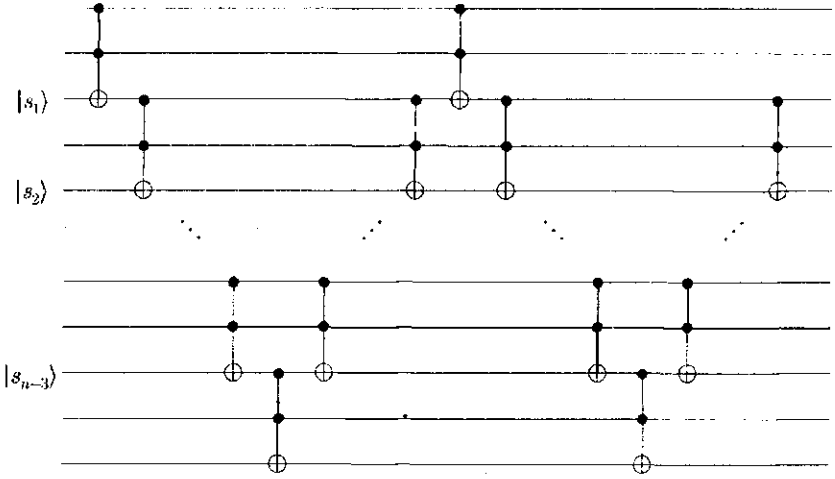
где s_i помечают вспомогательные биты, а x_i — контрольные линии.

Если все вспомогательные биты первоначально равны нулю, то заключенное в квадратные скобки слагаемое тождественно обращается в нуль.

Чтобы нейтрализовать влияние некоторых вспомогательных битов, не равных нулю в начальном состоянии, мы можем вычислить слагаемое в квадратных скобках $[\cdot]$ отдельно и подставить его XOR (то есть $[\cdot] \Rightarrow [\cdot] \oplus [\cdot]$) в окончательный результат

$$\begin{aligned} y \;&\rightarrow\; [\cdot] \oplus [\cdot] \oplus x_1 \ldots x_{n-1} \oplus y = \\ &= x_1 \ldots x_{n-1} \oplus y = \\ &= \theta^{(n)}(y, x_1 \ldots x_{n-1}). \end{aligned}$$

Конечно, мы должны быть внимательны к XOR в этом новом слагаемом *после* того, как восстановили значения вспомогательных битов. Наконец, как и в конструкции части **(а)**, нам нужно восстановить вспомогательные биты, на которые повлияла эта новая схема. Таким образом, модифицированный массив вентилей напоминает банни Бробдингнега[1]



Следовательно, количество вентилей в этом семействе схем равно $2n - 5 + 2(n-1) - 5 = 4n - 12$.

6.2. Набор универсальных квантовых вентилей

а) Используя указание, мы можем записать каждое из преобразований A, B, C, и U в виде произведений трех поворотов. Более того, нам известно,

[1] Бробдингнег – придуманная Джонатаном Свифтом фантастическая страна великанов, в которую попал Гулливер во время своего второго путешествия. – *Прим. перев.*

что сопряжение матрицей σ_x меняет знак угла матрицы поворота. Чтобы не переписывать всякий раз букву $\mathbf{R}$, я буду обозначать $\mathbf{R}_z(\psi)$ символом $\mathbf{Z}_\psi$, $\mathbf{R}_y(\theta)$ — символом $\mathbf{Y}_\theta$, а σ_x — символом $\mathbf{X}$. Принимая эти обозначения, можно записать

$$\begin{aligned}\mathbf{A} &= \mathbf{Z}_\alpha\mathbf{Y}_\beta\mathbf{Z}_\gamma,\\ \mathbf{B} &= \mathbf{Z}_\delta\mathbf{Y}_\varepsilon\mathbf{Z}_\varphi,\\ \mathbf{C} &= \mathbf{Z}_\lambda\mathbf{Y}_\mu\mathbf{Z}_\nu,\\ \mathbf{U} &= \mathbf{Z}_\psi\mathbf{Y}_\theta\mathbf{Z}_\phi.\end{aligned}$$

Подобно этому уравнения связи представляются в виде

$$\begin{aligned}\mathbf{1} &= \mathbf{Z}_\alpha\mathbf{Y}_\beta\mathbf{Z}_\gamma\mathbf{Z}_\delta\mathbf{Y}_\varepsilon\mathbf{Z}_\varphi\mathbf{Z}_\lambda\mathbf{Y}_\mu\mathbf{Z}_\nu =\\ &= \mathbf{Z}_\alpha\mathbf{Y}_\beta\mathbf{Z}_{\gamma+\delta}\mathbf{Y}_\varepsilon\mathbf{Z}_{\varphi+\lambda}\mathbf{Y}_\mu\mathbf{Z}_\nu,\\ \mathbf{Z}_\psi\mathbf{Y}_\theta\mathbf{Z}_\phi &= (\mathbf{Z}_\alpha\mathbf{Y}_\beta\mathbf{Z}_\gamma)\mathbf{X}(\mathbf{Z}_\delta\mathbf{X}\mathbf{X}\mathbf{Y}_\varepsilon\mathbf{X}\mathbf{X}\mathbf{Z}_\varphi)\mathbf{X}(\mathbf{Z}_\lambda\mathbf{Y}_\mu\mathbf{Z}_\nu) =\\ &= \mathbf{Z}_\alpha\mathbf{Y}_\beta\mathbf{Z}_{\gamma-\delta}\mathbf{Y}_{-\varepsilon}\mathbf{Z}_{-\varphi+\lambda}\mathbf{Y}_\mu\mathbf{Z}_\nu.\end{aligned}$$

С этого момента у нас два уравнения с девятью неизвестными. Найти решение будет большой удачей, не правда ли? Конечно, на самом деле это *матричные* уравнения, следовательно, имеется самое большее восемь уравнений, связывающих эти переменные, при условии, что ни одно из них не эквивалентно другому.

Тем не менее посмотрим, сможем ли мы найти решение для углов, не копаясь в матричных элементах. Непосредственно проверяется, что одно из решений первого уравнения связи возникает, если соседние матрицы определяют повороты на равные углы в противоположных друг другу направлениях, то есть когда углы удовлетворяют условиям

$$\begin{aligned}\varphi &= -\lambda, \qquad \varepsilon + \mu = -\beta,\\ \gamma &= -\delta, \qquad \alpha = -\nu.\end{aligned}$$

Этот выбор позволяет записать второе уравнение связи в виде

$$\mathbf{Z}_\alpha\mathbf{Y}_{-(\varepsilon+\mu)}\mathbf{Z}_{2\gamma}\mathbf{Y}_{-\varepsilon}\mathbf{Z}_{2\lambda}\mathbf{Y}_\mu\mathbf{Z}_{-\alpha} = \mathbf{Z}_\psi\mathbf{Y}_\theta\mathbf{Z}_\phi.$$

Выбор $\alpha = \psi$ согласует друг с другом последние повороты вокруг оси z в обеих частях этого равенства. Тогда $\mathbf{Z}_\phi$ согласуется при дополнительном выборе $\gamma = \mu = 0$:

$$\mathbf{Z}_\alpha\mathbf{Y}_{-2\varepsilon}\mathbf{Z}_{2\lambda-\alpha} = \mathbf{Z}_\psi\mathbf{Y}_\theta\mathbf{Z}_\phi.$$

Теперь мы имеем три уравнения с тремя неизвестными (прогресс!), которые имеют решение

$$\alpha = \psi,$$
$$\lambda = \frac{1}{2}(\phi + \psi),$$
$$\varepsilon = -\frac{\theta}{2}.$$

Подставляя их все вместе в матрицы $\mathbf{A}$, $\mathbf{B}$ и $\mathbf{C}$, мы находим, что они удовлетворяют условиям

$$\mathbf{A} = \mathbf{Z}_{\psi}\mathbf{Y}_{\theta/2},$$
$$\mathbf{B} = \mathbf{Y}_{\theta/2}\mathbf{Z}_{-(\phi+\psi)/2},$$
$$\mathbf{C} = \mathbf{Z}_{(\phi-\psi)/2}$$

или в более общепринятых обозначениях:

$$\mathbf{A} = \mathbf{R}_z(\psi)\mathbf{R}_y(\theta/2),$$
$$\mathbf{B} = \mathbf{R}_y(-\theta/2)\mathbf{R}_z(-(\phi+\psi)/2),$$
$$\mathbf{C} = \mathbf{R}_z((\phi-\psi)/2).$$

b) Чтобы показать, что вентиль контролируемой фазы $\mathbf{P} = \operatorname{diag}(1, 1, e^{i\alpha}, e^{i\alpha})$ является однокубитовым вентилем, мы должны показать, что $\mathbf{P} = \mathbf{V} \otimes \mathbf{W}$ для унитарных матриц $\mathbf{V}$ и $\mathbf{W}$. Непосредственной проверкой можно убедиться в том, что $\mathbf{P}$ разлагается как $\mathbf{P} = \mathbf{Z}(\alpha) \otimes \mathbf{1}$, где $\mathbf{Z}(\alpha) \equiv \operatorname{diag}(1, e^{i\alpha})$.

Однако при более строгом подходе мы должны доказать следующее. Матрица $\mathbf{P}$ диагональна, и если она разлагается, то тоже на диагональные матрицы. Более того, поскольку $\{\{\mathbf{R}_z(\theta), \mathbf{1}\} | \theta \in [0, 2\pi]\}$ образует базис для диагональных матриц в $\mathrm{SU}(2)$, то в самом общем виде разложимая матрица $\mathbf{P}$ может быть записана в виде суммы четырех слагаемых:

$$\mathbf{P} = \lambda_1 \mathbf{1} \otimes \mathbf{1} + \lambda_2 \mathbf{1} \otimes \mathbf{R}_z(\theta) + \lambda_3 \mathbf{R}_z(\varphi) \otimes \mathbf{1} + \lambda_4 \mathbf{R}_z(\psi) \otimes \mathbf{R}_z(\chi),$$

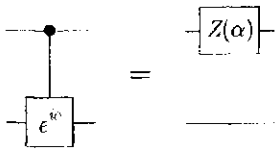
где $\lambda_i \in \mathbb{C}$ удовлетворяют условию $|\sum_i \lambda_i|^2 = 1$.

Так как $\mathbf{P}$ одинаково действует на состояния $|00\rangle$ и $|11\rangle$, то $\lambda_2 = \lambda_4 = 0$. Аналогично, так как $\mathbf{P}$ по разному действует на состояния $|00\rangle$ и $|01\rangle$, то $\lambda_1 = 0$. Следовательно, оставшийся коэффициент в наиболее общем случае представляет собой фазовый множитель $e^{i\eta}$. Таким образом,

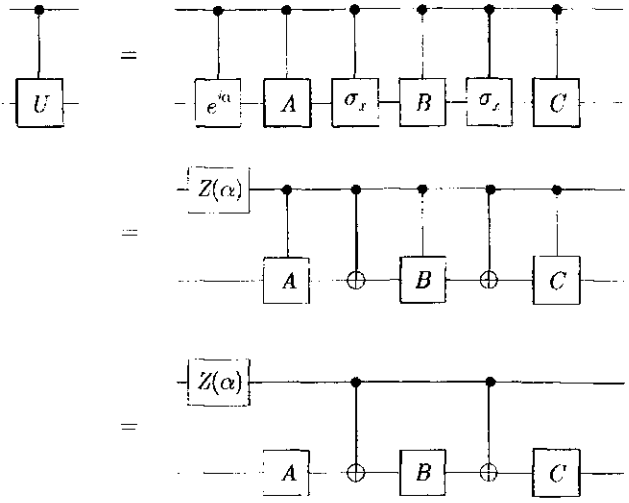
$$\mathbf{P} = e^{i\eta}\mathbf{R}_z(\varphi) \otimes \mathbf{1},$$
$$\operatorname{diag}(1, 1, e^{i\alpha}, e^{i\alpha}) = e^{i\eta}\operatorname{diag}(e^{i\varphi/2}, e^{i\varphi/2}, e^{-i\varphi/2}, e^{-i\varphi/2}),$$

что имеет решение $2\eta = -\varphi = \alpha$.

Следовательно, мы видим, что $\mathbf{P}$ действительно разложима, причем $\mathbf{P} = e^{i\alpha/2}\mathbf{R}_z(-\alpha) \otimes \mathbf{1}$. Общую фазу можно включить в состав другого сомножителя тензорного произведения, но, вероятно, самым естественным является упомянутый выше способ $\mathbf{P} = \mathbf{Z}(\alpha) \otimes \mathbf{1}$:



c) Произвольный элемент $SU(2)$ может быть записан как $\mathbf{A}\sigma_x\mathbf{B}\sigma_x\mathbf{C}$, где $\mathbf{A}$, $\mathbf{B}$ и $\mathbf{C}$ определены в части **(a)**. Следовательно, *произвольное* унитарное 2×2-преобразование $\mathbf{U} \in U(2)$ может быть записано как $\mathbf{U} = e^{i\alpha/2}\mathbf{A}\sigma_x\mathbf{B}\sigma_x\mathbf{C}$. Используя полученные в части **(a)** тождества, мы можем записать контролируемое $\mathbf{U}$ как последовательность однокубитовых вентилей и вентилей контролируемое NOT:



6.3. Точность

a) В этой задаче имеются некоторые трудности, поскольку на операторы не наложено никаких ограничений[1]. В частности, соотношения из условия

[1]Формально, для того чтобы нормы были определены, мы должны потребовать, чтобы $\mathbf{A}$ был компактным, а $\mathbf{B}$ – ограниченным. [Фактически приводимое ниже решение требует компактности обоих операторов, $\mathbf{A}$ и $\mathbf{B}$. – *Прим. ред.*]

задачи справедливы и для *недиагонализуемых* матриц, для которых мы не можем рассматривать $\|\mathbf{A}\|_{\mathrm{tr}}$ как сумму модулей собственных значений $\mathbf{A}$. Например, матрица

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

имеет $\|\mathbf{A}\|_{\mathrm{tr}} = 1$.

Тем не менее мы можем добиться успеха, покопавшись поглубже в закоулках нашей памяти и вспомнив *теорему о полярном разложении* из линейной алгебры, которая утверждает, что для любой матрицы $\mathbf{A}$ существует единственная унитарная матрица $\boldsymbol{U}$ такая, что

$$\begin{aligned} \mathbf{A} &= \mathbf{U}\sqrt{\mathbf{A}^\dagger\mathbf{A}}, \\ &= \mathbf{U}|\mathbf{A}|, \end{aligned}$$

где введено краткое обозначение $|\mathbf{A}|$ для $\sqrt{\mathbf{A}^\dagger\mathbf{A}}$. Это разложение полезно, поскольку по построению $|\mathbf{A}|$ является самосопряженной и, следовательно, диагонализуемой ортогональным базисом $\{|i\rangle\}$ с соответствующими собственными значениями (так называемые сингулярные числа матрицы $\mathbf{A}$) $|a_i|$. Используя это разложение, мы найдем:

$$\begin{aligned} |\operatorname{tr}\mathbf{A}| = |\operatorname{tr}\mathbf{U}|\mathbf{A}|\,| &= \left|\sum_i \langle i|\mathbf{U}|a_i|\,|i\rangle\right| \leqslant \\ &\leqslant \sum_i \left|\langle i|\mathbf{U}|a_i|\,|i\rangle\right| = \sum_i |a_i|\cdot|\langle i|\mathbf{U}|i\rangle| \leqslant \\ &\leqslant \sum_i |a_i| = \sum_i \langle i|\,|\mathbf{A}|\,|i\rangle = \operatorname{tr}|\mathbf{A}| = \|\mathbf{A}\|_{\mathrm{tr}}, \end{aligned}$$

где мы использовали тот факт, что унитарная матрица преобразует один ортогональный базис в другой, так что $|\langle i|\mathbf{U}|i\rangle| \leqslant 1$.

Другую часть этой задачи сложно проверить из-за порядка матриц $\mathbf{A}$ и $\mathbf{B}$ под знаком следовой нормы. Проще сначала показать, что $\|\mathbf{AB}\|_{\mathrm{tr}} \leqslant \|\mathbf{A}\|\cdot\|\mathbf{B}\|_{\mathrm{tr}}$, а затем обратиться к полярному разложению, чтобы получить искомое ограничение.

Вычисляя след в базисе, диагонализующем $\mathbf{B}^\dagger\mathbf{B}$ (то есть $\mathbf{B}^\dagger\mathbf{B} = \sum_i \lambda_i |i\rangle\langle i|$), мы находим

$$\begin{aligned} \|\mathbf{AB}\|_{\mathrm{tr}} &= \sum_i \langle i|\,|\mathbf{AB}|\,|i\rangle = \sum_i |\langle i|\,|\mathbf{AB}|\,|i\rangle| = \\ &= \sum_i \sqrt{|\langle i|\,|\mathbf{AB}|\,|i\rangle|^2} \leqslant \end{aligned}$$

$$\leqslant \sum_i \sqrt{\sum_j |\langle j||\mathbf{AB}||i\rangle|^2} = \sum_i \sqrt{\sum_j \langle i||\mathbf{AB}||j\rangle\langle j||\mathbf{AB}||i\rangle} =$$

$$= \sum_i \sqrt{\langle i||\mathbf{AB}|^2|i\rangle} = \sum_i \sqrt{\langle i|\mathbf{B}^\dagger\mathbf{A}^\dagger\mathbf{AB}|i\rangle} =$$

$$= \sum_i \sqrt{\frac{\langle i|\mathbf{B}^\dagger\mathbf{A}^\dagger\mathbf{AB}|i\rangle}{\langle i|\mathbf{B}^\dagger\mathbf{B}|i\rangle}\langle i|\mathbf{B}^\dagger\mathbf{B}|i\rangle} = \sum_i \frac{\|\mathbf{AB}|i\rangle\|}{\|\mathbf{B}|i\rangle\|}\sqrt{\langle i|\mathbf{B}^\dagger\mathbf{B}|i\rangle} \leqslant$$

$$\leqslant \sum_i \sup_{\mathbf{B}|i\rangle} \frac{\|\mathbf{AB}|i\rangle\|}{\|\mathbf{B}|i\rangle\|}\sqrt{\langle i|\mathbf{B}^\dagger\mathbf{B}|i\rangle} = \|\mathbf{A}\| \sum_i \sqrt{\langle i|\mathbf{B}^\dagger\mathbf{B}|i\rangle} =$$

$$= \|\mathbf{A}\| \sum_i \sqrt{\lambda_i} = \|\mathbf{A}\| \sum_i \langle i|\sqrt{\lambda_i}|i\rangle =$$

$$= \|\mathbf{A}\| \sum_i \langle i|\sqrt{\mathbf{B}^\dagger\mathbf{B}}|i\rangle = \|\mathbf{A}\| \cdot \|\mathbf{B}\|_{\mathrm{tr}}.$$

Как уже говорилось, это почти то, что требуется получить, но в неправильном порядке! Но мы можем привлечь полярное разложение, чтобы получить правильный порядок. Если положить $\mathbf{AB} = \mathbf{U}|\mathbf{AB}|$, то

$$\begin{aligned}\|\mathbf{AB}\|_{\mathrm{tr}} &= \mathrm{tr}\,|\mathbf{AB}| = \\ &= \mathrm{tr}(\mathbf{ABU}^{-1}) = \\ &= \mathrm{tr}(\mathbf{BU}^{-1}\mathbf{A}) = \\ &= \left|\,\mathrm{tr}(\mathbf{ABU}^{-1})\right|.\end{aligned}$$

Но согласно нашему первому результату

$$\left|\,\mathrm{tr}(\mathbf{ABU}^{-1})\right| \leqslant \|\mathbf{BU}^{-1}\mathbf{A}\|_{\mathrm{tr}},$$

а согласно второму —

$$\begin{aligned}\|\mathbf{BU}^{-1}\mathbf{A}\|_{\mathrm{tr}} &\leqslant \|\mathbf{BU}^{-1}\| \cdot \|\mathbf{A}\|_{\mathrm{tr}} = \\ &= \|\mathbf{B}\| \cdot \|\mathbf{A}\|_{\mathrm{tr}},\end{aligned}$$

то есть

$$\|\mathbf{AB}\|_{\mathrm{tr}} \leqslant \|\mathbf{B}\| \cdot \|\mathbf{A}\|_{\mathrm{tr}},$$

что и требовалось показать.

b) Этот результат вытекает из совершенных в части **(a)** подвигов Геракла

$$\begin{aligned}
\sum_a |P_a - \tilde{P}_a| &= \sum_a |\langle a|\boldsymbol{\rho}|a\rangle - \langle a|\tilde{\boldsymbol{\rho}}|a\rangle| = \sum_a |\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle| - \\
&= \sum_a \langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big] = \\
&= \sum_a \operatorname{tr}\left[(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}})|a\rangle\langle a|\right] \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big] = \\
&= \left|\sum_a \operatorname{tr}\left[(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}})|a\rangle\langle a|\right] \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big]\right| = \\
&= \left|\operatorname{tr}\left[\sum_a (\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}})|a\rangle\langle a| \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big]\right]\right| \leqslant \\
&\leqslant \left\|\sum_a (\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}) \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big] |a\rangle\langle a|\right\|_{\mathrm{tr}} = \\
&= \left\|(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}) \sum_a \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big] |a\rangle\langle a|\right\|_{\mathrm{tr}} \leqslant \\
&\leqslant \|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}} \cdot \left\|\sum_a \operatorname{sign}\big[\langle a|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}|a\rangle\big] |a\rangle\langle a|\right\| \leqslant \\
&\leqslant \|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}}.
\end{aligned}$$

Здесь при переходе к последней строке мы воспользовались тем, что норма оператора, собственные значения которого равны ± 1, ограничена сверху единицей.

c) Без потери общности можно записать $|\tilde{\psi}\rangle = \alpha|\psi\rangle + \beta|\psi^{\perp}\rangle$, где $\alpha, \beta \in \mathbb{C}$ и удовлетворяют условию $|\alpha|^2 + |\beta|^2 = 1$. Выражая $\boldsymbol{\rho}$ и $\tilde{\boldsymbol{\rho}}$ в базисе $\{|\psi\rangle, |\psi^{\perp}\rangle\}$, получим

$$\begin{aligned}
\|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}} &= \operatorname{tr}\sqrt{(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}})^{\dagger}(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}})} = \\
&= \operatorname{tr}\sqrt{(\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}})^2} = \\
&= \operatorname{tr}\sqrt{\begin{bmatrix} 1 - |\alpha|^2 & -\alpha\beta^* \\ -\alpha^*\beta & -|\beta|^2 \end{bmatrix}^2} =
\end{aligned}$$

$$
\begin{aligned}
&= \operatorname{tr}\sqrt{\begin{bmatrix} |\beta|^2 & -\alpha\beta^* \\ -\alpha^*\beta & -|\beta|^2 \end{bmatrix}^2} = \\
&= \operatorname{tr}\sqrt{\begin{bmatrix} |\beta|^4 + |\alpha|^2|\beta|^2 & 0 \\ 0 & |\beta|^4 + |\alpha|^2|\beta|^2 \end{bmatrix}} = \\
&= \operatorname{tr}\sqrt{\begin{bmatrix} |\beta|^2 & 0 \\ 0 & |\beta|^2 \end{bmatrix}} = \\
&= 2|\beta|.
\end{aligned}
$$

Однако при $\beta \neq 0$ расстояние между состояниями $|\psi\rangle$ и $|\tilde{\psi}\rangle$

$$
\begin{aligned}
\| |\psi\rangle - |\tilde{\psi}\rangle \| &= \|(1-\alpha)|\psi\rangle - \beta|\psi^{\perp}\rangle\| = \\
&= \sqrt{|1-\alpha|^2 + |\beta|^2} = \\
&= |\beta|\sqrt{1 + \frac{|1-\alpha|^2}{|\beta|^2}} \geqslant |\beta|,
\end{aligned}
$$

Таким образом, если $\beta \neq 0$, то

$$\|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}} \leqslant 2\| |\psi\rangle - |\tilde{\psi}\rangle\|.$$

А поскольку при $\beta = 0$

$$\|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}} = 0 \leqslant 2\| |\psi\rangle - |\tilde{\psi}\rangle\|,$$

то, используя результат части **(b)**, мы находим, что

$$\sum_a |P_a - \tilde{P}_a| \leqslant \|\boldsymbol{\rho} - \tilde{\boldsymbol{\rho}}\|_{\mathrm{tr}} \leqslant 2\| |\psi\rangle - |\tilde{\psi}\rangle\|.$$

6.4. Поиск в базе данных в непрерывном времени

a) Поскольку гамильтониан $\mathbf{H}$ не зависит от времени, формальное интегрирование нестационарного уравнения Шредингера дает

$$|\psi(T)\rangle = e^{-i\mathbf{H}T}|\psi_0\rangle.$$

В рассматриваемом случае $|\psi_0\rangle = |s\rangle$.

Гамильтониан $\mathbf{H}$ ограничивает эволюцию подпространством $\{|\omega\rangle, |s\rangle\}$ с (ненормированным) базисом собственных состояний $\{|s\rangle + |\omega\rangle, |s\rangle - |\omega\rangle\}$

(Представьте, например, сферу Блоха, чтобы убедиться в этом.) Вычисляя собственные значения $\mathbf{H}$, находим

$$\begin{aligned}\mathbf{H}(|s\rangle \pm |\omega\rangle) &= E(|\omega\rangle\langle\omega| + |s\rangle\langle s|)(|s\rangle \pm |\omega\rangle) = \\ &= E\big[|\omega\rangle\langle\omega|(|s\rangle \pm |\omega\rangle) + |s\rangle\langle s|(|s\rangle \pm |\omega\rangle)\big] = \\ &= E\big[(|s\rangle \pm |\omega\rangle) \pm 2^{-n/2}(|s\rangle \pm |\omega\rangle)\big] = \\ &= E(1 \pm 2^{-n/2})(|s\rangle \pm |\omega\rangle),\end{aligned}$$

где мы подставили перекрытие $\langle s|\omega\rangle = 2^{-n/2}$.

Записывая $|s\rangle = \frac{1}{2}(|s\rangle + |\omega\rangle) + \frac{1}{2}(|s\rangle - |\omega\rangle)$, представим состояние системы в момент времени T в виде

$$\begin{aligned}|\psi(T)\rangle &= e^{-i\mathbf{H}T}|s\rangle = \\ &= \frac{1}{2}e^{-iET(1+2^{-n/2})}(|s\rangle + |\omega\rangle) + \frac{1}{2}e^{-iET(1-2^{-n/2})}(|s\rangle - |\omega\rangle) = \\ &= e^{-iET}\left[\cos\left(\frac{ET}{2^{n/2}}\right)|s\rangle - i\sin\left(\frac{ET}{2^{n/2}}\right)|\omega\rangle\right].\end{aligned}$$

Чтобы оптимизировать вероятность успешного определения $|\omega\rangle$, необходимо максимизировать вероятность $p = |\langle\omega|\psi(T)\rangle|^2$. Непосредственная проверка показывает, что значение $p = 1$ достигается, если фазы собственных векторов различаются на 180°:

$$\begin{aligned}e^{-iET(1+2^{-n/2})} &= -e^{-iET(1-2^{-n/2})}, \\ ET \cdot 2^{-n/2} &= -ET \cdot 2^{-n/2} + (2k+1)\pi, \\ T &= \frac{(2k+1)\pi}{2E}2^{n/2}.\end{aligned}$$

Очевидно, нам хотелось бы определить $|\omega\rangle$ как можно быстрее, поэтому выбирем $k = 0$ и получим границу Гровера:

$$T = \frac{\pi}{2E}2^{n/2}.$$

b) Можно получить весьма общую границу квадратичного ускорения поиска в базе данных в непрерывном времени, подобно тому, как это было сделано для поиска в базе данных в дискретном времени. Фактически мы увидим, что это, по существу, та же самая граница.

Допустим, что мы имеем алгоритм $\mathcal{A}$, который применяет гамильтониан $\mathbf{H} = \mathbf{H}_\omega + \mathbf{H}'(t)$ к состоянию $|\psi_0\rangle$ и спустя время T со стопроцентной надежностью определяет состояние $|\omega\rangle$. Так как $|\omega\rangle$ может принять любое из 2^n различных значений, гильбертово пространство, содержащее результат вычисления $|\psi_T\rangle$, должно иметь размерность как минимум 2^n. Более того, так как $\mathcal{A}$ должен быть способным *идеально* различать все альтернативы, множество

$$\{|\psi_T^a\rangle \mid |\psi_T^a\rangle \text{ представляет ответ } \mathcal{A}(|\psi_0\rangle) = a\}$$

должно образовывать ортогональный базис.

Теперь рассмотрим (скорее «плохой») алгоритм $\mathcal{D}$, который пытается определить $|\omega\rangle$, путем применения лишь гамильтониана $\mathbf{H} = \mathbf{H}'(t)$ к состоянию $|\psi_0\rangle$ в течение времени T. Так как мы неявно предполагаем, что $\mathbf{H}'(t)$ не имеет определенной зависимости от ω, кажется невероятным, что алгоритм $\mathcal{D}$ будет успешным. Но в то же самое время похоже, что в среднем[1] результат $|\psi_T^\omega\rangle$ алгоритма $\mathcal{A}$ должен отличаться от результата $|\varphi_T\rangle$ алгоритма $\mathcal{D}$ на величину, ограниченную некоторой функцией от T. («В течение ограниченного интервала времени оракул [гамильтониан] может только увести нас от нашей плохой догадки».)

Мы можем усилить наше подозрение, выполнив такой же анализ, что и в дискретном случае: разобьем $\mathcal{A}$ на отдельные шаги и определим границу того, как далеко от $|\varphi_t\rangle$ берется $|\psi_t^\omega\rangle$ на каждом шаге. (Но теперь шаги инфинитезимальны!)

Унитарные операторы, которые применяются алгоритмами $\mathcal{A}$ и $\mathcal{D}$ на каждом «шаге времени», представляют собой

$$\mathcal{A} : |\psi_t^\omega\rangle \to d\mathbf{U}_t|\psi_t^\omega\rangle,$$
$$\mathcal{D} : |\varphi_t\rangle \to d\mathbf{U}_t'|\varphi_t\rangle.$$

Действие $\mathcal{A}$ на «плохое» состояние $|\varphi_t\rangle$ в момент времени t имеет вид

$$d\mathbf{U}_t|\varphi_t\rangle = |\varphi_t\rangle + |E_t\rangle,$$

где

$$\begin{aligned}|E_t\rangle &= (d\mathbf{U}_t - d\mathbf{U}_t')|\varphi_t\rangle = \\ &= \big((\mathbf{1} - i\mathbf{H}dt) - (\mathbf{1} - i\mathbf{H}'dt)\big)|\varphi_t\rangle = \\ &= -i(\mathbf{H} - \mathbf{H}')|\varphi_t\rangle dt = \\ &= -iE|\omega\rangle\langle\omega|\varphi_t\rangle dt.\end{aligned}$$

[1] Среднее здесь представляет собой среднее *по ансамблю* всех возможных значений $|\omega\rangle$.

Как и в дискретном случае, спустя время T мы получаем непрерывный аналог уравнения (6.65) из лекций:

$$
\begin{aligned}
|\psi_T^\omega\rangle &= |\varphi_T\rangle + |E_T\rangle + d\mathbf{U}_T|E_{T-dt}\rangle + \ldots + d\mathbf{U}_T \ldots d\mathbf{U}_0|E_0\rangle = \\
&= |\varphi_T\rangle + e^{-i\mathbf{H}\cdot 0}|E_T\rangle + e^{-i\mathbf{H}dt}|E_{T-dt}\rangle + \ldots + e^{-i\mathbf{H}T}|E_0\rangle = \\
&= |\varphi_T\rangle + \int_0^T e^{-i\mathbf{H}t}|E_{T-t}\rangle dt = \\
&= |\varphi_T\rangle - iE\int_0^T e^{-i\mathbf{H}t}|\omega\rangle\langle\omega|\psi_{T-t}\rangle dt = \\
&= |\varphi_T\rangle - iE\int_0^T e^{-i\mathbf{H}(T-t)}|\omega\rangle\langle\omega|\psi_t\rangle dt = \\
&= |\varphi_T\rangle - iE\int_0^T e^{-i\mathbf{H}(T-t)}|\omega\rangle\langle\omega|e^{-i\mathbf{H}'t}|\psi_0\rangle dt.
\end{aligned}
$$

Вооруженные этим выражением, мы можем вывести границу для расстояния между $|\psi_T^\omega\rangle$ и $|\varphi_T\rangle$. (Аналог «зловещего» уравнения за номером (6.66) в лекциях[1].)

$$
\begin{aligned}
\| |\psi_T^\omega\rangle - |\varphi_T\rangle \| &= \left\| -iE\int_0^T e^{-i\mathbf{H}(T-t)}|\omega\rangle\langle\omega|\psi_t\rangle dt \right\| \leqslant \\
&\leqslant E\int_0^T \left\| e^{-i\mathbf{H}(T-t)}|\omega\rangle\langle\omega|\psi_t\rangle \right\| dt = \\
&= E\int_0^T \| |\omega\rangle\langle\omega|\psi_t\rangle \| dt = \\
&= E\int_0^T |\langle\omega|\psi_t\rangle| dt.
\end{aligned}
$$

[1] Подходящее для задания на пятницу 13-го, не правда ли?

Возводя это соотношение в квадрат, находим

$$\| |\psi_T^\omega\rangle - |\varphi_T\rangle\|^2 \leqslant E^2 \left(\int_0^T |\langle\omega|\psi_t\rangle| dt \right)^2 \leqslant$$

$$\leqslant E^2 T \int_0^T |\langle\omega|\psi_t\rangle|^2 dt.$$

Усредняя этот результат по всем возможным оракулам, мы находим, что, в подтверждение наших ранних подозрений, среднеквадратичное расстояние между конечными состояниями алгоритмов $\mathcal{A}$ и $\mathcal{D}$ ограничено сверху:

$$\left\langle (d(\mathcal{A}, \mathcal{D}))^2 \right\rangle = \frac{1}{2^n} \sum_\omega \| |\psi_T^\omega\rangle - |\varphi_T\rangle\|^2 \leqslant$$

$$\leqslant \frac{1}{2^n} E^2 T \sum_\omega \int_0^T \langle\psi_t|\omega\rangle\langle\omega|\psi_t\rangle dt =$$

$$= \frac{1}{2^n} E^2 T^2.$$

К счастью для нас, это среднеквадратичное расстояние также ограничено и снизу. Так как состояния $\{|\psi_T^\omega\rangle\}$ образуют ортогональный базис, они не могут все сколь угодно близко сконцентрироваться вокруг некоторого определенного фиксированного состояния. В частности, они не могут все сконцентрироваться вокруг $|\varphi_T\rangle$ и, следовательно, среднеквадратичное расстояние между конечными состояниями алгоритмов $\mathcal{A}$ и $\mathcal{D}$ ограничено снизу уравнением (6.159) из лекций:

$$\left\langle (d(\mathcal{A}, \mathcal{D}))^2 \right\rangle \geqslant \frac{1}{2^n} \left(2 \cdot 2^n - 2\sqrt{2^n}\right).$$

Сравнивая эти верхнюю и нижнюю границы, мы, как и было обещано, получаем квадратичную по времени границу гроверовского типа:

$$E^2 T^2 \geqslant 2 \cdot 2^n - 2\sqrt{2^n},$$

$$T \geqslant \frac{\sqrt{2}}{E} 2^{n/2} \sqrt{1 - 2^{-n/2}},$$

$$T \geqslant \frac{\sqrt{2}}{E} 2^{n/2}.$$

Поскольку $\sqrt{2} \approx 1{,}41$, а $\frac{\pi}{2} \approx 1{,}57$, эта общая граница сильнее текущего времени явного алгоритма из части **(а)** на

$$\frac{\pi/2 - \sqrt{2}}{\sqrt{2}} \approx 11\%.$$

Это то же различие, что и найденное нами в исходном (то есть дискретном) алгоритме Гровера. Поскольку в дискретном случае более тонкие границы демонстрировали насыщение алгоритма, у нас есть все основания ожидать, что и для непрерывного алгоритма подобное улучшение границы также будет демонстрировать оптимальность.

Прескилл Джон

Квантовая информация и квантовые вычисления

Том 1

Дизайнер М. Баженова
Технический редактор А. В. Широбоков
Компьютерная верстка Д. П. Вакуленко, А. В. Моторин
Корректор Г. Г. Тетерина

Подписано в печать 21.02.2008. Формат 60 × 84 1/16.
Печать офсетная. Усл. неч. л. 26,97. Уч. изд. л. 25,21.
Гарнитура Таймс. Бумага офсетная №1. Заказ №10.

Научно-издательский центр «Регулярная и хаотическая динамика»
426034, г. Ижевск, ул. Университетская, 1.
http://shop.rcd.ru E-mail: mail@rcd.ru Тел./факс: (+73412) 500–295

Переплет выполнен в ГУП УР «Ижевский полиграфический комбинат»
426039, г. Ижевск, Воткинское шоссе, 180.

Уважаемые читатели!

Интересующие Вас книги нашего издательства можно заказать через наш Интернет-магазин **http://shop.rcd.ru** или по электронной почте **subscribe@rcd.ru**

Книги можно приобрести в наших представительствах:

МОСКВА
Институт машиноведения им. А. А. Благонравова РАН
ул. Бардина, д. 4, корп. 3, к. 414, тел.: 135–54–37

ИЖЕВСК
Удмуртский государственный университет
ул. Университетская, д. 1, корп. 4, 2 эт., к. 211, тел./факс: (3412) 500 295

Также книги можно приобрести:

МОСКВА
Московский государственный университет им. М.В. Ломоносова
ГЗ (1 эт.), Физический ф-т (1 эт.), Гуманитарный ф-т (0 и 1 эт.),
Биологический ф-т (1 эт.).

Российский государственный университет нефти и газа им. И. М. Губкина
ГЗ (3–4 эт.), книжные киоски фирмы «Аргумент».

Магазины:

МОСКВА:	«Дом научно-технической книги» Ленинский пр., 40. тел.: 137-06-33
	«Московский дом книги» ул. Новый Арбат, 8. тел.: 290-45-07
	«Библиоглобус» м. «Лубянка», ул. Мясницкая, 6. тел.: 928–87 -44
ДОЛГОПРУДНЫЙ:	Книжный магазин «Физматкнига» новый корп. МФТИ, 1 эт. тел.: 409-93-28
САНКТ-ПЕТЕРБУРГ:	«Санкт-Петербургский дом книги» Невский проспект, 28
	Издательство СПбГУ, Магазин №1 Университетская набережная, 7/9

Дж. Прескилл — известный физик-теоретик, профессор теоретической физики Отделения Физики, Математики и Астрономии Калифорнийского Технологического Института (КАЛТЕХ). Область научных интересов — физика элементарных частиц и космология, топологические дефекты, непертурбативные методы квантовой теории поля, квантовые аспекты ранней Вселенной и черных дыр. В середине 90-х годов увлекся теорией квантовой информации, квантовых вычислений и кодирования. В настоящее время — один из ведущих специалистов в этой области. Руководитель Института Квантовых Вычислений, а также Центра Физики Информации при КАЛТЕХе.

ISBN 978-5-93972-651-1

Дж. Прескилл

Квантовая информация и квантовые вычисления

Том 2

R&C
Dynamics


Институт
Компьютерных
Исследований


R&C
Dynamics

Lecture Notes for Physics 229: Quantum Information and Computation

John Preskill
California Institute of Technology

September, 1998

Дж. Прескилл

Квантовая информация и квантовые вычисления

Том 2

Перевод с английского
Т. С. Нечаевой

Под научной редакцией
С. Г. Новокшонова

Москва ♦ Ижевск

2011

УДК 22.314.1
ББК 517.958:530.145.6
П 73

Прескилл Дж.
Квантовая информация и квантовые вычисления. Том 2. — М.–Ижевск: НИЦ «Регулярная и хаотическая динамика», Ижевский институт компьютерных исследований, 2011. — 312 с.

Книга Дж. Прескилла представляет собой наиболее полное современное введение в новую, бурно развивающуюся область науки — теорию квантовой информации и квантовых вычислений.

Вопросы, рассматриваемые во втором томе, объединяет общая тема: защита квантовой информации от ошибок, возникающих как во время ее хранения и передачи, так и при оперировании с ней. В первой из двух основных глав излагаются принципы детектирования, диагностики и коррекции квантовых ошибок; основные типы и принципы организации и работы квантовых кодов коррекции ошибок. Кроме этого в Приложении помещены две обзорные статьи Дж. Прескилла, в которых обсуждается проблема реализации отказоустойчивых квантовых вычислений на основе схем, использующих «шумящие» вентили.

В отдельной большой главе впервые в русскоязычной литературе рассматривается принципиально новый *физический* подход к проблеме защиты квантовой информации от ошибок, в основе которого лежит топологическая устойчивость некоторых квантовых состояний, реализующихся в низкоразмерных многочастичных сильнокоррелированных системах. Несмотря на сложность обсуждаемых в этой главе физических и математических идей, ее содержание дает «ясное представление о предмете без доходящих до абсурда упрощений» и вполне доступно читателю, знакомому с нерелятивистской квантовой механикой, основами теории представлений групп и самыми элементарными сведениями из топологии.

ISBN 978-5-4344-0030-5 **ББК 517.958:530.145.6**

http://shop.rcd.ru
http://ics.org.ru

Оглавление

Глава 7

Коррекция квантовых ошибок

7.1. Квантовые коды коррекции ошибок

Изучая квантовые алгоритмы, мы нашли убедительные свидетельства того, что квантовый компьютер может обладать исключительными способностями. Но будет ли он действительно работать? Сможем ли мы когда-нибудь создать квантовый компьютер и управлять им?

Чтобы добиться этого, необходимо решить проблему защиты квантовой информации от ошибок. Как уже отмечалось в первой главе, у этой проблемы есть несколько аспектов. Между квантовым компьютером и его окружением неизбежно взаимодействие, приводящее к декогерентизации и, следовательно, к разрушению хранящейся в нем квантовой информации. Пока мы не сможем успешно противостоять декогерентизации, наш квантовый компьютер безусловно обречен. Даже если бы нам удалось предотвратить декогерентизацию, полностью изолировав компьютер от окружающей среды, ошибки по-прежнему представляли бы серьезные трудности. Квантовые вентили (в отличие от классических) представляют собой унитарные преобразования, множество которых образует континуум. Следовательно, идеально точное выполнение квантовых операций невозможно. Малые погрешности вентилей будут аккумулироваться, приводя в конце концов к серьезному сбою в вычислении. Любая эффективная стратегия борьбы с ошибками в квантовом компьютере должна обеспечивать защиту как от декогерентизации, так и от малых унитарных ошибок в квантовых схемах.

В этой и следующей главах мы увидим, как искусное кодирование квантовой информации может (в принципе) защитить ее от ошибок. В этой главе будет представлена теория квантовых кодов коррекции ошибок. Мы узнаем, что соответствующим образом закодированная квантовая информация может быть помещена в квантовое запоминающее устройство (квантовую память), подвергаемое разрушительному воздействию шума окружающей среды, и извлечена оттуда неповрежденной (если шум не слишком

силен). Затем в восьмой главе мы распространим эту теорию в двух важных направлениях. Мы увидим, что процедура восстановления информации, может эффективно работать, даже если в ходе ее время от времени случаются ошибки. Мы узнаем, как следует обращаться с закодированной информацией, чтобы квантовые вычисления могли успешно выполняться, несмотря на разрушительное действие декогерентизации и несовершенство квантовых логических вентилей.[1]

Квантовый код коррекции ошибок (КККО) можно рассматривать как отображение k кубитов (гильбертово пространство размерности 2^k) на n кубитов (гильбертово пространство размерности 2^n), где $n > k$. Эти k кубитов представляют собой «логические кубиты» или «закодированные кубиты», которые мы хотим защитить от ошибок. Дополнительные $n-k$ кубитов позволяют хранить k логических кубитов в избыточном виде, так чтобы закодированную информацию было нелегко разрушить.

Для того чтобы лучше понять идею КККО, вернемся к рассмотренному в первой главе примеру кода Шора с $n = 9$ и $k = 1$. Его можно описать, определив два базисных состояния подпространства кода; будем обозначать эти базисные состояния как $|\bar{0}\rangle$ — «логический нуль» и $|\bar{1}\rangle$ — «логическая единица». Они имеют вид

$$\begin{aligned} |\bar{0}\rangle &= \left[\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)\right]^{\otimes 3}, \\ |\bar{1}\rangle &= \left[\frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)\right]^{\otimes 3}; \end{aligned} \tag{7.1}$$

каждое базисное состояние представляет собой троекратно повторенное трехкубитовое «кот-состояние». Как вы помните из обсуждения «кот-состояния» в четвертой главе, два базисных состояния можно различить с помощью трехкубитовой наблюдаемой $\boldsymbol{\sigma}_x^1 \otimes \boldsymbol{\sigma}_x^2 \otimes \boldsymbol{\sigma}_x^3$ (где $\boldsymbol{\sigma}_x^i$ обозначает матрицу Паули $\boldsymbol{\sigma}_x$, действующую на i-й кубит); мы будем использовать обозначение $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$ для этого оператора. (Здесь подразумевается, что на остальные кубиты действует скрытый в этом обозначении оператор $1 \otimes 1 \otimes \ldots \otimes 1$.) Состояния $|\bar{0}\rangle$ и $|\bar{1}\rangle$ являются собственными векторами оператора $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$ с собственными значениями $+1$ и -1 соответственно. Однако невозможно

[1] Материал главы, посвященной отказоустойчивым квантовым вычислениям, до настоящего времени не опубликован. Чтобы как-то компенсировать этот пробел и познакомить читателей с основными принципами реализации отказоустойчивых квантовых вычислений, с разрешения автора в приложении предлагаются переводы двух его обзорных статей, посвященных этой теме. — *Прим. ред.*

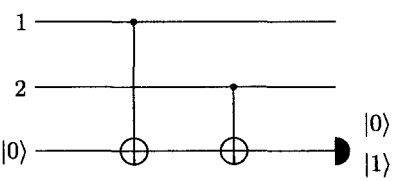
отличить $|\bar{0}\rangle$ от $|\bar{1}\rangle$ (извлечь любую информацию о значении логического кубита), наблюдая любые один или два кубита из имеющихся девяти. В этом смысле, логический кубит закодирован *нелокальным образом*; фактически, он записан в запутанности между кубитами блока. Это свойство нелокальности закодированной информации обеспечивает защиту от шума, если, конечно, он является локальным (то есть действует независимо, или почти независимо, на различные кубиты в блоке).

Предположим, что приготовлено неизвестное квантовое состояние и закодировано как $a|\bar{0}\rangle + b|\bar{1}\rangle$. Пусть теперь возникла ошибка; мы должны выявить ее и уничтожить. Как нам поступить? Допустим для начала, что происходит однократное инвертирование, действующее на один из трех первых кубитов. Тогда, как обсуждалось в первой главе, положение инвертированного кубита можно определить путем измерения двухкубитовых операторов

$$\mathbf{Z}_1\mathbf{Z}_2, \qquad \mathbf{Z}_2\mathbf{Z}_3. \tag{7.2}$$

Базисные логические состояния $|\bar{0}\rangle$ и $|\bar{1}\rangle$ являются собственными векторами этих операторов с собственным значением $+1$. Но инвертирование любого из трех кубитов меняет эти собственные значения. Например, если $\mathbf{Z}_1\mathbf{Z}_2 = -1$, а $\mathbf{Z}_2\mathbf{Z}_3 = 1$,[1] то мы делаем вывод, что инвертирован первый кубит относительно двух других. Мы можем исправить ошибку, еще раз инвертируя этот кубит.

Важно, что наше измерение, диагностирующее инвертированный кубит, является коллективным измерением двух кубитов одновременно — мы узнаем значение $\mathbf{Z}_1\mathbf{Z}_2$, но не должны определять индивидуальные значения $\mathbf{Z}_1$ и $\mathbf{Z}_2$, поскольку это может повредить закодированное состояние. Как выполнить такое коллективное измерение? Фактически, его можно осуществить, располагая квантовым компьютером, способным выполнять операции CNOT (контролируемое НЕ). Сначала мы вводим приготовленный в состоянии $|0\rangle$ дополнительный «служебный» кубит, затем выполняем квантовую схему



[1] Вновь (см. подстрочное примечание на с. 201 первого тома) подобные равенства следует понимать как символические. В данном случае, например, первое из них означает, что состояние с одним инвертированным кубитом является собственным состоянием оператора $\mathbf{Z}_1\mathbf{Z}_2$ с собственным значением -1. — *Прим. ред.*

и, наконец, измеряем служебный кубит. Если кубиты 1 и 2 находятся в состоянии с $\mathbf{Z}_1\mathbf{Z}_2 = -1$ ($|0\rangle_1|1\rangle_2$ или $|1\rangle_1|0\rangle_2$), то служебный кубит однократно инвертируется и результатом его измерения будет $|1\rangle$. Но если кубиты 1 и 2 находятся в состоянии с $\mathbf{Z}_1\mathbf{Z}_2 = +1$ ($|0\rangle_1|0\rangle_2$ или $|1\rangle_1|1\rangle_2$), то служебный кубит останется неизменным или инвертируется дважды, а результатом измерения будет $|0\rangle$. Аналогично, можно выполнить измерение и операторов других двухкубитовых наблюдаемых

$$\begin{array}{ll} \mathbf{Z}_4\mathbf{Z}_5, & \mathbf{Z}_5\mathbf{Z}_6, \\ \mathbf{Z}_7\mathbf{Z}_8, & \mathbf{Z}_8\mathbf{Z}_9, \end{array} \tag{7.3}$$

чтобы диагностировать ошибки инвертирования кубита в двух других кластерах из трех кубитов.

Трехкубитового кода достаточно для защиты от однократного инвертирования бита. Для защиты от фазовых ошибок требуется троекратное повторение трехкубитовых кластеров. Предположим, что возникает фазовая ошибка

$$|\psi\rangle \rightarrow \mathbf{Z}|\psi\rangle, \tag{7.4}$$

действующая на один из девяти кубитов. Мы можем определить, в каком из кластеров она возникла, измерив две шестикубитовые наблюдаемые

$$\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4\mathbf{X}_5\mathbf{X}_6, \qquad \mathbf{X}_4\mathbf{X}_5\mathbf{X}_6\mathbf{X}_7\mathbf{X}_8\mathbf{X}_9. \tag{7.5}$$

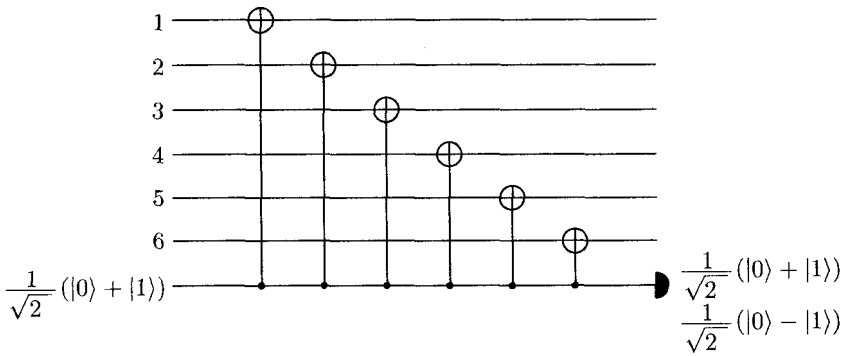
Оба базисных логических состояния $|\bar{0}\rangle$ и $|\bar{1}\rangle$ являются собственными векторами этих операторов с собственным значением одной из этих наблюдаемых (± 1). Фазовая ошибка, действующая на любой один из кубитов в некотором кластере, изменит в нем значение $\mathbf{XXX}$ относительно двух других кластеров; положение этого изменения можно определить путем измерения наблюдаемых (7.5). Как только поврежденный кластер определен, ошибку можно исправить, применяя $\mathbf{Z}$ к одному из кубитов этого кластера.

Как измерить шестикубитовую наблюдаемую $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4\mathbf{X}_5\mathbf{X}_6$? Заметим, что если начальным состоянием управляющего кубита является $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, а управляемым кубитом является собственное состояние $\mathbf{X}$ (то есть NOT), то вентиль CNOT действует в соответствии с правилом

$$\mathrm{CNOT} : \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |x\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + (-1)^x|1\rangle) \otimes |x\rangle; \tag{7.6}$$

он действует тривиально, если управляемое состояние соответствует собственному значению $\mathbf{X} = +1$ ($x = 0$), и обращает фазу управляющего

кубита, если управляемому соответствует собственное значение $\mathbf{X} = -1$ ($x = 1$).[1] Чтобы измерить произведение $\mathbf{X}$-операторов, мы выполняем схему



а затем измеряем служебный кубит в базисе $\frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$.

Таким образом, действующая на любой один из девяти кубитов в блоке ошибка не вызывает непоправимого повреждения. Но если в одном кластере из трех кубитов инвертируются два, то закодированная информация *будет* разрушена. Например, если в кластере одновременно инвертируются два первых кубита, то мы неверно определим ошибку и, пытаясь исправить ее, инвертируем третий кубит. Эти ошибки совместно с неправильной попыткой исправления действуют на кодовый блок оператором $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$. Поскольку $|\bar{0}\rangle$ и $|\bar{1}\rangle$ являются собственными состояниями оператора $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$ с различными собственными значениями, то в результате инвертирования двух битов в одном кластере в закодированном кубите возникает *фазовая ошибка*

$$\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3 \,:\, a|\bar{0}\rangle + b|\bar{1}\rangle \;\rightarrow\; a|\bar{0}\rangle - b|\bar{1}\rangle. \tag{7.7}$$

Закодированная информация также будет повреждена, если фазовые ошибки возникнут в двух разных кластерах. Тогда в результате неверной попытки исправления мы внесем фазовую ошибку в третий кластер, так что в итоге будет применен оператор $\mathbf{Z}_1\mathbf{Z}_4\mathbf{Z}_7$, который инвертирует закодированный кубит

$$\mathbf{Z}_1\mathbf{Z}_4\mathbf{Z}_7 \,:\, a|\bar{0}\rangle + b|\bar{1}\rangle \;\rightarrow\; a|\bar{1}\rangle + b|\bar{0}\rangle. \tag{7.8}$$

[1] Другими словами, в базисе собственных состояний оператора $\mathbf{X}$ управляющий (контролирующий, или источник) и управляемый (контролируемый или целевой) кубиты вентиля CNOT меняются ролями. В том, что это действительно так, нетрудно убедиться, используя его определение (4.11) в базисе собственных состояний оператора Z. Напомним, что в последнем случае под действием CNOT, в зависимости от состояния управляющего кубита изменяется (инвертируется) или остается неизменным управляемый кубит. – *Прим. ред.*

Если вероятность ошибки достаточно мала и если ошибки, действующие на разные кубиты, не сильно скоррелированы, то использование девятикубитового кода позволяет сохранить неизвестный кубит надежнее, чем в том случае, когда мы вообще не беспокоимся о его кодировании. Предположим, например, что на каждый из девяти кубитов окружающая среда действует как описанный в третьей главе деполяризующий канал с вероятностью ошибки p. Тогда вероятность инвертирования бита равна $\frac{2}{3}p$, а вероятность обращения фазы — $\frac{2}{3}p$. (Вероятность одновременного появления обеих ошибок равна $\frac{1}{3}p$.) Нетрудно видеть, что вероятность фазовой ошибки, непоправимо искажающей логический кубит, ограничена сверху величиной $4p^2$, а вероятность подобной ошибки инвертирования бита — величиной $12p^2$. Полная вероятность ошибки не превышает $16p^2$; то есть улучшение по сравнению с вероятностью ошибки p незащищенного кубита имеет место при условии $p < 1/16$.

Конечно, в приведенном выше анализе по умолчанию предполагалось, что кодирование, декодирование, измерение синдрома ошибки и ее исправление выполняются идеально точно. Более реалистический случай, когда ошибки возникают и во время этих операций, обсуждается в приложении.

7.2. Критерии исправления квантовых ошибок

При обсуждении исправления ошибок с помощью девятикубитового кода предполагалось, что каждый кубит подвержен либо ошибке инвертирования бита, либо ошибке обращения фазы (или им обеим). Это нереалистическая модель ошибок, и нам следует понять, как осуществлять коррекцию квантовых ошибок в более общих условиях.

Рассмотрим сначала один кубит, первоначально находящийся в чистом состоянии и произвольным образом взаимодействующий со своим окружением. Из третьей главы мы знаем, что без потери общности (мы все еще можем представлять, что на наш кубит действует самый общий супероператор) можно предполагать, что начальным состоянием окружения является чистое состояние, которое мы обозначим как $|0\rangle_E$. Тогда эволюция кубита и его окружения может быть описана унитарным преобразованием

$$\mathbf{U}: \begin{array}{l} |0\rangle \otimes |0\rangle_E \rightarrow |0\rangle \otimes |e_{00}\rangle_E + |1\rangle \otimes |e_{01}\rangle_E, \\ |1\rangle \otimes |0\rangle_E \rightarrow |0\rangle \otimes |e_{10}\rangle_E + |1\rangle \otimes |e_{11}\rangle_E; \end{array} \tag{7.9}$$

здесь $|e_{ij}\rangle_E$ — четыре состояния окружения, которым не обязательно быть нормированными или взаимно ортогональными (хотя они удовлетворяют

некоторым ограничениям, вытекающим из унитарности $\mathbf{U}$). Под действием $\mathbf{U}$ произвольное состояние кубита $|\psi\rangle = a|0\rangle + b|1\rangle$ эволюционирует как

$$\begin{aligned}\mathbf{U}\colon (a|0\rangle + b|1\rangle) \otimes |0\rangle_E \to\ & a(|0\rangle \otimes |e_{00}\rangle_E + |1\rangle \otimes |e_{01}\rangle_E) + \\ & + b(|0\rangle \otimes |e_{10}\rangle_E + |1\rangle \otimes |e_{11}\rangle_E) = \\ & = (a|0\rangle + b|1\rangle) \otimes \frac{1}{2}(|e_{00}\rangle_E + |e_{11}\rangle_E) + \\ & + (a|1\rangle + b|0\rangle) \otimes \frac{1}{2}(|e_{01}\rangle_E + |e_{10}\rangle_E) + \\ & + (a|1\rangle - b|0\rangle) \otimes \frac{1}{2}(|e_{01}\rangle_E - |e_{10}\rangle_E) + \\ & + (a|0\rangle - b|1\rangle) \otimes \frac{1}{2}(|e_{00}\rangle_E - |e_{11}\rangle_E) \equiv \\ & \equiv \mathbf{1}|\psi\rangle \otimes |e_I\rangle_E + \mathbf{X}|\psi\rangle \otimes |e_X\rangle_E + \\ & + \mathbf{Y}|\psi\rangle \otimes |e_Y\rangle_E + \mathbf{Z}|\psi\rangle \otimes |e_Z\rangle_E. \end{aligned} \tag{7.10}$$

Действие $\mathbf{U}$ может быть разложено по (унитарным) операторам Паули $\{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$, просто потому, что они образуют базис в векторном пространстве 2×2-матриц. Эвристически мы можем интерпретировать это разложение, говоря, что с кубитом происходит одно из четырех возможных событий: ничего ($\mathbf{1}$), инвертирование бита ($\mathbf{X}$), обращение фазы ($\mathbf{Z}$) или обе ошибки ($\mathbf{Y} = i\mathbf{XZ}$). Однако не следует понимать эту классификацию буквально, поскольку, пока состояния окружения $\{|e_I\rangle_E, |e_X\rangle_E, |e_Y\rangle_E, |e_Z\rangle_E\}$ не являются взаимно ортогональными, не существует мыслимого измерения, которое могло бы идеально различить эти четыре альтернативы.

Аналогично, действующая в n-кубитовом гильбертовом пространстве произвольная $2^n \times 2^n$-матрица может быть разложена по 2^{2n} операторам

$$\{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}^{\otimes n}; \tag{7.11}$$

то есть каждый такой оператор может быть представлен как тензорное произведение однокубитовых операторов, каждый из которых выбирается из единичного $\mathbf{1}$ и трех матриц Паули $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$. Таким образом, действие произвольного унитарного оператора на n кубитов и их окружение можно представить в виде разложения

$$|\psi\rangle \otimes |0\rangle_E \to \sum_a \mathbf{E}_a|\psi\rangle \otimes |e_a\rangle_E; \tag{7.12}$$

здесь индекс a пробегает 2^{2n} значений. $\{\mathbf{E}_a\}$ — множество всех линейно независимых операторов Паули, действующих на n кубитов, а $|e_a\rangle_E$ — со-

ответствующие состояния окружения (которые *не предполагаются* нормированными или взаимно ортогональными). Важной для дальнейшего особенностью этого разложения является то, что каждый оператор $\mathbf{E}_a$ является унитарным.

Уравнение (7.12) обеспечивает идейную основу коррекции квантовых ошибок. При разработке ККК0 мы определяем подмножество $\mathcal{E}$ всех операторов Паули

$$\mathcal{E} \subseteq \{\mathbf{E}_a\} \equiv \{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}^{\otimes n}; \tag{7.13}$$

это множество тех ошибок, которые мы хотим уметь исправлять. Нашей целью является выполнение коллективного измерения n кубитов в кодовом блоке, которое позволяет определить, какая из ошибок $\mathbf{E}_a \in \mathcal{E}$ возникла. Если $|\psi\rangle$ — состояние, принадлежащее кодовому подпространству, то для некоторых (но не для всех) кодов это измерение приготовит состояние $\mathbf{E}_a|\psi\rangle \otimes |e_a\rangle_E$, где значение a известно из результата измерения. Так как оператор $\mathbf{E}_a$ является унитарным (и одновременно самосопряженным), мы можем применить к кодовому блоку оператор $\mathbf{E}_a^\dagger$ $(= \mathbf{E}_a)$, восстанавливая неповрежденное состояние $|\psi\rangle$.

Каждому оператору Паули можно сопоставить *вес*, целое число t, удовлетворяющее неравенству $0 \leqslant t \leqslant n$; вес представляет собой количество кубитов, на которые действует нетривиальная матрица Паули ($\mathbf{X}$, $\mathbf{Y}$ или $\mathbf{Z}$). Тогда эвристически слагаемое разложения (7.12), в котором оператор $\mathbf{E}_a$ имеет вес t, можно интерпретировать как событие, состоящее в появлении ошибок в t кубитах (и вновь не следует принимать эту интерпретацию буквально, если состояния $\{|e_a\rangle_E\}$ не являются взаимно ортогональными). Как правило, в качестве $\mathcal{E}$ выбирается совокупность всех операторов Паули с весами вплоть до некоторого t включительно; тогда если удается восстановить исходное состояние после действия на него любого супероператора ошибки с носителем из множества $\mathcal{E}$, то мы говорим, что код может исправить t ошибок. Такой выбор множества $\mathcal{E}$ неявно предполагает, что ошибки, возмущающие разные кубиты, слабо коррелируют между собой, поэтому вероятность возникновения большего, чем t, количества ошибок на n кубитах относительно мала.

Каким необходимым и достаточным условиям должно удовлетворять кодовое подпространство, для того чтобы было возможно исправление заданного множества ошибок $\mathcal{E}$? Обозначим как $\{|\bar{i}\rangle\}$ ортонормированный базис в кодовом подпространстве. (Будем говорить об этих базисных элементах как о «кодовых словах».) Очевидно, *необходимо*, чтобы

$$\langle \bar{j}|\mathbf{E}_b^\dagger \mathbf{E}_a|\bar{i}\rangle = 0, \qquad i \neq j, \tag{7.14}$$

где $\mathbf{E}_{a,b} \in \mathcal{E}$. Если бы это условие не выполнялось для некоторых $i \neq j$, то ошибки могли бы разрушить идеальную различимость ортогональных кодовых слов и закодированной квантовой информации наверняка был бы нанесен ущерб. (Более подробный вывод этого необходимого условия будет представлен ниже.) Также нетрудно видеть, что *достаточным* условием является

$$\langle \bar{j}|\mathbf{E}_b^\dagger \mathbf{E}_a|\bar{i}\rangle = \delta_{ab}\delta_{ij}. \tag{7.15}$$

В этом случае операторы $\mathbf{E}_a$ разбивают кодовое подпространство на совокупность взаимно ортогональных «подпространств ошибок»

$$\mathcal{H}_a = \mathbf{E}_a \mathcal{H}_{code}. \tag{7.16}$$

Предположим, что в произвольное состояние $|\psi\rangle$, приготовленное в кодовом пространстве, вкралась ошибка. Тогда итоговым состоянием кодового блока и окружения является

$$\sum_{\mathbf{E}_a \in \mathcal{E}} \mathbf{E}_a|\psi\rangle \otimes |e_a\rangle_E, \tag{7.17}$$

где суммирование ведется по ошибкам, принадлежащим множеству $\mathcal{E}$. В этом случае можно выполнить ортогональное измерение, проецирующее кодовый блок на одно из пространств $\mathcal{H}_a$, так что состояние приобретает вид

$$\mathbf{E}_a|\psi\rangle \otimes |e_a\rangle_E. \tag{7.18}$$

Наконец, для завершения процедуры восстановления к кодовому блоку применяется унитарный оператор $\mathbf{E}_a^\dagger$.

Код, удовлетворяющий условию (7.15), называется *невырожденным*. Этот термин означает, что существует измерение, которое может однозначно выявить возникшую ошибку $\mathbf{E}_a \in \mathcal{E}$. Но пример девятикубитового кода только что показал, что возможны более общие коды. Девятикубитовый код *вырожден*, поскольку фазовые ошибки, действующие на разные кубиты одного и того же кластера из трех кубитов, одинаковым образом влияют на кодовое подпространство (например, $\mathbf{Z}_1|\psi\rangle = \mathbf{Z}_2|\psi\rangle$). Хотя ни одно измерение не может определить, в каком из кубитов возникла ошибка, это не является помехой для ее успешного исправления.

Нетрудно установить необходимое и достаточное условие возможности восстановления

$$\langle \bar{j}|\mathbf{E}_b^\dagger \mathbf{E}_a|\bar{i}\rangle = C_{ab}\delta_{ij}, \tag{7.19}$$

где $\mathbf{E}_{a,b} \in \mathcal{E}$, а $C_{ab} = \langle \bar{i}|\mathbf{E}_b^\dagger \mathbf{E}_a|\bar{i}\rangle$ — произвольная эрмитова матрица. Нетривиальное содержание этого условия, которое существенно сильнее более

слабого необходимого условия (7.14), состоит в том, что $\langle\bar{i}|\mathbf{E}_b^\dagger\mathbf{E}_a|\bar{i}\rangle$ не зависит от i. Природа этого условия очевидна — будь это иначе, при определении подпространства ошибки $\mathcal{H}_a$ мы получали бы некоторую информацию о закодированном состоянии, что неизбежно приводило бы к его возмущению.

Чтобы доказать необходимость и достаточность условия (7.19), обратимся к развитой в третьей главе теории супероператоров. Действующая на кодовый блок ошибка описывается супероператором, и проблема состоит в том, можно ли построить другой супероператор (процедура восстановления), аннулирующий ее действие. В третьей главе мы узнали, что обратить можно только те супероператоры, которые являются унитарными операторами. Однако от нас не требуется умение аннулировать действие супероператора ошибки на любое состояние в n-кубитовом кодовом блоке; вполне достаточно уметь исправлять ошибки в состояниях, первоначально принадлежавших k-кубитовому закодированному подпространству.

Альтернативным выражением действия ошибки на одно из кодовых базисных состояний $|\bar{i}\rangle$ (и на окружение) является

$$|\bar{i}\rangle\otimes|0\rangle_E \rightarrow \sum_\mu \mathbf{M}_\mu|\bar{i}\rangle\otimes|\mu\rangle_E, \tag{7.20}$$

где теперь состояния $|\mu\rangle_E$ представляют собой элементы *ортонормированного базиса* окружения, а матрицы $\mathbf{M}_\mu$ являются линейными комбинациями содержащихся в $\mathcal{E}$ операторов Паули $\mathbf{E}_a$ и удовлетворяют условию нормировки операторной суммы

$$\sum_\mu \mathbf{M}_\mu^\dagger\mathbf{M}_\mu = \mathbf{1}. \tag{7.21}$$

Ошибка может быть исправлена оператором восстановления, если существуют такие операторы $\mathbf{R}_\nu$, что

$$\sum_\nu \mathbf{R}_\nu^\dagger\mathbf{R}_\nu = \mathbf{1} \tag{7.22}$$

и

$$\sum_{\mu,\nu} \mathbf{R}_\nu\mathbf{M}_\mu|\bar{i}\rangle\otimes|\mu\rangle_E\otimes|\nu\rangle_A = |\bar{i}\rangle\otimes|\mathrm{stuff}\rangle_{E,A}; \tag{7.23}$$

здесь векторы $|\nu\rangle_A$ являются элементами ортонормированного базиса гильбертова пространства служебного кубита, привлекаемого для осуществления операции восстановления, а состояние окружающей среды и служебного кубита $|\mathrm{stuff}\rangle_{E,A}$ не должно зависеть от i. Отсюда следует, что для каждого μ и ν

$$\mathbf{R}_\nu\mathbf{M}_\mu|\bar{i}\rangle = \lambda_{\nu\mu}|\bar{i}\rangle; \tag{7.24}$$

в кодовом подпространстве действие произведения $\mathbf{R}_\nu \mathbf{M}_\mu$ эквивалентно умножению на число. Используя условие нормировки, которому удовлетворяют операторы $\mathbf{R}_\nu$, мы находим, что

$$\mathbf{M}_\delta^\dagger \mathbf{M}_\mu |\bar{i}\rangle = \mathbf{M}_\delta^\dagger \left(\sum_\nu \mathbf{R}_\nu^\dagger \mathbf{R}_\nu \right) \mathbf{M}_\mu |\bar{i}\rangle = \sum_\nu \lambda_{\nu\delta}^* \lambda_{\nu\mu} |\bar{i}\rangle, \tag{7.25}$$

так что действие $\mathbf{M}_\delta^\dagger \mathbf{M}_\mu$ в кодовом подпространстве также эквивалентно умножению на число. Другими словами,

$$\langle \bar{j} | \mathbf{M}_\delta^\dagger \mathbf{M}_\mu | \bar{i} \rangle = C_{\delta\mu} \delta_{ij}; \tag{7.26}$$

отсюда следует (7.19), поскольку каждый оператор $\mathbf{E}_a$ из $\mathcal{E}$ является линейной комбинацией операторов $\mathbf{M}_\mu$.

Другой поучительный способ понять, почему (7.26) является необходимым условием возможности исправления ошибки, — это обратить внимание на то, что если кодовый блок приготовлен в состоянии $|\psi\rangle$, а ошибка действует в соответствии (7.20), то получаемая путем вычисления следа по кодовому блоку матрица плотности окружения имеет вид

$$\boldsymbol{\rho}_E = \sum_{\mu\nu} |\mu\rangle_E \langle\psi| \mathbf{M}_\nu^\dagger \mathbf{M}_\mu |\psi\rangle \, {}_E\langle\nu|. \tag{7.27}$$

Ошибка может быть успешно исправлена только в том случае, если в процессе измерения окружения невозможно получить какую-либо информацию о состоянии $|\psi\rangle$. Следовательно, мы требуем, чтобы $\boldsymbol{\rho}_E$ не зависела от $|\psi\rangle$, если $|\psi\rangle$ — произвольное состояние из кодового подпространства; тогда отсюда следует уравнение (7.26).

Чтобы увидеть, что уравнение (7.26) как необходимо, так и достаточно, можно явно построить исправляющий ошибки супероператор. С этой целью достаточно выбрать базис окружения $\{|\mu\rangle_E\}$ таким образом, чтобы матрица $C_{\delta\mu}$ в уравнении (7.26) была диагональна

$$\langle \bar{j} | \mathbf{M}_\delta^\dagger \mathbf{M}_\mu | \bar{i} \rangle = C_\mu \delta_{\delta\mu} \delta_{ij}, \tag{7.28}$$

где $\sum_\mu C_\mu = 1$ вытекает из условия нормировки операторной суммы. Пусть для каждого ν с $C_\nu \neq 0$

$$\mathbf{R}_\nu = \frac{1}{\sqrt{C_\nu}} \sum_i |\bar{i}\rangle \langle \bar{i} | \mathbf{M}_\nu^\dagger, \tag{7.29}$$

так что $\mathbf{R}_\nu$ действует в соответствии с

$$\mathbf{R}_\nu\colon \mathbf{M}_\mu|\bar{i}\rangle \to \sqrt{C_\nu}\delta_{\mu\nu}|\bar{i}\rangle. \tag{7.30}$$

Тогда нетрудно понять, что

$$\sum_{\mu,\nu}\mathbf{R}_\nu\mathbf{M}_\mu|\bar{i}\rangle\otimes|\mu\rangle_E\otimes|\nu\rangle_A = |\bar{i}\rangle\otimes\sum_\nu\sqrt{C_\nu}|\nu\rangle_E\otimes|\nu\rangle_A; \tag{7.31}$$

определяемый $\mathbf{R}_\nu$ супероператор действительно исправляет ошибку. Остается лишь проверить, что $\mathbf{R}_\nu$ удовлетворяют условию нормировки. Имеем

$$\sum_\nu\mathbf{R}_\nu^\dagger\mathbf{R}_\nu = \sum_{\nu,i}\frac{1}{C_\nu}\mathbf{M}_\nu|\bar{i}\rangle\langle\bar{i}|\mathbf{M}_\nu^\dagger, \tag{7.32}$$

что представляет собой ортогональный проектор на пространство состояний, которые достигаются в результате действия ошибок на кодовые слова. Таким образом, мы можем завершить подробное описание супероператора восстановления, добавив к операторной сумме еще один элемент — проектор на дополнительное подпространство.

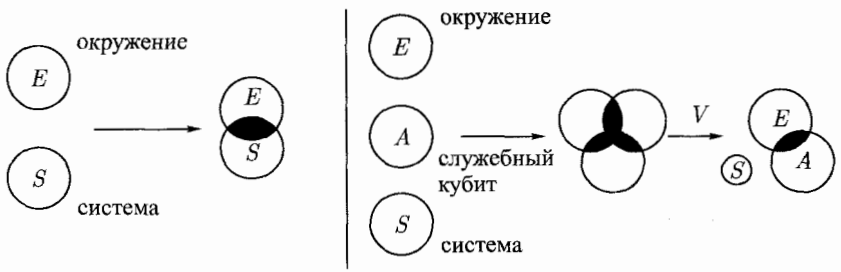
Итак, уравнение (7.19) является достаточным условием исправления ошибок, поскольку для операторов ошибок можно выбрать базис, диагонализирующий матрицу C_{ab} (не обязательно базис операторов Паули), а в этом базисе можно однозначно диагностировать ошибку, выполняя соответствующее ортогональное измерение. (Собственные моды C_{ab} с равными нулю собственными значениями, подобно $\mathbf{Z}_1-\mathbf{Z}_2$ в случае 9-кубитового кода, соответствуют ошибкам, вероятность появления которых равна нулю.) Таким образом, как только совокупность возможных ошибок $\mathcal{E}$ установлена, операция восстановления определена. В частности, не нужна никакая информация о связанных с ошибками $\mathbf{E}_a$ состояниях окружения $|e_a\rangle_E$. Следовательно, код одинаково эффективно борется как с унитарными ошибками, так и с *ошибками декогерентизации* (до тех пор, пока пренебрежимо мала вероятность появления ошибок, не принадлежащих множеству $\mathcal{E}$). Конечно, в случае невырожденного кода C_{ab} диагональна уже в базисе Паули, и мы можем представить базис восстановления в виде

$$\mathbf{R}_a = \sum_i|\bar{i}\rangle\langle\bar{i}|\mathbf{E}_a^\dagger; \tag{7.33}$$

каждому $\mathbf{E}_a$ из $\mathcal{E}$ соответствует $\mathbf{R}_a$.

Мы описали коррекцию ошибок как процедуру, состоящую из двух этапов: во-первых, для выявления ошибки проводится коллективное измерение, а во-вторых, для ее исправления осуществляется обусловленное результатом измерения унитарное преобразование. Эта точка зрения имеет много достоинств. В частности, именно процедура квантового измерения, по-видимому, позволяет укротить континуум возможных ошибок, поскольку измерение проецирует поврежденное состояние на один из дискретного множества результатов, для каждого из которых существует инструкция по восстановлению. Но в действительности измерение — не самый важный этап процесса коррекции квантовых ошибок. Конечно, супероператор восстановления (7.31) может рассматриваться как ортогональное преобразование, действующее на кодовый блок и служебный кубит. Этот супероператор может описывать следующее за унитарным оператором измерение, если мы представим, что служебный кубит подвергается ортогональному измерению, но измерение не является необходимым.

В отсутствие измерения мы можем взглянуть с другой стороны на достигаемое в процессе восстановления обращение декогерентизации. Когда кодовый блок взаимодействует с окружением, он запутывается с ним. В результате неймановская энтропия окружения (как и энтропия кодового блока) возрастает. Если мы не в состоянии управлять окружением, то рост его энтропии никогда не будет обращен; почему в таком случае возможна коррекция квантовых ошибок? Предоставляемый уравнением (7.31) ответ состоит в том, что мы можем применить унитарное преобразование к информации и служебному кубиту, которыми мы *действительно* управляем. Если критерии коррекции квантовых ошибок удовлетворены, то можно выбрать унитарное преобразование, позволяющее запутывание информации с окружением трансформировать в запутывание служебного кубита с окружением, восстанавливая тем самым чистоту информации, как это показано на рисунке.



В то время как измерение не является обязательной составной частью процедуры коррекции ошибок, служебный кубит абсолютно необходим. Он играет роль депозитария для энтропии, вносимой в кодовый блок ошибками — он «разогревается», тогда как информация «охлаждается». Если мы должны в течение длительного времени продолжать защиту квантовой информации, хранящейся в квантовой памяти, то для этой цели необходимо наладить непрерывную поставку служебных кубитов, которые можно отбрасывать после использования. Если же служебный кубит используется повторно, то для этого он должен быть предварительно очищен. Как обсуждалось в первой главе, удаление является диссипативным процессом. Следовательно, согласно принципам термодинамики, коррекция (квантовых) ошибок требует энергетических затрат. Ошибки являются причиной проникновения энтропии в информацию. С помощью обратимого процесса эту энтропию можно перенести на служебный кубит, но для того чтобы откачать ее из служебного кубита и вернуть в окружающую среду, необходимо совершить определенную работу.

7.3. Некоторые основные свойства ККК0

7.3.1. Расстояние

Говорят, что квантовый код является *двоичным*, если он может быть представлен на языке кубитов. В двоичном коде кодовое подпространство размерности 2^k погружено в пространство размерности 2^n, где k и $n\,(>k)$ — целые числа. В сущности, нет никакой необходимости требовать, чтобы размерности этих пространств были степенями двойки (смотри упражнения); тем не менее, мы здесь главным образом будем ограничиваться двоичным кодированием как наиболее простым.

В дополнение к размеру блока n и количеству закодированных кубитов k, еще одним важным параметром, характеризующим код, является расстояние d. Расстояние d представляет собой минимальный вес оператора Паули $\mathbf{E}$, такого, что

$$\langle \bar{i}|\mathbf{E}_a|\bar{j}\rangle \neq C_a\delta_{ij}. \tag{7.34}$$

Квантовый код с размером n, количеством закодированных кубитов k и расстоянием d будет кратко обозначаться символом $[[n,k,d]]$. Обозначение с двойными скобками используется для квантового кода, чтобы отличить его от обозначения $[n,k,d]$ для классического кода.

Будем говорить, что ККК0 может исправить t ошибок, если множество $\mathcal{E}$ допускающих исправление ошибок $\mathbf{E}_a$ включает все операторы Паули с весом, не превышающим t. Наше определение расстояния подразумевает,

что критерию коррекции ошибок

$$\langle \bar{i}|\mathbf{E}_a^\dagger \mathbf{E}_b|\bar{j}\rangle = C_{ab}\delta_{ij} \tag{7.35}$$

удовлетворяют все операторы Паули $\mathbf{E}_a$ с весом, не превышающим t, при условии, что $d \geqslant 2t+1$. Следовательно, ККК0 с расстоянием $d = 2t+1$ может исправить t ошибок.

7.3.2. Локализованные ошибки

Код с расстоянием $d \geqslant 2t+1$ может исправить t ошибок, независимо от их положения в кодовом блоке. Но иногда мы можем знать, что некоторые кубиты особенно предрасположены к появлению ошибок. Возможно, мы видели, как по ним ударили молотком. Или, может быть, вы послали мне блок из n кубитов, но $t\,(<n)$ из них оказались потерянными и уже никогда не будут получены. Я уверен, что остальные $n-t$ кубитов были хорошо упакованы и получены неповрежденными. Тогда я заменяю t недостающих кубитов (произвольно выбранным) состоянием $|00\ldots0\rangle$, вполне отдавая себе отчет в том, что эти кубиты могут содержать ошибки.

Тот же самый код может защитить от большего количества ошибок, если они появляются в известных местах. Фактически ККК0 с расстоянием $d = t+1$ может исправить t ошибок в известных положениях. В этом случае множество ошибок $\mathcal{E}$, которые необходимо исправить, представляет собой совокупность всех операторов Паули с *носителем* в t определенных местоположениях (каждый $\mathbf{E}_a$ действует тривиально на другие $n-t$ кубитов). Но тогда для каждого $\mathbf{E}_a$ и $\mathbf{E}_b$ из $\mathcal{E}$ произведение $\mathbf{E}_a^\dagger \mathbf{E}_b$ также имеет вес не больше, чем t. Следовательно, критерий коррекции ошибок удовлетворяется для всех $\mathbf{E}_{a,b} \in \mathcal{E}$, при условии, что код имеет расстояние, по крайней мере равное $t+1$.

В частности, ККК0, корректирующий t ошибок, находящихся в произвольных положениях, может исправить $2t$ ошибок в известных положениях.

7.3.3. *Обнаружение ошибок*

В некоторых случаях оказывается достаточным просто заметить ошибку, даже если мы не можем полностью диагностировать или исправить ее. Предназначенное для регистрации ошибок измерение имеет два возможных результата: «good» и «bad». Если получается результат «good», мы уверены, что квантовое состояние не повреждено. Если получается результат «bad», значит, состояние было повреждено, и от него следует избавиться.

Если носитель супероператора ошибки принадлежит множеству всех операторов Паули $\mathcal{E}$ с весом, не превышающим t, и возможно измерение, точно показывающее, возникла ошибка или нет, то в таком случае говорят, что мы можем обнаружить t ошибок. Обнаружение ошибок проще, чем их коррекция, поэтому один и тот же код может детектировать больше ошибок, чем исправить. Фактически, ККК0 с расстоянием $d = t + 1$ может обнаружить t ошибок.

Такой код обладает свойством

$$\langle \bar{i}|\mathbf{E}_a|\bar{j}\rangle = C_a\delta_{ij} \tag{7.36}$$

для каждого оператора Паули $\mathbf{E}_a$ с весом не выше t, или

$$\mathbf{E}_a|\bar{i}\rangle = C_a|\bar{i}\rangle + |\varphi_{ai}^{\perp}\rangle, \tag{7.37}$$

где $|\varphi_{ai}^{\perp}\rangle$ — ненормированный вектор, ортогональный кодовому подпространству. Следовательно, действие супероператора ошибки с носителем в $\mathcal{E}$ на состояние кодового подпространства $|\psi\rangle$ имеет вид

$$|\psi\rangle\otimes|0\rangle_E \to \sum_{\mathbf{E}_a\in\mathcal{E}} \mathbf{E}_a|\psi\rangle\otimes|e_a\rangle_E = |\psi\rangle\otimes\sum_{\mathbf{E}_a\in\mathcal{E}} C_a\otimes|e_a\rangle_E + |\mathrm{orthog}\rangle, \tag{7.38}$$

где $|\mathrm{orthog}\rangle$ обозначает вектор, ортогональный кодовому подпространству.

Теперь мы можем выполнить «грубое» ортогональное измерение информации, с двумя результатами: состояние проецируется либо на подпространство кодов, либо на дополнительное ему подпространство. Первый результат воспроизводит неповрежденное состояние $|\psi\rangle$, второй — сообщает об обнаружении ошибки. Таким образом, ККК0 с расстоянием d может регистрировать $d-1$ ошибок. В частности, если ККК0 может исправить t ошибок, то детектировать он может $2t$ ошибок.

7.3.4. Квантовые коды и запутывание

ККК0 защищает квантовую информацию от ошибок, кодируя ее *нелокальным* образом, то есть распределяя ее между несколькими кубитами в блоке. Таким образом, квантовое кодовое слово представляет собой сильно запутанное состояние.

Фактически, невырожденный код с расстоянием $d = t + 1$ обладает следующим свойством. Выберем любое, принадлежащее кодовому подпространству, состояние $|\psi\rangle$ и любые t кубитов в блоке. Возьмем след по оставшимся $n - t$ кубитам, чтобы получить матрицу плотности t кубитов

$$\boldsymbol{\rho}^{(t)} = \mathrm{tr}_{(n-t)}\,|\psi\rangle\langle\psi|. \tag{7.39}$$

Тогда эта матрица плотности абсолютно случайна

$$\boldsymbol{\rho}^{(t)} = \frac{1}{2^t}\mathbf{1}. \tag{7.40}$$

(Наблюдая любые t кубитов в блоке, мы ничего не можем узнать об информации, закодированной кодом с расстоянием $t+1$; то есть $\boldsymbol{\rho}^{(t)}$ — независимая от кодового слова константа. Но матрица плотности t кубитов действительно будет кратной единичному оператору, если только код невырожден.)

Чтобы проверить свойство (7.40), заметим, что для невырожденного кода с расстоянием $t+1$

$$\langle \bar{i}|\mathbf{E}_a|\bar{j}\rangle = 0 \tag{7.41}$$

для любого $\mathbf{E}_a$ ненулевого веса, не превышающего t. Так что

$$\mathrm{tr}(\boldsymbol{\rho}^{(t)}\mathbf{E}_a) = 0, \tag{7.42}$$

для любого, отличного от единичного, t-кубитового оператора Паули. Теперь $\boldsymbol{\rho}^{(t)}$, подобно любой эрмитовой $2^t \times 2^t$-матрице, может быть разложена по операторам Паули:

$$\boldsymbol{\rho}^{(t)} = \frac{1}{2^t}\mathbf{1} + \sum_{\mathbf{E}_a \neq \mathbf{1}} \rho_a \mathbf{E}_a. \tag{7.43}$$

Так как операторы $\mathbf{E}_a$ удовлетворяют условию

$$\frac{1}{2^t}\,\mathrm{tr}(\mathbf{E}_a\mathbf{E}_b) = \delta_{ab}, \tag{7.44}$$

мы находим, что все $\rho_a = 0$, и приходим к выводу, что $\boldsymbol{\rho}^{(t)}$ пропорциональна единичному оператору.

7.4. Вероятность сбоя

7.4.1. Нижняя граница точности воспроизведения

Если носитель супероператора ошибок содержит только операторы Паули из множества $\mathcal{E}$, способ исправления которых нам известен, то закодированная квантовая информация может быть восстановлена с идеальной точностью воспроизведения. Однако в реальной ситуации всегда существует небольшая, но отличная от нуля, вероятность появления ошибок, которые не входят в $\mathcal{E}$, так что восстановленное состояние не будет идеальным. Что можно сказать о точности воспроизведения восстановленного состояния?

Разложение супероператора ошибок по операторам Паули можно разбить на сумму «хороших» (входящих в $\mathcal{E}$) и «плохих» (не входящих в $\mathcal{E}$) операторов. В соответствии с этим результат его действия на состояние кодового подпространства $|\psi\rangle$ можно представить в виде

$$\begin{aligned} |\psi\rangle \otimes |0\rangle_E \to & \sum_a \mathbf{E}_a |\psi\rangle \otimes |e_a\rangle_E = \\ = & \sum_{\mathbf{E}_a \in \mathcal{E}} \mathbf{E}_a |\psi\rangle \otimes |e_a\rangle_E + \sum_{\mathbf{E}_b \notin \mathcal{E}} \mathbf{E}_b |\psi\rangle \otimes |e_b\rangle_E = \\ = & \, |\mathrm{GOOD}\rangle + |\mathrm{BAD}\rangle. \end{aligned} \tag{7.45}$$

Тогда операция восстановления (унитарное преобразование, действующее на информацию и служебный кубит) отображает $|\mathrm{GOOD}\rangle$ на состояние информации, окружающей среды и служебного кубита $|\mathrm{GOOD}'\rangle$, а $|\mathrm{BAD}\rangle$ — на состояние $|\mathrm{BAD}'\rangle$, так что после восстановления мы получаем состояние

$$|\mathrm{GOOD}'\rangle + |\mathrm{BAD}'\rangle; \tag{7.46}$$

здесь (поскольку, действуя на «хорошее» состояние, восстановление работает идеально)

$$|\mathrm{GOOD}'\rangle = |\psi\rangle \otimes |s\rangle_{EA}, \tag{7.47}$$

где $|s\rangle_{EA}$ — некоторое состояние окружения и служебного кубита.

Предположим, что состояния $|\mathrm{GOOD}\rangle$ и $|\mathrm{BAD}\rangle$ взаимно ортогональны. Это справедливо, если, в частности, все «хорошие» состояния окружения ортогональны всем «плохим» состояниям, то есть если

$$\langle e_a | e_b \rangle = 0 \quad \text{при} \quad \mathbf{E}_a \in \mathcal{E},\ \mathbf{E}_b \notin \mathcal{E}. \tag{7.48}$$

Пусть $\boldsymbol{\rho}_{\mathrm{rec}}$ обозначает матрицу плотности восстановленного состояния, полученную путем вычисления следа по состояниям окружения и служебного кубита, и пусть

$$F = \langle \psi | \boldsymbol{\rho}_{\mathrm{rec}} | \psi \rangle \tag{7.49}$$

— его точность воспроизведения. Теперь, поскольку $|\mathrm{BAD}'\rangle$ ортогонально $|\mathrm{GOOD}'\rangle$ (то есть $|\mathrm{BAD}'\rangle$ не имеет ни одной компоненты вдоль $|\psi\rangle \otimes |s\rangle_{EA}$), точность воспроизведения равна

$$F = \langle \psi | \boldsymbol{\rho}_{\mathrm{GOOD}'} | \psi \rangle + \langle \psi | \boldsymbol{\rho}_{\mathrm{BAD}'} | \psi \rangle, \tag{7.50}$$

где

$$\boldsymbol{\rho}_{\mathrm{GOOD}'} = \mathrm{tr}_{EA}(|\mathrm{GOOD}'\rangle\langle\mathrm{GOOD}'|), \qquad \boldsymbol{\rho}_{\mathrm{BAD}'} = \mathrm{tr}_{EA}(|\mathrm{BAD}'\rangle\langle\mathrm{BAD}'|). \tag{7.51}$$

Следовательно, точность воспроизведения восстановленного состояния удовлетворяет неравенству

$$F \geqslant \langle\psi|\rho_{\mathrm{GOOD}'}|\psi\rangle = \||s\rangle_{EA}\|^2 = \||\mathrm{GOOD}'\rangle\|^2. \tag{7.52}$$

Более того, в силу унитарности операции восстановления $\||\mathrm{GOOD}'\rangle\| = \||\mathrm{GOOD}\rangle\|$ и, следовательно,

$$F \geqslant \||\mathrm{GOOD}\rangle\|^2 = \left\| \sum_{\mathbf{E}_a \in \mathcal{E}} \mathbf{E}_a|\psi\rangle \otimes |e_a\rangle_E \right\|^2. \tag{7.53}$$

Однако в общем случае $|\mathrm{BAD}\rangle$ не обязательно ортогонально $|\mathrm{GOOD}\rangle$, так что $|\mathrm{BAD}'\rangle$ не должно быть ортогонально $|\mathrm{GOOD}'\rangle$. Тогда $|\mathrm{BAD}'\rangle$ может иметь компоненту вдоль $|\mathrm{GOOD}'\rangle$, которая деструктивно интерферирует с $|\mathrm{GOOD}'\rangle$ и, следовательно, понижает точность воспроизведения. Тем не менее мы можем получить нижнюю границу точности воспроизведения и в этом, более общем, случае, разлагая $|\mathrm{BAD}'\rangle$ на компоненту вдоль $|\mathrm{GOOD}'\rangle$ и ортогональную компоненту

$$|\mathrm{BAD}'\rangle = |\mathrm{BAD}'_{\|}\rangle + |\mathrm{BAD}'_{\perp}\rangle. \tag{7.54}$$

Тогда, рассуждая так же, как и выше, получаем

$$F \geqslant \||\mathrm{GOOD}'\rangle + |\mathrm{BAD}'_{\|}\rangle\|^2. \tag{7.55}$$

Конечно, так как и операция ошибки, и операция восстановления унитарно действуют на информацию, окружение и служебный кубит, полное состояние $|\mathrm{GOOD}'\rangle + |\mathrm{BAD}'\rangle$ нормировано, или

$$\||\mathrm{GOOD}'\rangle + |\mathrm{BAD}'_{\|}\rangle\|^2 + \||\mathrm{BAD}'_{\perp}\rangle\|^2 = 1, \tag{7.56}$$

а неравенство (7.55) приобретает вид

$$F \geqslant 1 - \||\mathrm{BAD}'_{\perp}\rangle\|^2. \tag{7.57}$$

Наконец, норма вектора $|\mathrm{BAD}'_{\perp}\rangle$ не может превысить норму вектора $|\mathrm{BAD}'\rangle$ и, следовательно,

$$1 - F \leqslant \||\mathrm{BAD}'\rangle\|^2 = \||\mathrm{BAD}\rangle\|^2 \equiv \left\| \sum_{\mathbf{E}_b \notin \mathcal{E}} \mathbf{E}_b|\psi\rangle \otimes |e_b\rangle_E \right\|^2. \tag{7.58}$$

Это наиболее общая нижняя граница «вероятности сбоя» операции восстановления. Уравнение (7.53) следует отсюда в частном случае, когда $|\mathrm{GOOD}\rangle$ и $|\mathrm{BAD}\rangle$ являются взаимно ортогональными состояниями.

7.4.2. Некоррелированные ошибки

Рассмотрим теперь некоторые приложения этих результатов для случая, когда действующие на отдельные кубиты ошибки полностью некоррелированы. В этом случае супероператор ошибок является тензорным произведением однокубитовых супероператоров. Если на самом деле ошибки одинаково действуют на все кубиты, то мы можем представить n-кубитовый супероператор как

$$\$_{\text{error}}^{(n)} = \left[\$_{\text{error}}^{(1)}\right]^{\otimes n}, \tag{7.59}$$

где $\$_{\text{error}}^{(1)}$ — однокубитовый супероператор, действие которого (в его унитарном представлении) имеет вид

$$|\psi\rangle\otimes|0\rangle_E \to |\psi\rangle\otimes|e_I\rangle_E + \mathbf{X}|\psi\rangle\otimes|e_X\rangle_E + \mathbf{Y}|\psi\rangle\otimes|e_Y\rangle_E + \mathbf{Z}|\psi\rangle\otimes|e_Z\rangle_E. \tag{7.60}$$

Влияние ошибок на закодированную информацию особенно легко анализировать, если предположить, что каждое из трех состояний окружения $|e_{X,Y,Z}\rangle$ ортогонально состоянию $|e_I\rangle$. В этом случае запись того, возникла ошибка или нет, для каждого кубита постоянно отпечатывается на окружении, и разумно говорить о вероятности ошибки p_{error} для каждого кубита, где

$$\langle e_I|e_I\rangle = 1 - p_{\text{error}}. \tag{7.61}$$

Если наш квантовый код может исправить t ошибок, то «хорошие» операторы Паули имеют вес, не превышающий t, а «плохие» — вес, превышающий t. Тогда восстановление несомненно будет успешным, по крайней мере пока ошибкам не подвергнутся $t+1$ кубитов. Отсюда следует, что точность воспроизведения подчиняется условию

$$1 - F \leqslant \sum_{s=t+1}^{n} \binom{n}{s} p_{\text{error}}^s (1 - p_{\text{error}})^{n-s} \leqslant \binom{n}{t+1} p_{\text{error}}^{t+1}. \tag{7.62}$$

(Для каждого из $\binom{n}{t+1}$ способов выбора $t+1$ положений, вероятность возникновения ошибки во всех этих позициях равна p_{error}^{t+1}, где мы пренебрегаем вероятностью возникновения дополнительных ошибок в остальных $n-t-1$ позициях. Следовательно, окончательное выражение (7.62) определяет верхний предел вероятности возникновения по крайней мере $t+1$ ошибок в блоке из n кубитов.) При малой p_{error} и большом t точность воспроизведения закодированной информации существенно улучшается по сравнению с точностью воспроизведения $F = 1 - O(p)$ незащищенного кубита.

Для действующего на один кубит общего супероператора ошибок не существует четкого понятия «вероятности ошибки»; состояние кубита и окружения, полученное в результате действия оператора Паули $\mathbf{1}$, не ортогонально (и, следовательно, его нельзя полностью отличить) состоянию, полученному в результате действия операторов Паули $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$. В предельном случае, когда декогерентизация вообще отсутствует, «ошибки» возникают в результате действия на кубиты неизвестных унитарных преобразований. (Если бы действующее на кубит унитарное преобразование $\mathbf{U}$ было известно, то мы могли бы исправить «ошибку», просто применив $\mathbf{U}^\dagger$.)

Рассмотрим некоррелированные унитарные ошибки, действующие на n кубитов в кодовом блоке, каждая из которых (с точностью до несущественной фазы) имеет вид

$$\mathbf{U}^{(1)} = \sqrt{1-p} + i\sqrt{p}\mathbf{W}, \tag{7.63}$$

где $\mathbf{W}$ — (бесследовая, эрмитова) линейная комбинация операторов $\mathbf{X}$, $\mathbf{Y}$ и $\mathbf{Z}$, удовлетворяющая условию $\mathbf{W}^2 = \mathbf{1}$. Если приготовлено состояние кубита $|\psi\rangle$, а затем возникает унитарная ошибка (7.63), то точность воспроизведения итогового состояния

$$F = |\langle\psi|\mathbf{U}^{(1)}|\psi\rangle|^2 = 1 - p(1 - (\langle\psi|\mathbf{W}|\psi\rangle)^2) \geqslant 1 - p. \tag{7.64}$$

Если унитарная ошибка (7.63) действует на каждый из n кубитов в кодовом блоке, а результирующее состояние разложено по операторам Паули, как в уравнении (7.45), тогда состояние $|\mathrm{BAD}\rangle$ (возникающее из слагаемых, в которых $\mathbf{W}$ действует по крайней мере на $t+1$ кубитов) имеет норму порядка $(\sqrt{p})^{t+1}$, а неравенство (7.58) приобретает вид

$$1 - F \leqslant O(p^{t+1}). \tag{7.65}$$

Таким образом, кодирование обеспечивает увеличение точности воспроизведения одного и того же порядка независимо от того, возникают ли некоррелированные ошибки вследствие декогерентизации или неизвестного унитарного преобразования.

Чтобы избежать путаницы, подчеркнем значение слова «некоррелированный» для ясного понимания предыдущего обсуждения. Мы рассматриваем действующую на n кубитов унитарную ошибку как «некоррелированную», если она является тензорным произведением однокубитовых унитарных преобразований, независимо от того, как могут быть связаны друг с другом унитарные преобразования, действующие на различные кубиты. Например, коррекция квантовых ошибок эффективно справляется с ошибкой, приводящей к повороту всех кубитов на угол θ вокруг общей оси. Если код может защитить от t некоррелированных ошибок, то точность

воспроизведения после восстановления равна $F = 1 - O(\theta^{2(t+1)})$. Напротив, больше трудностей вызвала бы унитарная ошибка вида $\mathbf{U}^{(n)} \sim \mathbf{1} + i\theta\mathbf{E}_{\mathrm{bad}}^{(n)}$, где $\mathbf{E}_{\mathrm{bad}}^{(n)}$ — n-кубитовый оператор Паули с весом, бóльшим t. В этом случае $|\mathrm{BAD}\rangle$ имеет норму порядка θ, а типичная точность воспроизведения после восстановления равна $F = 1 - O(\theta^2)$.

7.5. Классические линейные коды

Квантовые коды коррекции ошибок впервые были изобретены менее четырех лет тому назад,[1] но классические коды коррекции ошибок имеют гораздо более длинную историю. За последние пятьдесят лет была построена удивительно красивая и мощная теория классического кодирования. Многое из нее может быть использовано при создании КККО. Здесь мы сделаем беглый обзор лишь некоторых элементов классической теории, акцентируя наше внимание на двоичных линейных кодах.

В двоичном коде k битов кодируются двоичной строкой длины n. То есть из 2^n строк длины n мы выбираем подмножество, содержащее 2^k строк — кодовых слов; k-битовое сообщение кодируется путем отбора одного из этих 2^k кодовых слов.

В частном случае двоичного линейного кода кодовые слова образуют k-мерное замкнутое линейное подпространство C двоичного векторного пространства F_2^n. То есть побитовое исключающее ИЛИ (XOR) двух кодовых слов является другим кодовым словом. Пространство кода C натянуто на базис из k векторов $v_1, v_2, \ldots, v_k$; произвольное кодовое слово можно представить в виде линейной комбинации этих базисных векторов

$$v(\alpha_1, \alpha_2, \ldots, \alpha_k) = \sum_i \alpha_i v_i, \tag{7.66}$$

где каждое $\alpha_i \in \{0, 1\}$, а сложение выполняется по модулю 2. Можно сказать, что вектор $v(\alpha_1, \alpha_2, \ldots, \alpha_k)$ длины n кодирует k-битовое сообщение $\alpha = (\alpha_1, \ldots, \alpha_k)$.

k базисных векторов $v_1, v_2, \ldots, v_k$ можно скомпоновать в $k \times n$-матрицу

$$G = \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}, \tag{7.67}$$

[1] Отсчитывая от 1998 года, времени опубликования английского издания этих лекций. — *Прим. ред.*

которая называется *генерирующей матрицей* кода. Тогда в матричных обозначениях уравнение (7.66) может быть переписано как

$$v(\alpha) = \alpha G; \tag{7.68}$$

действующая налево матрица G кодирует сообщение α.

Альтернативный способ описания k-мерного кодового подпространства векторного пространства F_2^n состоит в определении $n-k$ линейных условий. Существует такая $(n-k) \times n$-матрица H, что

$$Hv = 0 \tag{7.69}$$

для всех тех и только тех векторов v, которые принадлежат коду C. Эта матрица H называется матрицей контроля четности кода C. Строки матрицы H образуют $n-k$ линейно независимых векторов, а кодовым пространством является пространство *ортогональных* им векторов. Ортогональность определяется относительно побитового внутреннего произведения по модулю 2; две двоичные строки длины n ортогональны, если они «сталкиваются» («collide» — обе принимают значение 1) в четном количестве позиций. Отметим, что

$$HG^T = 0, \tag{7.70}$$

где G^T — транспонированная матрица G; строки G ортогональны строкам H.

Единственным типом ошибки классического бита является его инвертирование. Возникшую в n-битовой строке ошибку можно характеризовать n-компонентным вектором e, где единицы в e показывают позиции, в которых появились ошибки. Возмущенная ошибкой e строка v принимает вид

$$v \rightarrow v + e. \tag{7.71}$$

Ошибки можно обнаружить, применяя матрицу контроля четности. Если v является кодовым словом, то

$$H(v+e) = Hv + He = He. \tag{7.72}$$

Вектор He называется синдромом ошибки e. Обозначим через $\mathcal{E}$ множество ошибок $\{e_i\}$, которые мы хотим уметь корректировать. Исправление ошибок будет возможно, если и только если все они имеют различные синдромы. Только при этом условии можно однозначно выявить ошибку с синдромом He, а затем исправить ее, инвертируя отмеченные в e биты,

$$v + e \rightarrow (v+e) + e = v. \tag{7.73}$$

С другой стороны, если $He_1 = He_2$ при $e_1 \neq e_2$, то мы можем неправильно интерпретировать ошибку e_1 как e_2. Тогда попытка исправления произведет следующий эффект:

$$v + e_1 \rightarrow v + (e_1 + e_2) \neq v. \tag{7.74}$$

Восстановленное сообщение $v + e_1 + e_2$ принадлежит подпространству кодов, но оно отличается от исходного сообщения v; закодированная информация повреждена.

Расстояние d кода C представляет собой минимальное значение веса ненулевых векторов $v \in C$, где *вес* равен количеству единиц в строке v. Линейный код с расстоянием $d = 2t + 1$ может исправить t ошибок; код сопоставляет конкретный синдром каждой $e \in \mathcal{E}$, где $\mathcal{E}$ содержит все векторы с весом, не превышающим t. Это так, поскольку если $He_1 = He_2$, то

$$0 = He_1 + He_2 = H(e_1 + e_2) \tag{7.75}$$

и, следовательно, $e_1 + e_2 \in C$. Но если e_1 и e_2 не равны между собой и каждый имеет вес, не превышающий t, то вес $e_1 + e_2$ больше нуля, но не превышает $2t$. Но поскольку $d = 2t + 1$, то вектор $e_1 + e_2$ не может принадлежать C. Следовательно, He_1 и He_2 не могут быть равны друг другу.

Полезной конструкцией классической теории кодирования является *дуальный* (*двойственный*) *код*. Мы видели, что генерирующая $k \times n$-матрица G и $(n-k) \times n$-матрица контроля четности H кода C связаны соотношением $HG^T = 0$. Транспонируя это равенство, получим $GH^T = 0$. Таким образом, мы можем рассматривать H как генератор, а G — как матрицу контроля четности $(n-k)$-мерного кода (этот код обозначается как $C^\perp$ и называется дуальным по отношению к C). Другими словами, $C^\perp$ представляет собой ортогональное дополнение C в F_2^n. Вектор является самоортогональным, если он имеет четный вес, следовательно, возможно пересечение C и $C^\perp$. Код содержит дуальный ему код, если все его кодовые слова имеют четный вес и взаимно ортогональны. Если $n = 2k$, то возможно, что $C = C^\perp$, в этом случае говорят, что код C является *самодуальным* (или *самодвойственным*).

В следующем разделе окажется полезным тождество, связывающее код C с дуальным ему кодом $C^\perp$

$$\sum_{v \in C} (-1)^{v \cdot u} = \begin{cases} 2^k & u \in C^\perp, \\ 0 & u \notin C^\perp. \end{cases} \tag{7.76}$$

Нетривиальным содержанием этого тождества является утверждение, что сумма обращается в нуль для $u \notin C^{\perp}$. Это непосредственно вытекает из знакомого тождества

$$\sum_{v \in \{0,1\}^k} (-1)^{v \cdot w} = 0, \qquad w \neq 0, \tag{7.77}$$

где v и w — строки длины k. Мы можем представить $v \in C$ как

$$v = \alpha G, \tag{7.78}$$

где α представляет собой k-мерный вектор. Тогда

$$\sum_{v \in C} (-1)^{v \cdot u} = \sum_{\alpha \in \{0,1\}^k} (-1)^{\alpha \cdot Gu} = 0, \tag{7.79}$$

для $Gu \neq 0$. Так как G, генерирующая матрица кода C, является матрицей контроля четности кода $C^{\perp}$, то мы приходим к выводу, что сумма обращается в нуль при $u \notin C^{\perp}$.

7.6. Коды КШС

При создании квантовых кодов коррекции ошибок можно применять принципы теории классических линейных кодов. Здесь мы опишем семейство ККК0 — кодов Колдербэнка–Шора–Стина (КШС), при построении которых используется понятие дуального кода.

Пусть C_1 — классический линейный код с $(n - k_1) \times n$-матрицей контроля четности H_1, и пусть C_2 — *субкод* кода C_1 с $(n - k_2) \times n$-матрицей контроля четности H_2, где $k_2 < k_1$. Первые $n - k_1$ строк матрицы H_2 совпадают с первыми $n - k_1$ строками матрицы H_1, но в H_2 содержится $k_1 - k_2$ дополнительных линейно независимых строк; таким образом, каждое слово из C_2 содержится в C_1, но слова из C_2 подчиняются некоторым дополнительным линейным условиям.

Субкод C_2 определяет отношение эквивалентности в C_1; мы говорим, что $u, w \in C_1$ эквивалентны ($u \equiv w$), если и только если в C_2 существует v, такое, что $u = w + v$. Классы эквивалентности являются *смежными классами*, порождаемыми субкодом C_2 в C_1.[1]

КШС-код представляет собой квантовый код с $k = k_1 - k_2$, в котором каждому классу эквивалентности, определяемому элементами $v \in C_2$, соответствует кодовое слово. Каждый элемент базиса кодового подпространства может быть представлен в виде

$$|\bar{w}\rangle = \frac{1}{\sqrt{2^{k_2}}} \sum_{v \in C_2} |v + w\rangle \tag{7.80}$$

[1]Или короче, *смежными классами* C_1/C_2. — *Прим. ред.*

— равновзвешенной суперпозиции всех слов смежного класса, которому принадлежит элемент w. Существует $2^{k_1-k_2}$ смежных классов и, следовательно, $2^{k_1-k_2}$ линейно независимых кодовых слов. Состояния $|\bar{w}\rangle$ очевидно нормированы и взаимно ортогональны; то есть $\langle\bar{w}|\bar{w}'\rangle = 0$, если $\bar{w}$ и $\bar{w}'$ принадлежат разным смежным классам.

Рассмотрим теперь, что произойдет с кодовым словом $|\bar{w}\rangle$, если мы применим к нему побитовое преобразование Адамара $\mathbf{H}^{(n)}$

$$\begin{aligned}\mathbf{H}^{(n)} : |\bar{w}\rangle_F \equiv \frac{1}{\sqrt{2^{k_2}}}\sum_{v\in C_2}|v+w\rangle \to \\ \to |\bar{w}\rangle_P \equiv \frac{1}{\sqrt{2^n}}\sum_u \frac{1}{\sqrt{2^{k_2}}}\sum_{v\in C_2}(-1)^{u\cdot v}(-1)^{u\cdot w}|u\rangle = \\ = \frac{1}{\sqrt{2^{n-k_2}}}\sum_{u\in C_2^\perp}(-1)^{u\cdot w}|u\rangle;\end{aligned} \tag{7.81}$$

мы получаем взвешенную фазовыми множителями когерентную суперпозицию слов, принадлежащих дуальному коду $C_2^\perp$ [на последнем этапе мы использовали тождество (7.76)]. И снова в этом последнем выражении очевидно, что кодовое слово зависит лишь от смежного класса C_2, который представляет w, — сдвиг w на элемент, принадлежащий C_2, не влияет на $(-1)^{u\cdot w}$, если u принадлежит дуальному по отношению к C_2 коду.

Теперь предположим, что код C_1 имеет расстояние d_1, а код $C_2^\perp$ — расстояние $d_2^\perp$, такие, что

$$d_1 \geqslant 2t_F + 1, \qquad d_2^\perp \geqslant 2t_P + 1. \tag{7.82}$$

Тогда нетрудно видеть, что соответствующий КШС-код может корректировать t_F инвертированных битов и t_P обращений фазы. Пусть e — двоичная строка длины n, а $\mathbf{E}_e^{\text{flip}}$ обозначает оператор Паули с $\mathbf{X}$, действующими в каждой позиции i, в которой $e_i = 1$; он действует на состояние $|v\rangle$ по правилу

$$\mathbf{E}_e^{\text{flip}} : |v\rangle \to |v+e\rangle. \tag{7.83}$$

Пусть также $\mathbf{E}_e^{\text{phase}}$ обозначает оператор Паули с $\mathbf{Z}$, действующими в каждой позиции i, в которой $e_i = 1$; его действие представляет собой

$$\mathbf{E}_e^{\text{phase}} : |v\rangle \to (-1)^{v\cdot e}|v\rangle, \tag{7.84}$$

что в базисе, повернутом преобразованием Адамара, приобретает вид

$$\mathbf{E}_e^{\text{phase}} : |u\rangle \to |u+e\rangle. \tag{7.85}$$

Теперь, в исходном базисе (F, или «flip»-базис), каждое базисное состояние КШС-кода $|\bar{w}\rangle_F$ представляет собой суперпозицию слов кода C_1. Чтобы диагностировать ошибку инвертирования бита, мы применяем к информации и служебному кубиту унитарное преобразование

$$|v\rangle \otimes |0\rangle_A \to |v\rangle \otimes |H_1 v\rangle_A, \tag{7.86}$$

а затем измеряем служебный кубит. Результат измерения $H_1 e_F$ представляет собой *синдром инвертирования бита*. Если количество инвертированных битов не превышает t_F, то из этого синдрома мы можем сделать правильный вывод о том, что инвертирования возникли в позициях, отмеченных e_F. Применяя $\mathbf{X}$ к кубитам в этих позициях, мы восстанавливаем закодированную информацию.

Для исправления фазовых ошибок мы предварительно выполняем побитовое преобразование Адамара, чтобы перейти от базиса F к базису P («phase»). В базисе P каждое базисное состояние КШС-кода $|\bar{w}\rangle_P$ представляет собой суперпозицию слов кода $C_2^\perp$. Чтобы диагностировать фазовые ошибки, мы выполняем унитарное преобразование

$$|v\rangle \otimes |0\rangle_A \to |v\rangle \otimes |G_2 v\rangle_A \tag{7.87}$$

и измеряем служебный кубит (G_2, генерирующая матрица кода C_2, одновременно является матрицей контроля четности кода $C_2^\perp$). Результат измерения $G_2 e_P$ представляет собой *синдром фазовой ошибки*. Если количество фазовых ошибок не превосходит t_P, то из этого синдрома мы можем сделать правильный вывод о том, что фазовые ошибки возникли в позициях, отмеченных e_P. Применяя $\mathbf{X}$ (в базисе P) к кубитам в этих позициях, мы восстанавливаем закодированную информацию. Наконец, мы еще раз применяем побитовое преобразование Адамара, чтобы вернуть кодовые слова в исходный базис. (Эквивалентно, мы можем исправить фазовые ошибки, применив $\mathbf{Z}$ к поврежденным кубитам, после возвращения в базис F.)

Если e_F имеет вес меньше, чем d_1, а e_P — меньше, чем $d_2^\perp$, то

$$\langle \bar{w} | \mathbf{E}_{e_P}^{\text{phase}} \mathbf{E}_{e_F}^{\text{flip}} | \bar{w}' \rangle = 0 \tag{7.88}$$

(за исключением случая $e_P = e_F = 0$). Любой оператор Паули может быть представлен в виде произведения фазового оператора и оператора инвертирования. С этой точки зрения $\mathbf{Y}$ представляет собой инвертирование бита и обращение фазы, одновременно возмущающие один и тот же кубит. Итак, расстояние d КШС-кода удовлетворяет условию

$$d \geqslant \min(d_1, d_2^\perp). \tag{7.89}$$

КШС-коды обладают особым свойством (отсутствующим у более общих ККХО): процедуру восстановления можно разбить на две отдельные операции, одна корректирует инвертирование битов, а вторая — фазовые ошибки.

Унитарные преобразования (7.86) [или (7.87)] можно осуществить, выполняя простую квантовую схему. Нужно извлечь бит синдрома, связанный с каждой из $n-k_1$ строк матрицы контроля четности H_1. Чтобы найти a-й бит синдрома, мы готовим служебный бит в состоянии $|0\rangle_{A,a}$ и для каждого значения λ с $(H_1)_{a\lambda}=1$ осуществляем вентиль CNOT со служебным битом в качестве цели и кубитом λ в блоке данных в качестве управляющего. После измерения служебный кубит показывает значение контрольного разряда четности $\sum_\lambda (H_1)_{a\lambda} v_\lambda$. Чтобы выявить ошибки инвертирования бита и фазовые ошибки, измеряются различные синдромы. Важный частный случай конструкции КШС возникает, когда код C содержит свой дуальный код $C^\perp$. Тогда мы можем выбрать $C_1=C$, а $C_2=C^\perp\subseteq C$; в обоих базисах, F и P, вычисляется матрица контроля четности кода C, чтобы определить два синдрома.

На рисунке изображена полная квантовая схема простейшего из КШС-кодов — 7-кубитового кода Стина, рассматриваемого в следующем разделе.

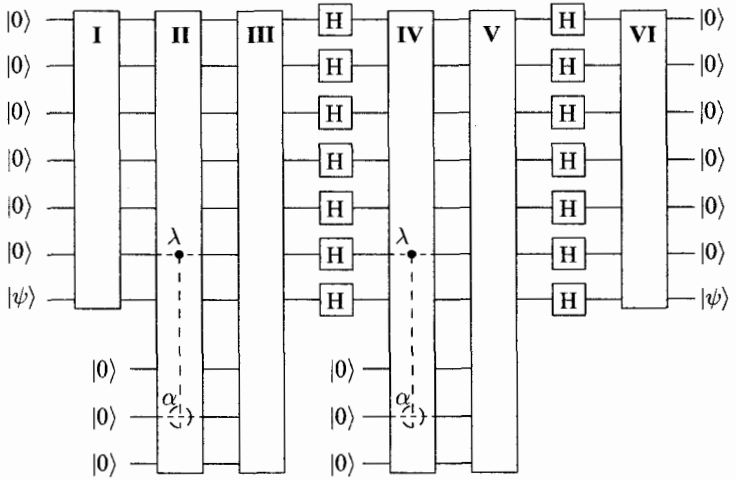
7.7. 7-кубитовый код

Простейшим из КШС-кодов является впервые сформулированный Эндрю Стином квантовый код $[[n,k,d]]=[[7,1,3]]$. Он построен по аналогии с классическим 7-битовым кодом Хэмминга.

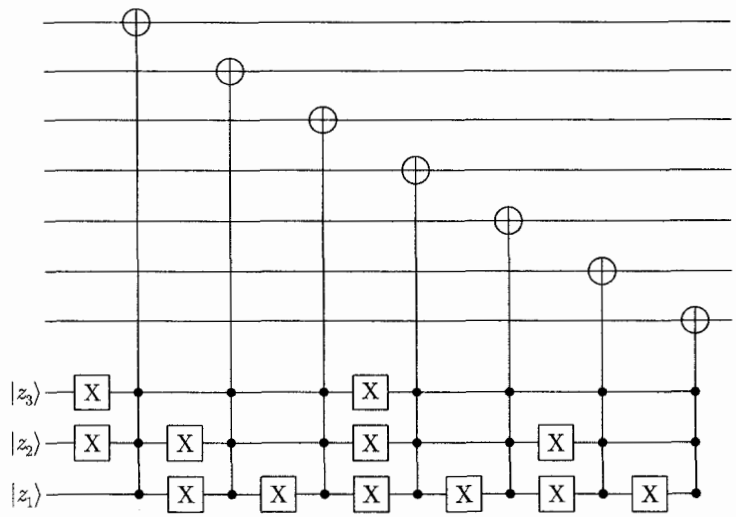
Код Хэмминга представляет собой классический код $[n,k,d]=[7,4,3]$ с 3×7-матрицей контроля четности:

$$H=\begin{pmatrix}1&0&1&0&1&0&1\\0&1&1&0&0&1&1\\0&0&0&1&1&1&1\end{pmatrix}. \tag{7.90}$$

Чтобы увидеть, что расстояние кода $d=3$, заметим сначала, что строка (1110000) с весом, равным трем, проходит контроль четности и, следовательно, принадлежит коду. Теперь нам нужно показать, что в коде нет векторов с меньшими весами. Если e_1 имеет единичный вес, то He_1 является одним из столбцов H. Но ни один из столбцов H не является тривиальным (все нули), поэтому e_1 не может принадлежать коду. Любой вектор с весом два может быть представлен как e_1+e_2, где e_1 и e_2 — различные векторы



I: Кодирование состояния $|\psi\rangle$. См. упражнение 7.6 или рис. 8 на странице 218.
II (IV): Измерение синдрома инвертирования бита (обращения фазы). См. рис. 7 на странице 214.
III (V): Коррекция ошибки инвертирования бита (обращения фазы). См. схему, изображенную ниже.
VI: Декодирование — осуществляется схемой кодирования, выполняемой в обратном направлении.



с единичным весом. Но

$$H(e_1 + e_2) = He_1 + He_2 \neq 0, \tag{7.91}$$

поскольку все столбцы матрицы H различны. Следовательно, $e_1 + e_2$ не может принадлежать коду.

Сами строки матрицы H проходят контроль четности и, следовательно, также принадлежат коду. (Вопреки интуиции, опирающейся на обычную линейную алгебру, ненулевой вектор над конечным полем F_2 может быть ортогонален самому себе.) Генерирующая матрица G кода Хэмминга может быть записана в виде

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}; \tag{7.92}$$

ее первые три строки совпадают со строками матрицы H, а в качестве четвертой строки прилагается кодовое слово (1110000) с весом три.

Дуальным по отношению к коду Хэмминга является код $[7, 3, 4]$, генерируемый матрицей H. В этом случае дуальный код на самом деле содержится в исходном коде — фактически это четный субкод кода Хэмминга, содержащий все те и только те кодовые слова, которые имеют четный вес. Нечетное кодовое слово (1110000) является представителем нетривиального смежного класса четного субкода. Для построения КШС-кода, выберем в качестве C_1 код Хэмминга, а дуальный ему четный субкод — в качестве C_2. Следовательно, $C_2^{\perp} = C_1$ также является кодом Хэмминга; мы будем использовать контроль четности Хэмминга для обнаружения ошибок инвертирования битов в базисе F и обращения фаз в базисе P.

Два ортонормированных кодовых слова этого КШС-кода, каждое из которых связано со своим смежным классом четного субкода, можно выразить в базисе F как

$$\begin{aligned} |\bar{0}\rangle_F &= \frac{1}{\sqrt{8}} \sum_{\{\text{even } v\} \in \text{Hamming}} |v\rangle, \\ |\bar{1}\rangle_F &= \frac{1}{\sqrt{8}} \sum_{\{\text{odd } v\} \in \text{Hamming}} |v\rangle. \end{aligned} \tag{7.93}$$

Так как $|\bar{0}\rangle$ и $|\bar{1}\rangle$ являются суперпозициями кодовых слов Хэмминга, то инвертирование битов можно диагностировать в этом базисе, осуществляя

контроль четности H. После адамаровского поворота базиса эти кодовые слова принимают вид

$$\mathbf{H}^{(7)}: \begin{aligned} |\bar{0}\rangle_F &\to |\bar{0}\rangle_P \equiv \frac{1}{4} \sum_{v\in\text{Hamming}} |v\rangle = \frac{1}{\sqrt{2}}(|\bar{0}\rangle_F + |\bar{1}\rangle_F), \\ |\bar{1}\rangle_F &\to |\bar{1}\rangle_P \equiv \frac{1}{4} \sum_{v\in\text{Hamming}} (-1)^{\text{wt}(v)}|v\rangle = \frac{1}{\sqrt{2}}(|\bar{0}\rangle_F - |\bar{1}\rangle_F). \end{aligned} \tag{7.94}$$

В этом базисе состояния также являются суперпозициями кодовых слов Хэмминга, так что инвертирование битов в базисе P (обращения фаз в исходном базисе) снова можно выявить с помощью контроля четности H. [Попутно отметим, что для этого кода выполнение побитового преобразования Адамара также осуществляет поворот Адамара закодированной информации; этот момент будет важен при обсуждении помехоустойчивых квантовых вычислений (см. приложение)].

Квантовый код Стина может исправить одно инвертирование бита и одно обращение фазы любого одного из семи кубитов в блоке. Но восстановление не удастся, если инвертирование бита или обращение фазы одновременно возникает в двух разных кубитах. Если e_1 и e_2 — две различные строки единичного веса, то $He_1 + He_2$ представляет собой сумму двух различных столбцов матрицы H и, следовательно, совпадает с одним из ее пяти остальных столбцов (все семь нетривиальных строк длины 3 выступают в качестве столбцов H). Поэтому существует другая строка с единичным весом, такая, что $He_1 + He_2 = He_3$, или

$$H(e_1 + e_2 + e_3) = 0; \tag{7.95}$$

таким образом, $e_1 + e_2 + e_3$ является словом кода Хэмминга с весом 3. Будем интерпретировать синдром He_3 как признак возникновения ошибки $v \to v + e_3$ и попытаемся восстановить информацию, применяя операцию $v \to v + e_3$. Тогда совокупным эффектом от двух ошибок инвертирования битов и нашей неудачной попытки восстановления будет добавление к информации $e_1 + e_2 + e_3$ (кодовое слово Хэмминга с нечетным весом), что приведет к инвертированию *закодированного* кубита

$$|\bar{0}\rangle_F \leftrightarrow |\bar{1}\rangle_F. \tag{7.96}$$

Аналогично, обращение двух фазовых множителей в базисе F эквивалентно инвертированию двух битов в базисе P, что (после неудачного восстановления) вызывает в закодированном кубите

$$|\bar{0}\rangle_P \leftrightarrow |\bar{1}\rangle_P, \tag{7.97}$$

или, что эквивариантно,

$$|\bar{0}\rangle_F \to |\bar{0}\rangle_F,$$
$$|\bar{1}\rangle_F \to -|\bar{1}\rangle_F \tag{7.98}$$

— обращение фазы закодированного кубита в базисе F. Восстановление будет успешным только в случае одновременного инвертирования одного бита и обращения одного фазового множителя (в одном или разных кубитах).

7.8. Некоторые ограничения на параметры кода

Код Шора защищает один закодированный кубит от ошибки в любом одном из девяти кубитов блока, а код Стина уменьшает размер блока с девяти до семи. Можно ли добиться лучших результатов?

7.8.1. Квантовая граница Хэмминга

Чтобы понять, насколько лучших результатов мы можем добиться, посмотрим, можно ли при данных значениях n и k найти какие-либо границы для расстояния $d = 2t + 1$ квантового кода $[[n, k, d]]$. На первых порах ограничим наше внимание *невырожденными* кодами, сопоставляющими свой синдром каждой возможной ошибке. Для данного кубита существуют три возможные линейно независимые ошибки $\mathbf{X}$, $\mathbf{Y}$ или $\mathbf{Z}$. В блоке из n кубитов имеется $\binom{n}{j}$ способов выбрать j возмущенных ошибками кубитов и три возможные ошибки для каждого из них; следовательно, общее количество возможных ошибок с весом, не превышающим t, равно

$$N(t) = \sum_{j=0}^{t} 3^j \binom{n}{j}. \tag{7.99}$$

Если закодировано k кубитов, то существует 2^k линейно независимых кодовых слов. Если все векторы $\mathbf{E}_a|\bar{j}\rangle$ линейно независимы, где $\mathbf{E}_a$ — произвольная ошибка с весом, не превышающим t, а $|\bar{j}\rangle$ — произвольный элемент базиса кодовых слов, то размерность 2^n гильбертова пространства n кубитов должна быть достаточно большой, чтобы вместить $N(t) \cdot 2^k$ независимых векторов; следовательно,

$$N(t) = \sum_{j=0}^{t} 3^j \binom{n}{j} \leqslant 2^{n-k}. \tag{7.100}$$

Этот результат называется квантовой границей Хэмминга. Аналогичная граница, но без множителя 3^j, применяется к классическим блоковым кодам, поскольку в этом случае существует только один тип ошибки (инвертирование бита), который может повредить классический бит. Подчеркнем также, что квантовая граница Хэмминга применима только в случае невырожденного кодирования, в то время как классическая граница Хэмминга применима в общем случае. Однако до сих пор не было создано ни одного вырожденного квантового кода, нарушающего квантовую границу Хэмминга (по положению на январь 1999 года).

В частном случае кода с одним закодированным кубитом ($k = 1$), исправляющим одну ошибку ($t = 1$), квантовая граница Хэмминга выглядит как

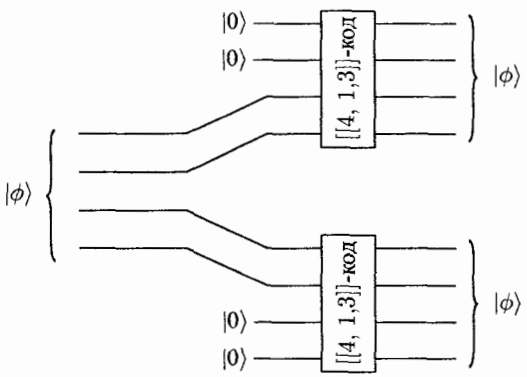
$$1 + 3n \leqslant 2^{n-1}, \tag{7.101}$$

что удовлетворяется при $n \geqslant 5$. Фактически при $n = 5$ это неравенство переходит в равенство ($1 + 15 = 16$). Невырожденный квантовый код $[[5, 1, 3]]$, если такой существует, является *совершенным*[1]. Необходимо все 32-мерное гильбертово пространство пяти кубитов, чтобы вместить все возможные действующие на кодовые слова однокубитовые ошибки — излишков пространства нет.

7.8.2. Граница невозможности клонирования

Было бы удивительно, если бы существовал *вырожденный* код $n = 4$, способный исправить одну ошибку. В действительности нетрудно увидеть, что существование такого кода невозможно. Мы уже знаем, что корректирующий t произвольно расположенных ошибок код можно также использовать для коррекции $2t$ ошибок в известных позициях. Предположим, что мы имеем некий квантовый код $[[4, 1, 3]]$. Тогда мы могли бы закодировать один кубит в 4-кубитовом блоке и разделить его на два субблока, по два кубита в каждом (см. рис. на с. 42). При добавлении к каждому из этих субблоков $|00\rangle$ исходный блок порождает двух потомков с двумя локализованными ошибками в каждом. Если бы мы могли исправить две локализованные ошибки в каждом из этих потомков, то получили бы две идентичные копии исходного блока, то есть клонировали неизвестное квантовое состояние, что невозможно. Следовательно, квантовый код $[[4, 1, 3]]$ невозможен. Таким образом, $n = 5$ представляет собой минимальный размер блока квантового кода коррекции ошибок, независимо от того, вырожден он или нет.

[1] По определению классический код называется *совершенным*, если он насыщает классическую границу Хэмминга. Классический совершенный код, исправляющий t ошибок, может исправить все ошибки с весом, не превышающим t, и не может исправить ни одной ошибки с весом, большим t. Здесь это понятие распространяется на квантовые коды. — *Прим. ред.*



Аналогичные рассуждения показывают, что код $[[n, k \geqslant 1, d]]$ возможен только при

$$n > 2(d-1). \tag{7.102}$$

7.8.3. Квантовая граница Синглтона

Сейчас мы увидим, что результат (7.102) можно усилить до

$$n - k \geqslant 2(d-1). \tag{7.103}$$

Неравенство (7.103) напоминает границу Синглтона для параметров классического кода

$$n - k \geqslant d - 1 \tag{7.104}$$

и поэтому называется «квантовой границей Синглтона». Для классического *линейного* кода неравенство (7.104) почти тривиально: код может иметь расстояние d, если только любые $d-1$ столбцов матрицы контроля четности линейно независимы. Так как столбцы имеют длину $n-k$, то линейно независимыми могут быть не более чем $n-k$ столбцов; следовательно, $d-1$ не может превосходить $n-k$. Классическая граница Синглтона справедлива и для нелинейных кодов.

Можно найти изящное доказательство квантовой границы Синглтона, использующее субаддитивность рассмотренной в разделе 5.2 энтропии фон Неймана. Для начала определим k-кубитовое вспомогательное (служебное) пространство и построим чистое состояние, в котором векторы этого служебного пространства масимально запутаны с 2^k кодовыми словами ККО

$$|\Psi\rangle_{AQ} = \frac{1}{\sqrt{2^k}} \sum |x\rangle_A |\bar{x}\rangle_Q, \tag{7.105}$$

где $\{|x\rangle_A\}$ обозначает ортонормированный базис 2^k-мерного служебного гильбертова пространства k кубитов, а $\{|\bar{x}\rangle_Q\}$ — ортонормированный базис 2^k-мерного кодового подпространства. Вычисляя след по кодовому блоку Q длины n, получим матрицу плотности служебных кубитов $\rho_A = \frac{1}{2^k}\mathbf{1}$, которой соответствует энтропия

$$S(A) = k = S(Q). \tag{7.106}$$

Теперь, если код имеет расстояние d, то может быть исправлено $d-1$ локализованных ошибок, или, как мы уже видели, ни одна наблюдаемая, действующая на $d-1$ кубитов из n, не может открыть никакой информации о закодированном состоянии. Или, что эквивалентно, наблюдаемая ничего не может сказать о состоянии служебных кубитов, закодированном в состоянии $|\Psi\rangle$.

Поскольку мы только что узнали, что $n > 2(d-1)$ (если $k \geqslant 1$), то мысленно разделим кодовый блок Q на три непересекающиеся части: множество $d-1$ кубитов $Q^{(1)}_{d-1}$, другое непересекающееся с предыдущим множество $d-1$ кубитов $Q^{(2)}_{d-1}$ и множество остальных кубитов $Q^{(3)}_{n-2(d-1)}$. Если вычислить след по состояниям кубитов из $Q^{(2)}$ и $Q^{(3)}$, то полученная в результате матрица плотности не должна содержать никаких корреляций между $Q^{(1)}$ и служебными кубитами A. Это означает, что энтропия системы $AQ^{(1)}$ аддитивна

$$S(Q^{(2)}Q^{(3)}) = S(AQ^{(1)}) = S(A) + S(Q^{(1)}). \tag{7.107}$$

Аналогично,

$$S(Q^{(1)}Q^{(3)}) = S(AQ^{(2)}) = S(A) + S(Q^{(2)}). \tag{7.108}$$

Более того, в общем случае энтропия фон Неймана субаддитивна, так что

$$\begin{aligned} S(Q^{(1)}Q^{(3)}) &\leqslant S(Q^{(1)}) + S(Q^{(3)}), \\ S(Q^{(2)}Q^{(3)}) &\leqslant S(Q^{(2)}) + S(Q^{(3)}). \end{aligned} \tag{7.109}$$

Объединяя эти неравенства с предыдущими уравнениями, получим

$$\begin{aligned} S(A) + S(Q^{(2)}) &\leqslant S(Q^{(1)}) + S(Q^{(3)}), \\ S(A) + S(Q^{(1)}) &\leqslant S(Q^{(2)}) + S(Q^{(3)}). \end{aligned} \tag{7.110}$$

Оба эти неравенства могут быть одновременно удовлетворены, если только

$$S(A) \leqslant S(Q^{(3)}). \tag{7.111}$$

Множество $Q^{(3)}$ имеет размерность $n-2(d-1)$, и его энтропия ограничена сверху этой величиной, так что

$$S(A) = k \leqslant n - 2(d-1), \tag{7.112}$$

что и представляет собой квантовую границу Синглтона.

Код $[[5,1,3]]$ насыщает эту границу, но большинству кодов с другими значениями n и k до нее далеко. Рейнс получил более сильный результат, согласно которому код $[[n,k,2t+1]]$ при $k \geqslant 1$ должен удовлетворять неравенству

$$t \leqslant \left[\frac{n+1}{6}\right], \tag{7.113}$$

где $[x]$ — «целая часть x», то есть наибольшее целое число, не превосходящее x. Таким образом, минимальная длина кода $k=1$, который может исправить $t=1,2,3,4,5$ ошибок, равна $n=5,11,17,23,29$, соответственно. Коды со всеми этими параметрами фактически построены, за исключением кода $[[23,1,9]]$.

7.9. Стабилизирующие коды

7.9.1. Общая формулировка

Мы сможем построить (невырожденный) квантовый код $[[5,1,3]]$, но для этого нам потребуется более мощная процедура построения квантовых кодов, нежели процедура КШС.

Как мы помним, чтобы установить критерий возможности исправления ошибок, оказалось весьма полезно разложить супероператор ошибок по n-кубитовым операторам Паули. Однако до сих пор никак не использовалась групповая структура множества этих операторов (произведение операторов Паули также является оператором Паули). Фактически, мы увидим, что теория групп является мощным инструментом ККК0.

Теперь нам удобнее, чтобы все операторы Паули, действующие на отдельные кубиты, были представлены вещественными матрицами, поэтому здесь символ $\mathbf{Y}$ будет обозначать антиэрмитову матрицу

$$\mathbf{Y} = \mathbf{ZX} = i\boldsymbol{\sigma}_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{7.114}$$

удовлетворяющую условию $\mathbf{Y}^2 = -\mathbf{1}$. Тогда операторы

$$\{\pm\mathbf{1}, \pm\mathbf{X}, \pm\mathbf{Y}, \pm\mathbf{Z}\} \equiv \pm\{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\} \tag{7.115}$$

являются элементами группы восьмого порядка.[1] n-кратное тензорное произведение однокубитовых операторов Паули также образует группу

$$G_n = \pm\{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}^{\otimes n}, \tag{7.116}$$

порядок которой $|G_n| = 2^{2n+1}$ (так как существует 4^n возможных тензорных произведений, а еще один множитель 2 учитывает знаки $\pm$); будем называть G_n *n-кубитовой группой Паули*. (По сути, мы будем использовать термин «группа Паули» для обозначения как самой абстрактной группы G_n, так и ее 2^n-мерного точного унитарного представления тензорными произведениями 2×2-матриц, ее единственного неприводимого представления с размерностью больше единицы.) Отметим, что G_n имеет состоящий из двух элементов центр $Z_2 = \{\pm\mathbf{1}^{\otimes n}\}$. С помощью центра как нормального делителя группы мы получаем фактор-группу $\bar{G}_n = G/Z_2$; эту группу можно также рассматривать как двоичное векторное пространство размерности 2^{2n}, этим свойством мы будем пользоваться ниже.

Группа Паули G_n (ее 2^n-мерное представление), очевидно, обладает следующими свойствами:

(i) Каждый элемент $\mathbf{M} \in G_n$ является унитарным $\mathbf{M}^{-1} = \mathbf{M}^\dagger$.

(ii) Для каждого элемента $\mathbf{M} \in G_n$, $\mathbf{M}^2 = \pm\mathbf{1} = \pm\mathbf{1}^{\otimes n}$. Более того, $\mathbf{M}^2 = +\mathbf{1}$, если количество операторов $\mathbf{Y}$ в соответствующем тензорном произведении четно, и $\mathbf{M}^2 = -\mathbf{1}$, если количество сомножителей $\mathbf{Y}$ нечетно.

(iii) Если $\mathbf{M}^2 = +\mathbf{1}$, то элемент $\mathbf{M}$ — эрмитов ($\mathbf{M} = \mathbf{M}^\dagger$); если $\mathbf{M}^2 = -\mathbf{1}$, то элемент $\mathbf{M}$ — антиэрмитов ($\mathbf{M} = -\mathbf{M}^\dagger$).

(iv) Любые два элемента $\mathbf{M}, \mathbf{N} \in G_n$ или коммутируют, или антикоммутируют: $\mathbf{MN} = \pm\mathbf{NM}$.

Мы будем использовать группу Паули, чтобы характеризовать ККК0 следующим образом: пусть S означает абелевую подгруппу n-кубитовой группы Паули G_n. Таким образом, все элементы S, действующие в $\mathcal{H}_{2^n}$, могут быть одновременно диагонализованы. Тогда связанный с S *стабилизирующий код* $\mathcal{H}_S \subseteq \mathcal{H}_{2^n}$ представляет собой общее собственное пространство всех элементов S с равным единице собственным значением.[2] То

[1] Это не группа кватернионов, а другая неабелева группа восьмого порядка — группа симметрии квадрата. Элемент четвертого порядка $\mathbf{Y}$ может рассматриваться как поворот плоскости на 90° тогда как $\mathbf{X}$ и $\mathbf{Z}$ представляют отражения в двух взаимно ортогональных осях.

[2] Стабилизирующие коды в литературе также называют *симплектическими*. — *Прим. ред.*

есть

$$|\psi\rangle \in \mathcal{H}_S, \quad \text{если} \quad \mathbf{M}|\psi\rangle = |\psi\rangle \quad \text{для всех} \quad \mathbf{M} \in S. \tag{7.117}$$

Группа S называется *стабилизатором* кода, так как она сохраняет все кодовые слова.

Группа S может быть охарактеризована ее генераторами. Это *независимые* элементы $\{\mathbf{M}_i\}$ (ни один не может быть представлен как произведение других), такие, что каждый элемент S может быть представлен как произведение элементов $\{\mathbf{M}_i\}$. Если S имеет $n-k$ генераторов, то можно показать, что кодовое пространство $\mathcal{H}_S$ имеет размерность 2^k — существует k закодированных кубитов.

Чтобы убедиться в этом, заметим прежде всего, что каждый элемент $\mathbf{M} \in S$ должен удовлетворять уравнению $\mathbf{M}^2 = +1$; если $\mathbf{M}^2 = -1$, то $\mathbf{M}$ не может иметь собственное значение $+1$. Более того, для каждого $\mathbf{M} \neq \pm 1$ из G_n, квадрат которого равен единице, собственные значения $+1$ и -1 имеют одинаковое вырождение. Это так, поскольку для каждого $\mathbf{M} \neq \pm 1$ существует антикоммутирующий с ним элемент $\mathbf{N} \in G_n$

$$\mathbf{MN} = -\mathbf{NM}; \tag{7.118}$$

следовательно, $\mathbf{M}|\psi\rangle = |\psi\rangle$, если и только если $\mathbf{M}(\mathbf{N}|\psi\rangle) = -\mathbf{N}|\psi\rangle$. Таким образом, действие унитарного оператора $\mathbf{N}$ устанавливает взаимно однозначное соответствие между собственными состояниями оператора $\mathbf{M}$ с собственными значениями $+1$ и -1. Следовательно, существует $\frac{1}{2}2^n = 2^{n-1}$ взаимно ортогональных состояний, удовлетворяющих уравнению

$$\mathbf{M}_1|\psi\rangle = |\psi\rangle, \tag{7.119}$$

где $\mathbf{M}_1$ — один из генераторов S.

Пусть теперь $\mathbf{M}_2$ — другой (коммутирующий с $\mathbf{M}_1$) элемент G_n, такой, что $\mathbf{M}_2 \neq \pm 1, \pm\mathbf{M}_1$. Мы можем найти $\mathbf{N} \in G_n$, коммутирующий с $\mathbf{M}_1$, но антикоммутирующий $\mathbf{M}_2$; следовательно, $\mathbf{N}$ сохраняет собственное пространство оператора $\mathbf{M}_1$ с собственным значением $+1$, но внутри этого пространства обменивает собственные состояния $\mathbf{M}_2$ с собственными значениями $+1$ и -1. Отсюда следует, что пространство, удовлетворяющее уравнению

$$\mathbf{M}_1|\psi\rangle = \mathbf{M}_2|\psi\rangle = |\psi\rangle, \tag{7.120}$$

имеет размерность 2^{n-2}.

Продолжая эту процедуру, заметим, что если $\mathbf{M}_j$ не зависит от $\{\mathbf{M}_1, \mathbf{M}_2, \ldots, \mathbf{M}_{j-1}\}$, то существует оператор $\mathbf{N}$, коммутирующий с $\mathbf{M}_1$, $\ldots, \mathbf{M}_{j-1}$, но антикоммутирующий $\mathbf{M}_j$ (ниже мы обсудим более детально, как можно найти такой оператор $\mathbf{N}$). Следовательно, ограниченный пространством с $\mathbf{M}_1 = \mathbf{M}_2 = \ldots = \mathbf{M}_{j-1} = 1$ оператор $\mathbf{M}_j$ имеет одинаковое количество собственных векторов, соответствующих собственным значениям $+1$ и -1. Поэтому добавление еще одного генератора всегда сокращает размерность общего собственного пространства вдвое. В случае $n-k$ генераторов размерность оставшегося пространства равна $2^n(1/2)^{n-k} = 2^k$.

Язык стабилизатора полезен, поскольку он предоставляет простой способ охарактеризовать ошибки, которые код может обнаружить и исправить. Мы можем рассматривать $n-k$ генераторов стабилизатора $\mathbf{M}_1, \ldots, \mathbf{M}_{n-k}$ как *контролирующие операторы* кода, коллективные наблюдаемые, которые мы измеряем, чтобы диагностировать ошибки. Если закодированная информация не повреждена, то мы найдем $\mathbf{M}_i = 1$ для каждого генератора; но если для некоторого i окажется $\mathbf{M}_i = -1$, тогда данные ортогональны кодовому подпространству и обнаружена ошибка.

Вспомним, что супероператор ошибки может быть разложен по элементам $\mathbf{E}_a$ группы Паули. Конкретный элемент $\mathbf{E}_a$ или коммутирует, или антикоммутирует с конкретным генератором стабилизатора $\mathbf{M}$. Если $\mathbf{E}_a$ и $\mathbf{M}$ коммутируют, то

$$\mathbf{M}\mathbf{E}_a|\psi\rangle = \mathbf{E}_a\mathbf{M}|\psi\rangle = \mathbf{E}_a|\psi\rangle \tag{7.121}$$

для $|\psi\rangle \in \mathcal{H}_S$, так что ошибка сохраняет значение $\mathbf{M} = 1$. Но если $\mathbf{E}_a$ и $\mathbf{M}$ антикоммутируют, то

$$\mathbf{M}\mathbf{E}_a|\psi\rangle = -\mathbf{E}_a\mathbf{M}|\psi\rangle = -\mathbf{E}_a|\psi\rangle, \tag{7.122}$$

так что ошибка инвертирует значение $\mathbf{M}$ и, следовательно, может быть обнаружена путем измерения $\mathbf{M}$.

Для генераторов стабилизатора $\mathbf{M}_i$ и ошибок $\mathbf{E}_a$ мы можем написать

$$\mathbf{M}_i\mathbf{E}_a = (-1)^{s_{ia}}\mathbf{E}_a\mathbf{M}_i. \tag{7.123}$$

Величины s_{ia}, $i = 1, \ldots, n-k$, образуют *синдром* ошибки $\mathbf{E}_a$, поскольку при ее появлении результатом измерения $\mathbf{M}_i$ является $(-1)^{s_{ia}}$. В случае невырожденного кода величины s_{ia} различны для всех $\mathbf{E}_a \in \mathcal{E}$, так что измерение $n-k$ генераторов стабилизатора полностью диагностирует ошибку.

Найдем в более общем виде достаточное условие, которому должен удовлетворять стабилизатор, чтобы обеспечивать возможность коррекции ошибок. Вспомним, что для этого достаточно, чтобы равенство

$$\langle\psi|\mathbf{E}_a^\dagger\mathbf{E}_b|\psi\rangle = C_{ab}, \tag{7.124}$$

где C_{ab} не зависит от $|\psi\rangle$, выполнялось для любых $\mathbf{E}_a, \mathbf{E}_b \in \mathcal{E}$ и нормированного $|\psi\rangle$ из кодового подпространства. Нетрудно видеть, что это условие удовлетворяется, если для любых $\mathbf{E}_a, \mathbf{E}_b \in \mathcal{E}$ выполняется одно из нижеследующих условий:

1) $\mathbf{E}_a^\dagger\mathbf{E}_b \in S$,

2) Существует $\mathbf{M} \in S$, антикоммутирующий с $\mathbf{E}_a^\dagger\mathbf{E}_b$.

Доказательство: В случае (1) $\langle\psi|\mathbf{E}_a^\dagger\mathbf{E}_b|\psi\rangle = \langle\psi|\psi\rangle = 1$ для $|\psi\rangle \in \mathcal{H}_S$. В случае (2) предположим, что $\mathbf{M} \in S$ и $\mathbf{M}\mathbf{E}_a^\dagger\mathbf{E}_b = -\mathbf{E}_a^\dagger\mathbf{E}_b\mathbf{M}$. Тогда

$$\langle\psi|\mathbf{E}_a^\dagger\mathbf{E}_b|\psi\rangle = \langle\psi|\mathbf{E}_a^\dagger\mathbf{E}_b\mathbf{M}|\psi\rangle = -\langle\psi|\mathbf{M}\mathbf{E}_a^\dagger\mathbf{E}_b|\psi\rangle = -\langle\psi|\mathbf{E}_a^\dagger\mathbf{E}_b|\psi\rangle \tag{7.125}$$

и, следовательно, $\langle\psi|\mathbf{E}_a^\dagger\mathbf{E}_b|\psi\rangle = 0$.

Таким образом, *стабилизирующий код*, который корректирует $\{\mathcal{E}\}$, представляет собой пространство $\mathcal{H}_S$, фиксированное абелевой подгруппой S группы Паули, где любой оператор $\mathbf{E}_a^\dagger\mathbf{E}_b$ ($\mathbf{E}_a, \mathbf{E}_b \in \mathcal{E}$) удовлетворяет (1) или (2). Код является *невырожденным*, если условие (1) не выполняется ни для одного из операторов $\mathbf{E}_a^\dagger\mathbf{E}_b$.

Очевидно, мы также вполне могли бы выбрать в качестве кодового подпространства любое из 2^{n-k} совместных собственных пространств $n-k$ независимых коммутирующих элементов группы G_n. Но фактически все эти коды эквивалентны. Мы можем считать два стабилизирующих кода *эквивалентными*, если они отличаются только способом маркировки кубитов и выбором базиса для каждого однокубитового гильбертова пространства, то есть стабилизатор одного кода преобразуется в стабилизатор другого кода путем перестановки кубитов, сопровождаемой тензорным произведением унитарных однокубитовых преобразований. Если мы разделим генераторы стабилизатора на два множества $\{\mathbf{M}_1, \ldots, \mathbf{M}_j\}$ и $\{\mathbf{M}_{j+1}, \ldots, \mathbf{M}_{n-k}\}$, то существует оператор $\mathbf{N} \in G_n$, коммутирующий с каждым элементом первого множества и антикоммутирующий с каждым элементом второго множества. Применение $\mathbf{N}$ к $|\psi\rangle \in \mathcal{H}_S$ сохраняет собственные значения первого множества, одновременно инвертируя собственные значения второго. Поскольку $\mathbf{N}$ явля-

ется тензорным произведением однокубитовых унитарных преобразований, то без потери общности (с точностью до эквивалентности) все собственные значения можно выбрать равными единице. Более того, поскольку знаки минус на самом деле не имеют значения, когда стабилизатор определен, мы вполне можем сказать, что два кода эквивалентны, если, с точностью до фаз, стабилизаторы отличаются перестановкой n кубитов, а также перестановками всех индивидуальных кубитов в операторах $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$.

В процессе восстановления может произойти сбой, если существует оператор $\mathbf{E}_a^\dagger \mathbf{E}_b$, коммутирующий со стабилизатором, но не принадлежащий ему. Этот оператор сохраняет кодовое подпространство $\mathcal{H}_S$, но может действовать в нем нетривиально; таким образом, он может изменять закодированную информацию. Так как $\mathbf{E}_a|\psi\rangle$ и $\mathbf{E}_b|\psi\rangle$ имеют одинаковый синдром, мы можем неправильно принять одну ошибку $\mathbf{E}_a$ за другую — $\mathbf{E}_b$; тогда в результате действия ошибки и попытки ее исправления к информации будет применен оператор $\mathbf{E}_b^\dagger \mathbf{E}_a$, что может оказаться причиной ее повреждения.

Стабилизирующий код с расстоянием d обладает таким свойством, что любой $\mathbf{E} \in G_n$ с меньшим, чем d, весом или принадлежит стабилизатору, или антикоммутирует с его некоторым элементом. Код является невырожденным, если стабилизатор не содержит ни одного элемента с меньшим, чем d, весом. Код с расстоянием $d = 2t + 1$ может исправить t ошибок, а код с расстоянием $s + 1$ может обнаружить s ошибок или исправить s ошибок в известных позициях.

7.9.2. Симплектическая запись

Свойства стабилизирующих кодов часто лучше объясняются и выражаются на языке линейной алгебры. Стабилизатор кода S — абелева подгруппа группы Паули, имеющая порядок 2^{n-k} и состоящая из элементов, квадраты которых равны единице, — может рассматриваться как $(n-k)$-мерное замкнутое линейное подпространство пространства F_2^{2n}, самоортогональное относительно некоторого (симплектического) внутреннего произведения.

Группа $\bar{G}_n = G_n/Z_2$ изоморфна двоичному векторному пространству F_2^{2n}. Мы утверждаем это, замечая, что поскольку $\mathbf{Y} = \mathbf{ZX}$, то любой элемент $\mathbf{M}$ группы Паули (с точностью до знака $\pm$) можно представить в виде произведения операторов $\mathbf{Z}$ и $\mathbf{X}$; мы можем написать

$$\mathbf{M} = \mathbf{Z}_M \cdot \mathbf{X}_M, \tag{7.126}$$

где $\mathbf{Z}_M$ и $\mathbf{X}_M$ — тензорные произведения степеней операторов $\mathbf{Z}$ и $\mathbf{X}$ соответственно. В более явном виде оператор Паули можно записать как

$$(\alpha|\beta) \equiv \mathbf{Z}(\alpha)\mathbf{X}(\beta) = \bigotimes_{i=1}^{n} \mathbf{Z}^{\alpha_i} \cdot \bigotimes_{i=1}^{n} \mathbf{X}^{\beta_i}, \tag{7.127}$$

где α и β — двоичные строки длины n. Тогда операторы $\mathbf{Y}$ действуют в тех позициях, в которых «сталкиваются» α и β.[1] Умножение в $\bar{G}_n$ отображается на сложение в F_2^{2n}:

$$(\alpha|\beta)(\alpha'|\beta') = (-1)^{\alpha' \cdot \beta}(\alpha + \alpha'|\beta + \beta'); \tag{7.128}$$

показатель фазы $\alpha' \cdot \beta$ подсчитывает количество перестановок $\mathbf{Z}$ и $\mathbf{X}$, в процессе преобразования произведения к стандартной форме (7.127).

Из уравнения (7.128) следует, что коммутационные свойства операторов Паули можно представить в виде

$$(\alpha|\beta)(\alpha'|\beta') = (-1)^{\alpha' \cdot \beta + \alpha \cdot \beta'}(\alpha'|\beta')(\alpha|\beta). \tag{7.129}$$

Таким образом, два оператора Паули коммутируют, если и только если соответствующие векторы ортогональны относительно симплектического внутреннего произведения

$$\alpha \cdot \beta' + \alpha' \cdot \beta. \tag{7.130}$$

Отметим также, что квадрат оператора Паули равен

$$(\alpha|\beta)^2 = (-1)^{\alpha \cdot \beta}\mathbf{1}, \tag{7.131}$$

так как $\alpha \cdot \beta$ подсчитывает количество множителей $\mathbf{Y}$ в операторе; квадрат оператора Паули равен единице, если и только если

$$\alpha \cdot \beta = 0. \tag{7.132}$$

Заметим, что замкнутое подпространство, каждый элемент которого обладает этим свойством, автоматически самоортогонально, поскольку

$$\alpha \cdot \beta' + \alpha' \cdot \beta = (\alpha + \alpha') \cdot (\beta + \beta') - \alpha \cdot \beta - \alpha' \cdot \beta' = 0; \tag{7.133}$$

то есть, на языке теории групп, подгруппа G_n, квадрат каждого элемента которой равен единице, автоматически является абелевой.

[1] То есть в тех позициях, в которых в обеих двоичных строках α и β записаны единицы. При записи операторов Паули в виде (7.126), (7.127) использовано свойство тензорного произведения $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$, которое по индукции можно обобщить на произвольное количество сомножителей. — *Прим. ред.*

На языке линейной алгебры, некоторые из сделанных ранее утверждений относительно группы Паули легко проверяются путем подсчета линейных условий. Элементы являются независимыми, если линейно независимы соответствующие векторы в F_2^{2n}, так что мы можем рассматривать $n-k$ генераторов стабилизатора как базис в линейном подпространстве размерности $n-k$. Будем использовать обозначение S для линейного пространства и соответствующей абелевой группы. Тогда $S^\perp$ обозначает векторное пространство размерности $n+k$, ортогональное каждому вектору из S (относительно симплектического внутреннего произведения). Отметим, что $S^\perp$ содержит S, так как все векторы из S взаимно ортогональны. На языке теории групп, пространству $S^\perp$ соответствует нормализующая (или централизующая) группа $N(S)(\equiv S^\perp)$ группы $S \subset G_n$, то есть подгруппа группы G_n, содержащая все элементы, коммутирующие с каждым элементом S. Так как S – абелева группа, она содержится в своем собственном нормализаторе, в который входят и другие элементы (которые мы обсудим ниже). Стабилизатор кода с расстоянием d обладает свойством, согласно которому каждый элемент $(\alpha|\beta)$, вес которого $\sum_i(\alpha_i \vee \beta_i)$ меньше, чем d, или принадлежит подпространству стабилизатора S, или лежит *за пределами* ортогонального пространства $S^\perp$.

Код можно характеризовать его стабилизатором, стабилизатор — его генераторами, а $n-k$ генераторов можно представить матрицей размерности $(n-k)\times 2n$

$$H = (H_Z|H_X). \tag{7.134}$$

Здесь каждая строка представляет собой оператор Паули, записанный в виде $(\alpha|\beta)$. Синдром ошибки $\mathbf{E}_a = (\alpha_a|\beta_a)$ определяется его коммутационными свойствами с генераторами $\mathbf{M}_i = (\alpha_i'|\beta_i')$; то есть

$$s_{ia} = (\alpha_a|\beta_a)\cdot(\alpha_i'|\beta_i') = \alpha_a\cdot\beta_i' + \alpha_i'\cdot\beta_a. \tag{7.135}$$

В случае невырожденного кода каждая ошибка имеет свой собственный синдром. Если код вырожден, то возможно несколько ошибок с одним синдромом, но для их исправления мы можем применить любой из соответствующих наблюдаемому синдрому операторов $\mathbf{E}_a^\dagger$.

7.9.3. Несколько примеров стабилизирующих кодов

(а) Девятикубитовый код. Этот $[[9,1,3]]$-код имеет восемь генераторов стабилизатора, которые можно представить в виде

$$\begin{aligned} &\mathbf{Z}_1\mathbf{Z}_2, \quad \mathbf{Z}_2\mathbf{Z}_3, \quad \mathbf{Z}_4\mathbf{Z}_5, \quad \mathbf{Z}_5\mathbf{Z}_6, \quad \mathbf{Z}_7\mathbf{Z}_8, \quad \mathbf{Z}_8\mathbf{Z}_9, \\ &\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4\mathbf{X}_5\mathbf{X}_6, \quad \mathbf{X}_4\mathbf{X}_5\mathbf{X}_6\mathbf{X}_7\mathbf{X}_8\mathbf{X}_9. \end{aligned} \tag{7.136}$$

В записи (7.134) они принимают вид

$$\left(\begin{array}{cccccccccc|cccccccccc} 1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1 \end{array}\right).$$

(b) 7-кубитовый код. Этот $[[7,1,3]]$-код имеет шесть генераторов стабилизатора, которые можно записать как

$$\tilde{H} = \begin{pmatrix} H_{\text{ham}} & 0 \\ 0 & H_{\text{ham}} \end{pmatrix}, \tag{7.137}$$

где H_{ham} — 3×7-матрица контроля четности классического $[7,4,3]$-кода Хэмминга. Три контрольных оператора

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{Z}_1\mathbf{Z}_3\mathbf{Z}_5\mathbf{Z}_7, \\ \mathbf{M}_2 &= \mathbf{Z}_2\mathbf{Z}_3\mathbf{Z}_6\mathbf{Z}_7, \\ \mathbf{M}_3 &= \mathbf{Z}_4\mathbf{Z}_5\mathbf{Z}_6\mathbf{Z}_7 \end{aligned} \tag{7.138}$$

обнаруживают инвертирования битов, а три контрольных оператора

$$\begin{aligned} \mathbf{M}_4 &= \mathbf{X}_1\mathbf{X}_3\mathbf{X}_5\mathbf{X}_7, \\ \mathbf{M}_5 &= \mathbf{X}_2\mathbf{X}_3\mathbf{X}_6\mathbf{X}_7, \\ \mathbf{M}_6 &= \mathbf{X}_4\mathbf{X}_5\mathbf{X}_6\mathbf{X}_7 \end{aligned} \tag{7.139}$$

обнаруживают фазовые ошибки. Пространство с $\mathbf{M}_1 = \mathbf{M}_2 = \mathbf{M}_3 = 1$ натянуто на кодовые слова, удовлетворяющие контролю четности Хэмминга. Вспоминая, что адамаровское преобразование базиса обменивает операторы $\mathbf{Z}$ и $\mathbf{X}$, мы видим, что пространство с $\mathbf{M}_4 = \mathbf{M}_5 = \mathbf{M}_6 = 1$ натянуто на кодовые слова, удовлетворяющие контролю четности Хэмминга в базисе, полученном преобразованием Адамара. Действительно, мы построили семикубитовый код, требуя, чтобы контроль четности Хэмминга удовлетворялся в обоих базисах. Генераторы коммутируют, потому что код Хэмминга содержит дуальный ему код; то есть каждая строка матрицы H_{ham} удовлетворяет контролю четности Хэмминга.

(c) КШС-коды. Вспомним, что если классический $[n, k, d]$-код C содержит дуальный ему код $C^{\perp}$, мы можем выполнить КШС-конструкцию, чтобы получить квантовый $[[n, 2k-n, d]]$-код. Стабилизатор этого кода можно записать как

$$\tilde{H} = \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix}, \tag{7.140}$$

где $H - (n-k) \times n$-матрица контроля четности кода C. Как и для семикубитового кода, стабилизаторы коммутируют, поскольку C содержит $C^{\perp}$, а кодовое подпространство натянуто на состояния, удовлетворяющие контролю четности H в F- и P-базисах. Или, что эквивалентно, кодовые слова удовлетворяют контролю четности H и инвариантны относительно

$$|v\rangle \rightarrow |v + w\rangle, \tag{7.141}$$

где $w \in C^{\perp}$.

(d) Более общие КШС-коды. Рассмотрим более общий стабилизатор, каждый генератор которого можно выбрать в виде произведения операторов $\mathbf{Z}$ $(= (\alpha|0))$ или операторов $\mathbf{X}$ $(= (0|\beta))$. Тогда генераторы имеют вид

$$\tilde{H} = \begin{pmatrix} H_Z & 0 \\ 0 & H_X \end{pmatrix}. \tag{7.142}$$

Какому условию должны удовлетворять H_X и H_Z, если $\mathbf{Z}$- и $\mathbf{X}$-генераторы коммутируют между собой? Так как операторы $\mathbf{Z}$ должны сталкиваться с операторами $\mathbf{X}$ в четном количестве позиций, мы имеем

$$H_X H_Z^T = H_Z H_X^T = 0. \tag{7.143}$$

Но это всего лишь требование того, чтобы дуальный код $C_X^{\perp}$ с матрицей контроля четности H_X содержался в коде C_Z с матрицей контроля четности H_Z. Другими словами, этот ККК0 входит в семейство КШС-кодов с

$$C_2 = C_X^{\perp} \subseteq C_1 = C_Z. \tag{7.144}$$

Итак, мы можем характеризовать КШС-коды как такие и только такие, стабилизаторы которых имеют генераторы вида (7.142).

Однако следует предостеречь: определяемый уравнением (7.142) код невырожден, если ошибки ограничены весами, меньшими, чем $d = \min(d_Z, d_X)$ (где d_Z — расстояние кода C_Z, а d_X — расстояние кода C_X). Но истинное расстояние ККК0 может превышать d. Например,

9-кубитовый код в этом обобщенном смысле является КШС-кодом. Но в этом случае классический код C_X имеет единичное расстояние, отражая, например, то, что $\mathbf{Z}_1\mathbf{Z}_2$ содержится в стабилизаторе. Тем не менее, расстояние КШС-кода $d = 3$, так как ни один оператор Паули с весом два не принадлежит $S^{\perp} \setminus S$.

7.9.4. Закодированные кубиты

Мы видели, что причиняющие беспокойство ошибки находятся в $S^{\perp} \setminus S$ — те, что коммутируют со стабилизатором, но лежат за его пределами. Эти операторы Паули интересны также и по другой причине: их можно рассматривать как «логические» операторы, действующие на закодированные данные, которые защищены кодом.

С точки зрения линейной алгебры, мы можем видеть, что нормализатор $S^{\perp}$ стабилизатора содержит $n + k$ независимых генераторов. Действительно, подпространство в $2n$-мерном пространстве векторов $(\alpha|\beta)$, содержащее векторы, ортогональные каждому из $n - k$ линейно независимых векторов, имеет размерность $2n - (n - k) = n + k$. Из $n + k$ векторов, образующих линейную оболочку этого пространства, $n - k$ можно выбрать в качестве генераторов самого стабилизатора. Оставшиеся $2k$ генераторов сохраняют кодовое пространство, так как они коммутируют со стабилизатором, но нетривиально действуют на k закодированных кубитов.

Фактически в качестве этих $2k$ операций могут быть выбраны однокубитовые операторы $\bar{\mathbf{Z}}_i$, $\bar{\mathbf{X}}_i$, $i = 1, 2, \ldots, k$, где $\bar{\mathbf{Z}}_i$, $\bar{\mathbf{X}}_i$ — операторы Паули $\mathbf{Z}$ и $\mathbf{X}$, действующие на закодированный кубит, обозначенный индексом i. Во-первых, отметим, что мы можем расширить $n - k$ генераторов стабилизатора до максимального набора n коммутирующих операторов. Добавляемые в наш набор k операторов можно обозначить как $\bar{\mathbf{Z}}_1, \ldots, \bar{\mathbf{Z}}_k$. Тогда мы можем рассматривать общие собственные состояния $\bar{\mathbf{Z}}_1, \ldots, \bar{\mathbf{Z}}_k$ (в кодовом подпространстве $\mathcal{H}_S$) как логические базисные состояния $|\bar{z}_1, \ldots, \bar{z}_k\rangle$, с $\bar{z}_j = 0$, соответствующим $\bar{\mathbf{Z}}_j = 1$, и с $\bar{z}_j = 1$, соответствующим $\bar{\mathbf{Z}}_j = -1$.

Оставшиеся k генераторов нормализатора можно выбрать взаимно коммутирующими, а также коммутирующими со стабилизатором, но тогда они не будут коммутировать с любым из операторов $\bar{\mathbf{Z}}_i$. Осуществляя процедуру ортонормирования Грамма–Шмидта, мы можем выбрать эти генераторы, обозначенные как $\bar{\mathbf{X}}_i$, чтобы диагонализовать симплектическую форму, так что

$$\bar{\mathbf{Z}}_i\bar{\mathbf{X}}_j = (-1)^{\delta_{ij}}\bar{\mathbf{X}}_j\bar{\mathbf{Z}}_i. \tag{7.145}$$

Таким образом, каждый $\bar{\mathbf{X}}_j$ обращает собственное значение соответствующего оператора $\bar{\mathbf{Z}}_j$ и, следовательно, может рассматриваться как оператор Паули $\mathbf{X}$, действующий на j-й закодированный кубит.

(a) 9-кубитовый код. Как мы обсуждали выше, в качестве логических операторов можно выбрать

$$\bar{\mathbf{X}} = \mathbf{X}_1\mathbf{X}_2\mathbf{X}_3, \qquad \bar{\mathbf{Z}} = \mathbf{Z}_1\mathbf{Z}_4\mathbf{Z}_7. \tag{7.146}$$

Они антикоммутируют между собой ($\mathbf{X}$ и $\mathbf{Z}$ сталкиваются в позиции 1), коммутируют с генераторами стабилизатора и не зависят от генераторов (ни один из элементов стабилизатора не содержит три оператора $\mathbf{X}$ или три оператора $\mathbf{Z}$).

(b) 7-кубитовый код. Мы видели, что

$$\bar{\mathbf{X}} = \mathbf{X}_1\mathbf{X}_2\mathbf{X}_3, \qquad \bar{\mathbf{Z}} = \mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_3. \tag{7.147}$$

Тогда $\mathbf{X}$ добавляет нечетное кодовое слово Хэмминга, а $\mathbf{Z}$ обращает его фазу. Эти операции осуществляют инвертирование бита, соответственно, и обращение фазы в базисе $\{|0\rangle_F, |1\rangle_F\}$, определенном в уравнении (7.93).

7.10. 5-кубитовый код

Все рассмотренные до сих пор ККК0 относятся к КШС-типу — каждый генератор стабилизатора является произведением операторов $\mathbf{Z}$ или $\mathbf{X}$. Но не все стабилизирующие коды обладают этим свойством. Примером стабилизирующего кода, не относящиеся к КШС-типу, является совершенный невырожденный $[[5, 1, 3]]$-код.

Его четыре генератора стабилизатора можно представить в виде

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{XZZX1}, \\ \mathbf{M}_2 &= \mathbf{1XZZX}, \\ \mathbf{M}_3 &= \mathbf{X1XZZ}, \\ \mathbf{M}_4 &= \mathbf{ZX1XZ}. \end{aligned} \tag{7.148}$$

Генераторы $\mathbf{M}_{2,3,4}$ получены из $\mathbf{M}_1$ путем выполнения циклической перестановки кубитов. (Полученный с помощью циклической перестановки кубитов пятый оператор $\mathbf{M}_5 = \mathbf{ZZX1X} = \mathbf{M}_1\mathbf{M}_2\mathbf{M}_3\mathbf{M}_4$ зависит от четырех

других.) Поскольку результатом циклической перестановки сомножителей генератора является другой генератор, код сам по себе цикличен — результатом циклической перестановки кубитов кодового слова является кодовое слово.

Очевидно, что каждый $\mathbf{M}_i$ не содержит ни одного $\mathbf{Y}$ и, следовательно, при возведении в квадрат дает $\mathbf{1}$. Для каждой пары генераторов имеет место по два столкновения между $\mathbf{X}$ и $\mathbf{Z}$, так что генераторы коммутируют. Можно быстро проверить, что каждый оператор Паули с весом единица или два антикоммутирует по крайней мере с одним генератором, так что расстояние кода равно трем.

Рассмотрим, например, существуют ли коммутирующие со всеми четырьмя генераторами операторы ошибок с носителями на первой паре кубитов. Чтобы коммутировать с $\mathbf{1X}$ в $\mathbf{M}_2$ и с $\mathbf{X1}$ в $\mathbf{M}_3$, оператор с весом два должен быть равен $\mathbf{XX}$. Но $\mathbf{XX}$ антикоммутирует с $\mathbf{XZ}$ в $\mathbf{M}_1$ и с $\mathbf{ZX}$ в $\mathbf{M}_4$. В симплектической записи стабилизатор (7.148) имеет вид

$$\tilde{H} = \left(\begin{array}{c|c} 01100 & 10010 \\ 00110 & 01001 \\ 00011 & 10100 \\ 10001 & 01010 \end{array}\right). \tag{7.149}$$

Эта матрица имеет изящную интерпретацию, так как каждый из ее столбцов можно рассматривать как синдром однокубитовой ошибки. Например, оператор однокубитового инвертирования бита $\mathbf{X}_j$ коммутирует с $\mathbf{M}_i$, если в позиции j оператор $\mathbf{M}_i$ имеет $\mathbf{1}$ или $\mathbf{X}$, и антикоммутирует, если в позиции j оператор $\mathbf{M}_i$ имеет $\mathbf{Z}$. Таким образом, таблица

	$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$
$\mathbf{M}_1$	0	1	1	0	0
$\mathbf{M}_2$	0	0	1	1	0
$\mathbf{M}_3$	0	0	0	1	1
$\mathbf{M}_4$	1	0	0	0	1

составляет список результатов измерения $\mathbf{M}_{1,2,3,4}$ в случае инвертирования бита. (Например, если инвертирован первый бит, результаты измерения $\mathbf{M}_1 = \mathbf{M}_2 = \mathbf{M}_3 = 1$, $\mathbf{M}_4 = -1$ выявляют ошибку). Аналогично, правую часть $\tilde{H}$ можно рассматривать как таблицу синдромов фазовых ошибок.

	$\mathbf{Z}_1$	$\mathbf{Z}_2$	$\mathbf{Z}_3$	$\mathbf{Z}_4$	$\mathbf{Z}_5$
$\mathbf{M}_1$	1	0	0	1	0
$\mathbf{M}_2$	0	1	0	0	1
$\mathbf{M}_3$	1	0	1	0	0
$\mathbf{M}_4$	0	1	0	1	0

Так как $\mathbf{Y}$ антикоммутирует с $\mathbf{Z}$ и $\mathbf{X}$, синдром ошибки $\mathbf{Y}_i$ дает сумма i-х столбцов таблиц $\mathbf{X}$ и $\mathbf{Z}$:

	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	$\mathbf{Y}_4$	$\mathbf{Y}_5$
$\mathbf{M}_1$	1	1	1	1	0
$\mathbf{M}_2$	0	1	1	1	1
$\mathbf{M}_3$	1	0	1	1	1
$\mathbf{M}_4$	1	1	0	1	1

Путем непосредственной проверки можно убедиться в том, что все 15 столбцов таблиц синдромов $\mathbf{X}$, $\mathbf{Y}$ и $\mathbf{Z}$ различны, и, следовательно, мы вновь подтверждаем, что рассматриваемый код невырожден и корректирует одну ошибку. Действительно, код совершенен — каждая из пятнадцати нетривиальных двоичных строк длины четыре выступает в качестве столбца в одной из этих таблиц.

Благодаря свойству цикличности кода, нетрудно охарактеризовать все 15 нетривиальных элементов его стабилизатора. Помимо $\mathbf{M}_1 = \mathbf{XZZX1}$ и четырех операторов, получаемых из него путем циклических перестановок кубитов, стабилизатор включает

$$\mathbf{M}_3\mathbf{M}_4 = -\mathbf{YXXY1} \tag{7.150}$$

плюс все его циклические перестановки, а также

$$\mathbf{M}_2\mathbf{M}_5 = -\mathbf{ZYYZ1} \tag{7.151}$$

и все его циклические перестановки. Очевидно, что все элементы стабилизатора являются операторами Паули с весом четыре.

В качестве логических операторов можно выбрать

$$\bar{\mathbf{Z}} = \mathbf{ZZZZZ}, \qquad \bar{\mathbf{X}} = \mathbf{XXXXX}; \tag{7.152}$$

они коммутируют с $\mathbf{M}_{1,2,3,4}$, дают в квадрате единицу $\mathbf{1}$ и антикоммутируют между собой. Имея вес пять, они сами не содержатся в стабилизаторе. Следовательно, если нас не беспокоит разрушение закодированного состояния, то мы можем определить значение $\bar{\mathbf{Z}}$ для закодированного кубита, измеряя $\mathbf{Z}$ каждого кубита и вычисляя четность результатов. Фактически, поскольку код имеет расстояние три, существуют элементы множества $S^{\perp} \setminus S$ с весом три; альтернативные выражения для $\bar{\mathbf{Z}}$ и $\bar{\mathbf{X}}$ можно получить путем умножения на элементы стабилизатора. Например, мы можем

выбрать

$$\bar{\mathbf{Z}} = (\mathbf{ZZZZZ}) \cdot (-\mathbf{ZYYZ1}) = -\mathbf{1XX1Z} \tag{7.153}$$

(или одну из его циклических перестановок) и

$$\bar{\mathbf{X}} = (\mathbf{XXXXX}) \cdot (-\mathbf{YXXY1}) = -\mathbf{Z11ZX} \tag{7.154}$$

(или одну из его циклических перестановок). Следовательно, возможно установить значение $\bar{\mathbf{X}}$ или $\bar{\mathbf{Z}}$, измеряя $\mathbf{X}$ или $\mathbf{Z}$ только трех из пяти кубитов в блоке и вычисляя четности результатов.

Если угодно, ортонормированный базис кодового подпространства можно построить следующим образом. Начиная с любого состояния $|\psi_0\rangle$, можно получить

$$|\Psi_0\rangle = \sum_{\mathbf{M} \in S} \mathbf{M}|\psi_0\rangle. \tag{7.155}$$

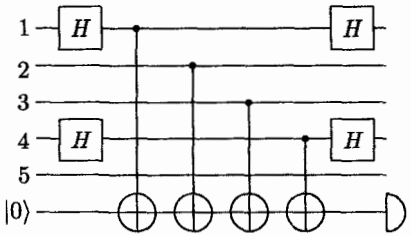
Это (ненормированное) состояние удовлетворяет условию $\mathbf{M}'|\Psi_0\rangle = |\Psi_0\rangle$ для каждого $\mathbf{M}' \in S$, так как умножение на элемент стабилизатора лишь переставляет слагаемые в сумме. Чтобы получить закодированное состояние $|\bar{0}\rangle$, отвечающее собственному значению $\bar{\mathbf{Z}} = 1$, можно начать с состояния $|00000\rangle$, которое также является собственным состоянием с $\bar{\mathbf{Z}} = 1$, но не принадлежит стабилизатору; в итоге находим (с точностью до нормировки)

$$\begin{aligned}
|\bar{0}\rangle = \sum_{\mathbf{M} \in S} \mathbf{M}|00000\rangle = \\
= |00000\rangle + (\mathbf{M}_1 + \text{циклические перестановки})|00000\rangle + \\
+ (\mathbf{M}_3\mathbf{M}_4 + \text{циклические перестановки})|00000\rangle + \\
+ (\mathbf{M}_2\mathbf{M}_5 + \text{циклические перестановки})|00000\rangle = \\
= |00000\rangle + (|10010\rangle + \text{циклические перестановки}) - \\
- (|11110\rangle + \text{циклические перестановки}) - \\
- (|01100\rangle + \text{циклические перестановки}).
\end{aligned} \tag{7.156}$$

После этого, применяя $\bar{\mathbf{X}}$ к $|\bar{0}\rangle$, то есть инвертируя все пять кубитов, можно найти

$$\begin{aligned}
|\bar{1}\rangle = \bar{\mathbf{X}}|\bar{0}\rangle = |11111\rangle + (|01101\rangle + \text{циклические перестановки}) - \\
- (|00001\rangle + \text{циклические перестановки}) - \\
- (|10011\rangle + \text{циклические перестановки}).
\end{aligned} \tag{7.157}$$

Как измеряется синдром? Возможная для выполнения измерения $\mathbf{M}_1 = \mathbf{XZZX1}$ схема изображена на рисунке.



Повороты Адамара первого и четвертого кубитов преобразуют $\mathbf{M}_1$ в тензорное произведение $\mathbf{ZZZZ1}$, а затем вентили CNOT отпечатывают значение этого оператора на служебный кубит. Заключительные повороты Адамара возвращают закодированный блок в стандартное кодовое подпространство. Схемы для измерения $\mathbf{M}_{2,3,4}$ получаются из изображенного выше путем циклической перестановки пяти кубитов в кодовом блоке.

А что можно сказать о кодировании? Мы хотим построить унитарное преобразование

$$\mathbf{U}_{\text{encode}} : |0000\rangle \otimes (a|0\rangle + b|1\rangle) \rightarrow a|\bar{0}\rangle + b|\bar{1}\rangle. \tag{7.158}$$

Мы только что видели, что $|00000\rangle$ является собственным состоянием оператора $\bar{\mathbf{Z}}$, отвечающим собственному значению $\bar{\mathbf{Z}} = 1$, а $|00001\rangle$ — собственным состоянием, отвечающим собственному значению $\bar{\mathbf{Z}} = -1$. Следовательно, (с точностью до нормировки)

$$a|\bar{0}\rangle + b|\bar{1}\rangle = \sum_{\mathbf{M}\in S} \mathbf{M}|0000\rangle \otimes (a|0\rangle + b|1\rangle). \tag{7.159}$$

Итак, нам необходимо понять, как построить схему, применяющую операцию $\sum \mathbf{M}$ к начальному состоянию.

Так как генераторы независимы, каждый элемент стабилизатора можно единственным способом представить в виде произведения генераторов и, следовательно, записать

$$\sum_{\mathbf{M}\in S} \mathbf{M} = (\mathbf{1} + \mathbf{M}_4)(\mathbf{1} + \mathbf{M}_3)(\mathbf{1} + \mathbf{M}_2)(\mathbf{1} + \mathbf{M}_1). \tag{7.160}$$

Теперь, чтобы двигаться дальше, удобно представить стабилизатор в альтернативной форме. Отметим, что, не изменяя стабилизатор, генератор $\mathbf{M}_i$ можно заменить на $\mathbf{M}_i\mathbf{M}_j$. Эта замена эквивалентна добавлению j-й строки к i-й строке матрицы $\tilde{H}$. Используя подобные операции со строками, можно выполнить процедуру Гаусса в матрице H_X размерности 4×5 и,

таким образом, получить новое представление для стабилизатора

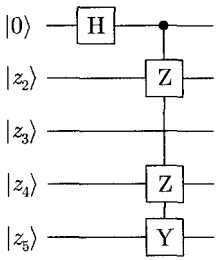
$$\tilde{H} = \left(\begin{array}{c|c} 11011 & 10001 \\ 00110 & 01001 \\ 11000 & 00101 \\ 10111 & 00011 \end{array}\right) \tag{7.161}$$

или

$$\begin{aligned} \mathbf{M}_1 &= -\mathbf{YZ1ZY}, \\ \mathbf{M}_2 &= +\mathbf{1XZZX}, \\ \mathbf{M}_3 &= +\mathbf{ZZX1X}, \\ \mathbf{M}_4 &= -\mathbf{Z1ZYY}. \end{aligned} \tag{7.162}$$

В таком виде $\mathbf{M}_i$ применяет $\mathbf{X}$ (инвертирование) только к i-у и пятому кубиту в блоке.[1]

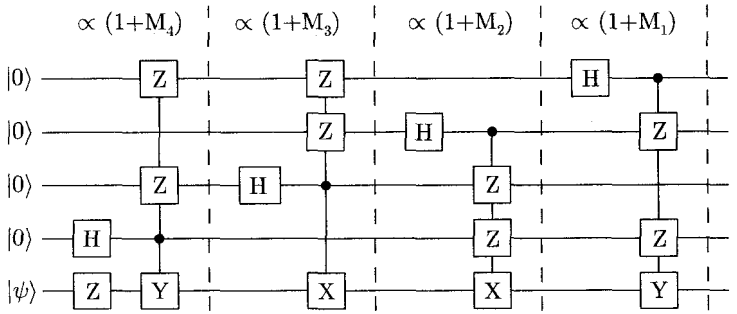
Выбирая стабилизатор в таком виде, мы можем применить $\frac{1}{\sqrt{2}}(\mathbf{1}+\mathbf{M}_1)$ к состоянию $|0, z_2, z_3, z_4, z_5\rangle$, выполняя схему



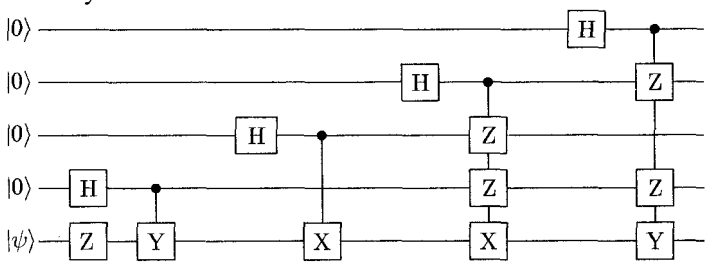
Преобразование Адамара готовит состояние $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Если первый кубит находится в состоянии $|0\rangle$, то другие операции ничего не делают, то есть результатом действия этой схемы является тождественное преобразование **11111**. Но если после адамаровского поворота первый кубит оказывается в состоянии $|1\rangle$, то остальные вентили этой схемы применяют **1Z1ZY**. В этом случае результатом является применение операции $\mathbf{M}_1 = -\mathbf{YZ1ZY}$. Можно построить подобные схемы, применяющие $\frac{1}{\sqrt{2}}(\mathbf{1}+\mathbf{M}_2)$ к состоянию $|z_1, 0, z_3, z_4, z_5\rangle$, и так далее. Кроме вентилей Адамара, каждая из этих схем подвергает действию операций $\mathbf{Z}$ и контролируемого $\mathbf{Z}$ от одного до четырех кубитов; эти кубиты никогда не инвертируют-

[1] В оригинале были пропущены знаки минус в выражениях для генераторов $\mathbf{M}_1$ и $\mathbf{M}_4$. Именно учет *правильных* знаков генераторов (7.162) приводит к *правильным* относительным знакам в выражениях для базисных векторов (7.156), (7.157). Это исправление учтено в следующих ниже рисунках и пояснениях к ним в тексте данного раздела. — *Прим. ред.*

ся. (Это убеждает нас в том, что мы выполнили процедуру Гаусса в матрице H_X.) Таким образом, можно сконструировать следующую кодирующую схему:



Более того, каждый действующий на $|0\rangle$ вентиль $\mathbf{Z}$ можно заменить на единичный, следовательно, эту схему можно упростить, исключив все такие вентили и получив



Эту процедуру можно обобщить для построения кодирующих схем для любых стабилизирующих кодов.[1]

Поскольку кодирующее преобразование унитарно, для декодирования можно использовать его сопряжение. А так как квадрат каждого вентиля равен $\pm\mathbf{1}$, то декодирующая схема представляет собой ту же самую кодирующую схему, лишь работающую в обратном направлении.

7.11. Распределение квантового секрета

Код $[[5, 1, 3]]$ является прекрасной иллюстрацией возможного применения корректирующих ошибки квантовых кодов.[2]

[1] *Естественно, что вид кодирующей схемы зависит от выбора генераторов стабилизатора. Альтернативную кодирующую схему для 5-кубитового кода (а также для других известных кодов) можно найти на сайте* http://iaks-www.ira.uka.de/home/grassl/QECC/ — *Прим. ред.*

[2] R. Cleve, D. Gottesman, and H.-K. Lo, *How to Share a Quantum Secret*, Phys. Rev. Lett. **83**, 648–651 (1999); quant-ph/9901025.

Предположим, что некоторую совершенно секретную информацию необходимо доверить n партнерам. Так как никому из них нельзя доверять полностью, секрет делится на n частей, так что каждый партнер имеет доступ только к своей части и не может узнать о секрете в целом. Но если достаточное количество партнеров соберутся и объединят свои части, то они смогут расшифровать секрет или какую-то его часть.

В частности, пороговая (m, n)-схема имеет такое свойство, что для реконструкции всей секретной информации достаточно m частей. Но из $m-1$ частей невозможно извлечь никакой информации. (Это называется пороговой схемой, потому что при собранных воедино $1, 2, 3, \ldots, m-1$ частях узнать ничего нельзя, но следующая часть позволяет переступить порог и раскрыть всю информацию.)

Следует различать два вида секретов: классический секрет представляет собой *a priori* неизвестную строку битов, в то время как квантовым секретом является *a priori* неизвестное квантовое состояние. Секреты каждого типа можно поделить. В частности, мы можем распределить классический секрет между несколькими партнерами, выбрав одно из ансамбля взаимно ортогональных (запутанных) квантовых состояний и распределив это состояние между партнерами.

Например, нетрудно видеть, что код $[[5, 1, 3]]$ может использоваться в пороговой $(3, 5)$-схеме, где разделенная информация является классической. Один классический бит кодируется одним из двух ортогональных состояний $|\bar{0}\rangle$ или $|\bar{1}\rangle$, а затем пять кубитов распределяются между пятью партнерами. Как мы уже видели, если объединятся любые два партнера, то матрицей плотности $\boldsymbol{\rho}$ их двух кубитов будет

$$\boldsymbol{\rho}^{(2)} = \frac{1}{4}\mathbf{1} \tag{7.163}$$

(поскольку код невырожден). Следовательно, из любого измерения их двух кубитов они ничего не узнают о квантовом состоянии. Но мы также видели, что код $[[5, 1, 3]]$ может исправить две локализованных ошибки или два стирания. Когда объединятся любые три участника, они могут исправить две ошибки (или восстановить два недостающих кубита) и полностью восстановить закодированное состояние $|\bar{0}\rangle$ или $|\bar{1}\rangle$.

Ясно, что с помощью аналогичной процедуры можно поделить один кубит квантовой информации — код $[[5, 1, 3]]$ также является основой квантовой пороговой ((3,5))-схемы (мы используем обозначение $((m, n))$, если поделена квантовая информация, и (m, n), если поделена классическая информация). Как эту схему деления квантового секрета распространить на большее количество кубитов? Допустим, мы приготовили чистое n-кубито-

вое состояние $|\psi\rangle$. Может ли оно быть использовано в пороговой $((m,n))$-схеме?

Нам известно, что для реконструкции состояния должно быть достаточно m кубитов; следовательно, можно восстановить $n-m$ удалений. Из общего критерия коррекции ошибок следует, что математическое ожидание любой наблюдаемой с весом, не превышающим $n-m$, не должно зависеть от состояния $|\psi\rangle$

$$\langle\psi|\mathbf{E}|\psi\rangle \quad \text{не зависит от} \quad |\psi\rangle, \quad \text{если} \quad \mathrm{wt}(\mathbf{E}) \leqslant n-m. \tag{7.164}$$

Таким образом, если m партнеров имеют всю информацию, то другие $n-m$ партнеров не имеют никакой информации. Это справедливо, поскольку квантовую информацию нельзя клонировать.

С другой стороны, мы знаем, что $m-1$ частей ничего не откроют, или что

$$\langle\psi|\mathbf{E}|\psi\rangle \quad \text{не зависит от} \quad |\psi\rangle, \quad \text{если} \quad \mathrm{wt}(\mathbf{E}) \leqslant m-1. \tag{7.165}$$

Отсюда следует, что можно восстановить $m-1$ стирание, или что другие $n-m+1$ партнеров располагают всей информацией.

Из этих двух высказываний мы получаем два неравенства

$$\begin{aligned} n-m < m &\Rightarrow n < 2m, \\ m-1 < n-m+1 &\Rightarrow n > 2m-2. \end{aligned} \tag{7.166}$$

Отсюда следует, что в квантовой пороговой $((m,n))$-схеме чистого состояния, в которой каждый партнер имеет один кубит,

$$n = 2m-1. \tag{7.167}$$

Другими словами, порог будет достигнут, когда количество наличных кубитов превысит половину всех n кубитов.

Таким образом, если каждая часть представляет собой кубит, то квантовая пороговая схема чистого состояния представляет собой квантовый код $[[2m-1,k,m]]$ с $k \geqslant 1$. Но в действительности коды $[[3,1,2]]$ и $[[7,1,4]]$ не существуют, а из границы Рейнса следует, что не существуют коды с $m > 3$. Следовательно, код $[[5,1,3]]$ представляет собой единственную квантовую пороговую схему.

Здесь следует сделать несколько оговорок. Во-первых, ограничение $n = 2m-1$ остается справедливым, даже если каждая часть является q-мерной системой, а не кубитом. Но в случае $q > 2$ можно построить различные коды:

$$[[2m-1,1,k]]_q \tag{7.168}$$

(см., например, упражнения).

Во-вторых, поделенная информация может представлять собой смешанное состояние (в котором закодировано чистое состояние). Например, если мы отбрасываем один кубит из 5-кубитового блока, мы получаем $((3,4))$-схему. Вновь, как только у нас появляется три кубита, мы сможем восстановить два стертых (то есть недостающих), один из которых находится в руках другого партнера, а второй — только что был нами же выброшен.

Наконец, мы предположили, что поделенная информация является квантовой. Но если вместо этого мы делим только классическую информацию, тогда условия восстановления стертых кубитов становятся менее строгими. Например, пару Белла можно рассматривать как вид пороговой $(2,2)$-схемы для двух битов классической информации, которая закодирована выбором одного из четырех взаимно ортогональных состояний $|\phi^{\pm}\rangle$, $|\psi^{\pm}\rangle$. Располагающий одним из двух кубитов партнер не может получить доступ к этой классической информации. Но эта схема не подходит для распределения квантового секрета, поскольку линейные комбинации этих состояний Белла *не обладают* тем свойством, что $\rho = \frac{1}{2}1$ после вычисления следа по состояниям одного из двух кубитов.

7.12. Некоторые другие стабилизирующие коды

7.12.1. Код $[[6,0,4]]$

При $k = 0$ квантовый код имеет одномерное кодовое подпространство, то есть существует только одно закодированное состояние. Код нельзя использовать для хранения неизвестной квантовой информации; тем не менее, коды с $k = 0$ могут обладать интересными свойствами. Так как они могут обнаруживать и диагностировать ошибки, они могут быть полезными для изучения корреляций в декогерентизации, вызванной взаимодействием с окружением.

Если $k = 0$, то S и $S^{\perp}$ совпадают — оператор Паули, коммутирующий со всеми элементами стабилизатора, должен принадлежать этому стабилизатору. В этом случае расстояние d определяется как минимальный вес любого принадлежащего стабилизатору оператора Паули. Таким образом, код с расстоянием d может «обнаружить $d-1$ ошибок»; то есть если любой оператор Паули с весом, меньшим d, действует на кодовое состояние, то результат будет ортогонален этому состоянию.

Закодированное состояние кода $[[6,0,4]]$ (ассоциированного с кодом $[[5,1,3]]$) можно представить в виде

$$|0\rangle \otimes |\bar{0}\rangle + |1\rangle \otimes |\bar{1}\rangle, \tag{7.169}$$

где $|\bar{0}\rangle$ и $|\bar{1}\rangle$ представляют собой собственные состояния оператора $\bar{\mathbf{Z}}$ кода $[[5,1,3]]$. Вы можете проверить, что этот код имеет расстояние $d=4$ (см. упражнение 7.3).

Код $[[6,0,4]]$ интересен тем, что его кодовое состояние максимально запутано. Мы можем выбрать любые три кубита из шести. Матрица плотности $\boldsymbol{\rho}^{(3)}$ этих трех кубитов, полученная путем вычисления следа по состояниям трех остальных, совершенно случайна, $\boldsymbol{\rho}^{(3)}=\frac{1}{8}\mathbf{1}$. В этом смысле, кодовое состояние $[[6,0,4]]$ является естественным многочастичным аналогом двухкубитовых состояний Белла. Оно «гораздо сильнее запутано», нежели шестикубитовое кот-состояние $\frac{1}{\sqrt{2}}(|000000\rangle+|111111\rangle)$. Если мы измерим в базисе $\{|0\rangle,|1\rangle\}$ любой один из шести кубитов кот-состояния, мы узнаем все о приготовленном состоянии оставшихся пяти кубитов. Но мы можем измерить по своему усмотрению любую наблюдаемую, действующую на любые *три* кубита в состоянии $[[6,0,4]]$, и ничего не узнаем относительно состояния оставшихся трех кубитов, которое по-прежнему описывается матрицей плотности $\boldsymbol{\rho}^{(3)}=\frac{1}{8}\mathbf{1}$.

Код $[[6,0,4]]$ тем более интересен, что, оказывается (но не так просто доказывается), не существует его обобщения на большее количество кубитов, то есть не существует $[[2n,0,n+1]]$ двоичных квантовых кодов для $n>3$. Однако в упражнениях вы увидите, что существуют другие, недвоичные, максимально запутанные состояния, которые можно построить.

7.12.2. Детектирующие ошибки $[[2m,2m-2,2]]$-коды

Состояние Белла $|\phi^+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)$ представляет собой код $[[2,0,2]]$ с генераторами стабилизатора

$$\mathbf{ZZ}, \qquad \mathbf{XX}. \tag{7.170}$$

Этот код имеет расстояние два, поскольку не существует операторов Паули с единичным весом, коммутирующих с обоими генераторами (ни один из $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ одновременно не коммутирует с $\mathbf{X}$ и $\mathbf{Z}$). Соответственно, инвертирование бита ($\mathbf{X}$), обращение фазы ($\mathbf{Z}$) или обе эти ошибки ($\mathbf{Y}$), действующие на любой кубит в $|\phi^+\rangle$, преобразуют его в ортогональное состояние (одно из состояний Белла $|\phi^-\rangle$, $|\psi^+\rangle$, $|\psi^-\rangle$).

Единственный способ обобщить состояния Белла на большее количество кубитов — рассмотреть код $n=4$, $k=2$ с генераторами стабилизатора

$$\mathbf{ZZZZ}, \qquad \mathbf{XXXX}. \tag{7.171}$$

Это код с расстоянием $d = 2$ по той же причине, что и предыдущий. Кодовое подпространство натянуто на четные состояния (**ZZZZ**), инвариантные относительно одновременного инвертирования всех четырех кубитов (**XXXX**). Базис представляет собой

$$\begin{aligned} &|0000\rangle + |1111\rangle, \\ &|0011\rangle + |1100\rangle, \\ &|0101\rangle + |1010\rangle, \\ &|0110\rangle + |1001\rangle. \end{aligned} \tag{7.172}$$

Очевидно, что действующая на любой кубит ошибка **X** или **Z** преобразует каждое из этих состояний в ортогональное кодовому подпространству состояние; таким образом, можно обнаружить любую однокубитовую ошибку.

Дальнейшим обобщением является код $[[2m, 2m-2, 2]]$ с генераторами стабилизатора

$$\mathbf{ZZ}\ldots\mathbf{Z}, \quad \mathbf{XX}\ldots\mathbf{X} \tag{7.173}$$

(длина должна быть четной, чтобы генераторы коммутировали между собой). Кодовое подпространство натянуто на 2^{n-2} хорошо знакомых нам кот-состояний

$$\frac{1}{\sqrt{2}}(|x\rangle + |\neg x\rangle), \tag{7.174}$$

где x — строка длины $n = 2m$ и четного веса.

7.12.3. Код [[8, 3, 3]]

Как уже отмечалось при обсуждении кода $[[5, 1, 3]]$, стабилизирующий код с генераторами

$$\tilde{H} = (H_Z|H_X) \tag{7.175}$$

может исправить одну ошибку, если: (1) столбцы матрицы $\tilde{H}$ различны (свой синдром для каждой ошибки **X** и **Z**); (2) любая сумма столбца матрицы H_Z с соответствующим столбцом матрицы H_X отличается от любого столбца $\tilde{H}$ и от всех других таких сумм (каждую ошибку **Y** можно отличить от всех остальных однокубитовых ошибок).

Нетрудно построить матрицу H размерности 5×16 с этими свойствами и, таким образом, получить стабилизатор кода $[[8, 3, 3]]$; выберем

$$\tilde{H} = \left(\begin{array}{c|c} H & H^\sigma \\ 11111111 & 00000000 \\ 00000000 & 11111111 \end{array}\right). \tag{7.176}$$

Здесь H представляет собой матрицу размерности 3×8

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \tag{7.177}$$

столбцами которой являются все возможные различные двоичные последовательности длины три, а H^σ получается из H с помощью соответствующей перестановки столбцов. Эта перестановка выбирается таким образом, чтобы были различны все восемь сумм столбцов матрицы H с соответствующими столбцами H^σ. С помощью непосредственной проверки можно убедиться, что подходящим выбором служит

$$H^\sigma = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}, \tag{7.178}$$

тогда суммы столбцов равны

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{7.179}$$

Две последних строки матрицы H служат для того, чтобы отличать каждый синдром $\mathbf{X}$ от каждого синдрома $\mathbf{Y}$ или $\mathbf{Z}$, а вышеупомянутое свойство матрицы H^σ гарантирует, что все синдромы $\mathbf{Y}$ различны. Следовательно, мы построили код длины восемь с $k = 8 - 5 = 3$, который может исправить одну ошибку. Собственно, это простейший в бесконечном классе кодов $[[2^m, 2^m - m - 2, 3]]$ с $m \geqslant 3$, построенных Готтесманом.

Квантовый код $[[8, 3, 3]]$, который мы только что описали, является, так сказать, кузеном «расширенного кода Хэмминга», самодуального классического кода $[8, 4, 4]$, полученного из дуального коду Хэмминга $[7, 3, 4]$-кода путем добавления дополнительного бита четности. Его матрица контроля четности (которая также является его генерирующей матрицей) имеет вид

$$H_{\text{EH}} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{7.180}$$

Эта матрица H_{EH} обладает тем свойством, что различны не только все ее восемь столбцов, но от них также отличается и *сумма* любых двух столбцов; действительно, четвертым битом суммы двух столбцов является нуль, а не единица.

7.13. Коды над GF(4)

Мы построили код $[[5,1,3]]$, угадав генераторы стабилизатора и проверив, что $d=3$. Существует ли более систематический метод?

На самом деле да. Наше подозрение, что код $[[5,1,3]]$ может существовать, возникло из наблюдения, что его параметры насыщают квантовое неравенство упаковки сфер для кодов с $t=1$

$$1+3n=2^{n-k}, \tag{7.181}$$

($16=16$ при $n=5$ и $k=1$). Для теоретика в области кодирования это уравнение может показаться знакомым.

Помимо двоичных кодов, на которых до сих пор мы концентрировали внимание, классические коды можно также построить из строк (длины n) символов, принимающих значения не в $\{0,1\}$, а в конечном поле $\mathrm{GF}(q)$, содержащем q элементов. Такие конечные поля существуют для любого $q=p^m$, где p — простое число. (GF представляет собой аббревиатуру для «Galois Field» — поле Галуа, названное так в честь его первооткрывателя.)

Для таких недвоичных кодов можно смоделировать ошибку как добавление элемента поля, циклический сдвиг q символов. Тогда всего будет $q-1$ нетривиальных ошибок. Весом вектора в $\mathrm{GF}(q)^n$ является количество его ненулевых элементов, а расстояние между двумя векторами представляет собой вес их разности (количество несовпадающих элементов). Классический код $[n,k,d]_q$ состоит из q^k кодовых слов в $\mathrm{GF}(q)^n$, где минимальное расстояние между парами строк равно d. Граница упаковки сфер, которая должна быть насыщена, для того чтобы мог существовать код $[n,k,d]_q$, при $d=3$ имеет вид

$$1+(q+1)n \leqslant q^{n-k}. \tag{7.182}$$

Насыщающие эту границу при $q=2$ совершенные двоичные коды Хэмминга с параметрами

$$n=2^m-1, \quad k=n-m \tag{7.183}$$

допускают обобщение на любое $\mathrm{GF}(q)$; совершенные коды Хэмминга над $\mathrm{GF}(q)$ можно построить при

$$n=\frac{q^m-1}{q-1}, \quad k=n-m. \tag{7.184}$$

Квантовый код $[[5,1,3]]$ происходит от классического кода Хэмминга $[5,3,3]_4$ (случай $q=4$ и $m=2$).

Что общего между классическими кодами над GF(4) и двоичными квантовыми стабилизирующими кодами? Родство возникает, потому что стабилизатору можно сопоставить замкнутое относительно сложения множество векторов над GF(4).

Поле GF(4) имеет четыре элемента, которые можно обозначить как $0, 1, \omega, \bar{\omega}$, где

$$1+1=\omega+\omega=\bar{\omega}+\bar{\omega}=0,$$
$$1+\omega=\bar{\omega} \tag{7.185}$$

и $\omega^2=\bar{\omega}$, $\omega\bar{\omega}=1$. Следовательно, аддитивная структура GF(4) соответствует мультипликативной структуре операторов группы Паули $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$. Действительно, двоичная строка $(\alpha|\beta)$ длины $2n$, которую мы использовали для обозначения элемента группы Паули, может эквивалентно рассматриваться как вектор длины n в $\mathrm{GF}(4)^n$

$$(\alpha|\beta) \leftrightarrow \alpha+\beta\omega. \tag{7.186}$$

Стабилизатор с 2^{n-k} элементами можно рассматривать как субкод кода GF(4), замкнутый относительно сложения и содержащий 2^{n-k} кодовых слов.

Отметим, что код не должен быть векторным пространством над GF(4), поскольку от него не требуется замкнутость относительно умножения на скаляр принадлежащий GF(4). В частном случае, когда код является векторным пространством, он называется *линейным* кодом.

О кодах над GF(4) известно много, поэтому эта связь открыла возможность теоретикам в области (классического) кодирования построить множество квантовых кодов коррекции ошибок.[1] Однако не каждый субкод $\mathrm{GF}(4)^n$ связан с квантовым кодом; мы до сих пор не выдвинули требование, чтобы стабилизатор был абелев — векторы $(\alpha|\beta)$, образующие линейную оболочку кода, должны быть взаимно ортогональны относительно симплектического внутреннего произведения

$$\alpha\cdot\beta'+\alpha'\cdot\beta. \tag{7.187}$$

Это условие ортогональности может выглядеть странным для теоретика в области кодирования, которому более привычно определение внутреннего произведения двух векторов в $\mathrm{GF}(4)^n$ как элемента поля GF(4), заданного соотношением

$$v*u=\bar{v}_1u_1+\ldots+\bar{v}_nu_n, \tag{7.188}$$

[1] A. R. Calderbank, E. M. Rains, P. M. Shor, and N. J. A. Sloane, *Quantum error correction via codes over* GF(4), IEEE Transact. on Inform. Theor., **44**, 1369 (1998); quant-ph/9608006.

где сопряжение, обозначенное черточкой, меняет местами ω и $\bar{\omega}$. Если это «эрмитово» внутреннее $*$-произведение двух векторов v и u равно

$$v * u = a + b\omega \in \mathrm{GF}(4), \tag{7.189}$$

то наше симплектическое внутреннее произведение равно

$$v \cdot u = b. \tag{7.190}$$

Следовательно, обращение в нуль симплектического внутреннего произведения является более слабым условием, чем обращение в нуль эрмитова внутреннего произведения. В самом деле, в частном случае *линейного* кода самоортогональность относительно эрмитова внутреннего произведения фактически эквивалентна самоортогональности относительно симплектического внутреннего произведения. Отметим, что если $v * u = a + b\omega$, то ортогональность относительно симплектического внутреннего произведения требует $b = 0$. Но если u принадлежит линейному коду, то и $\bar{\omega}u$ тоже, где

$$v * (\bar{\omega}u) = b + a\bar{\omega}, \tag{7.191}$$

так что

$$v \cdot (\bar{\omega}u) = a. \tag{7.192}$$

Мы видим, что если v и u принадлежат линейному $\mathrm{GF}(4)$-коду и ортогональны относительно симплектического внутреннего произведения, то они также ортогональны относительно эрмитова внутреннего произведения. Тогда мы делаем вывод, что линейный $\mathrm{GF}(4)$-код определяет квантовый стабилизирующий код, если и только если этот код самоортогонален относительно эрмитова внутреннего произведения. Классические коды с такими свойствами очень хорошо изучены.[1].

В частности, рассмотрим снова код Хэмминга $[5, 3, 3]_4$. Его матрицу контроля четности (в нетрадиционном представлении) можно представить в виде

$$H = \begin{pmatrix} 1 & \omega & \omega & 1 & 0 \\ 0 & 1 & \omega & \omega & 1 \end{pmatrix}, \tag{7.193}$$

что также является генерирующей матрицей дуального ему, линейного самоортогонального кода $[5, 2, 4]_4$. По сути, этот $[5, 2, 4]_4$ код с $4^2 = 16$

[1] См., например, E. J. MacWilliams, N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland Publishing Company, Amsterdam, New York, Oxford (1977); [перевод: Ф. Дж. МакВильямс, Н. Дж. А. Слоэн, *Теория кодов, исправляющих ошибки*. — М.: Связь, 1979. — *Прим. ред.*]

кодовыми словами в точности является стабилизатором квантового кода $[[5,1,3]]$. Отождествляя $1 \equiv \mathbf{X}$, $\omega \equiv \mathbf{Z}$, мы принимаем две строки матрицы H в качестве генераторов стабилизатора $\mathbf{M}_1$, $\mathbf{M}_2$. Код, дуальный коду Хэмминга, является линейным, поэтому линейные комбинации строк принадлежат коду. Складывая строки и умножая результат на ω, получаем

$$\omega(1,\bar{\omega},0,\bar{\omega},1) = (\omega,1,0,1,\omega), \tag{7.194}$$

что представляет собой $\mathbf{M}_4$. А если к $\mathbf{M}_4$ мы добавим $\mathbf{M}_2$ и умножим на $\bar{\omega}$, то найдем

$$\bar{\omega}(\omega,0,\omega,\bar{\omega},\bar{\omega}) = (1,0,1,\omega,\omega), \tag{7.195}$$

что представляет собой $\mathbf{M}_3$.

Код $[[5,1,3]]$ является одним из примеров достаточно общей конструкции. Рассмотрим субкод C на $\mathrm{GF}(4)^n$, который является аддитивным (замкнутым относительно сложения) и самоортогональным (содержащимся в дуальном ему коде) относительно симплектического внутреннего произведения. Этот $\mathrm{GF}(4)$-код может быть отождествлен со стабилизатором двоичного ККК0 длины n. Если $\mathrm{GF}(4)$-код содержит 2^{n-k} кодовых слов, то ККК0 имеет k закодированных кубитов. Расстоянием d ККК0 является минимальный вес вектора из $C^{\perp} \setminus C$.

Другим примером самоортогонального линейного $\mathrm{GF}(4)$-кода является дуальный коду Хэмминга $m = 3$ с

$$n = \frac{1}{3}(4^3 - 1) = 21. \tag{7.196}$$

Код Хэмминга имеет 4^{n-m} кодовых слов, а дуальный ему код — $4^m = 2^6$ кодовых слов. Мы непосредственно получаем ККК0 с параметрами

$$[[21,15,3]], \tag{7.197}$$

который может исправить одну ошибку.

7.14. Хорошие квантовые коды

Семейство кодов $[[n,k,d]]$ является *хорошим*, если оно содержит коды, чья «скорость» $R = k/n$ и «вероятность возникновения ошибки» $p = t/n$ (где $t = (d-1)/2$) стремятся к ненулевым пределам при $n \to \infty$. Мы можем использовать формализм стабилизатора, чтобы доказать «квантовую границу Гилберта–Варшамова», которая демонстрирует существование хороших квантовых кодов. В сущности, хорошие коды можно выбрать невырожденными.

Мы дадим только набросок доказательства, не выполняя точно требуемые вычисления. Пусть $\mathcal{E} = \{\mathbf{E}_a\}$ — множество ошибок, которые необходимо исправить, а $\mathcal{E}^{(2)} = \{\mathbf{E}_a^\dagger \mathbf{E}_b\}$ — множество произведений пар элементов $\mathcal{E}$. Тогда, чтобы построить невырожденный код, который может исправлять ошибки из $\mathcal{E}$, мы должны найти такой набор генераторов стабилизатора, чтобы некоторый генератор антикоммутировал с каждым элементом $\mathcal{E}^{(2)}$.

Чтобы увидеть, может ли выполнить эту работу k-кубитовый код длины n, начнем с множества $\mathcal{S}^{(n-k)}$ всех абелевых подгрупп группы Паули с $n-k$ генераторами. Будем постепенно отбрасывать подгруппы, которые являются неподходящими стабилизаторами для исправления ошибок в $\mathcal{E}$, а затем посмотрим, останется ли что-нибудь.

Каждая нетривиальная ошибка $\mathbf{E}_a$ коммутирует с долей $\sim 1/2^{n-k}$ всех содержащихся в $\mathcal{S}^{(n-k)}$ групп, так как она должна коммутировать с каждым из $n-k$ генераторов группы. (Существует малая поправка к этой доле, которой можно пренебречь при больших n.) Всякий раз, когда мы добавляем к $\mathcal{E}^{(2)}$ еще один элемент, должна быть отброшена доля 2^{k-n} всех кандидатов в стабилизаторы. Когда $\mathcal{E}^{(2)}$ полностью собрана, мы в худшем случае отбросили долю

$$|\mathcal{E}^{(2)}| \cdot 2^{k-n} \tag{7.198}$$

всех содержащихся в $\mathcal{S}^{(n-k)}$ подгрупп (где $|\mathcal{E}^{(2)}|$ — количество элементов в $\mathcal{E}^{(2)}$). До тех пор, пока эта доля меньше единицы, выполняющий эту работу стабилизатор будет существовать при больших n.

Если мы хотим исправить $t = pn$ ошибок, то $\mathcal{E}^{(2)}$ должно содержать операторы с весом, не превышающим $2t$, что позволяет сделать оценку

$$\log_2 |\mathcal{E}^{(2)}| \lesssim \log_2 \left[\binom{n}{2pn} 3^{2pn} \right] \sim n[H_2(2p) + 2p \log_2 3]. \tag{7.199}$$

Следовательно, существуют стабилизирующие коды, исправляющие pn ошибок, с асимптотическим значением $R = k/n$, определяемым неравенством

$$\log_2 |\mathcal{E}^{(2)}| + k - n < 0, \quad \text{или} \quad R < 1 - H_2(2p) - 2p \log_2 3. \tag{7.200}$$

Это (асимптотическая) форма квантовой границы Гилберта–Варшамова.

Отсюда следует, что должны существовать коды с ненулевой скоростью, которые защищают от ошибок, возникающих с любой вероятностью $p < p_{GV} \simeq 0{,}0946$. Для кода, который может защитить от каждого ошибочного оператора с весом $\leqslant pn$, максимальная вероятность ошибки, допускаемая границей Рейнса, равна $1/6$.

Хотя хорошие квантовые коды существуют, явная конструкция семейств хороших кодов представляет собой совсем другую проблему. Действительно, ни одной такой конструкции не известно.

7.15. Некоторые коды, исправляющие многократные ошибки

7.15.1. Каскадные коды

Все ККК0, которые мы явно сконструировали до настоящего момента, имеют $d = 3$ (или $d = 2$) и, следовательно, могут исправить (в лучшем случае) одну ошибку. Теперь мы опишем несколько примеров кодов с большим расстоянием.

Наиболее простым способом построения кодов, которые могут исправить большее количество ошибок, является каскадное соединение кодов, способных исправить одну ошибку. Каскадный код представляет собой код внутри кода. Предположим, что мы имеем два ККК0 с $k = 1$, $[[n_1, 1, d_1]]$-код C_1 и $[[n_2, 1, d_2]]$-код C_2. Представим принадлежащее C_2 кодовое слово длины n_2, построенное в виде когерентной суперпозиции произведений состояний, в которых каждый кубит находится в одном из состояний $|0\rangle$ или $|1\rangle$. Теперь, используя код C_1, заменим каждый кубит закодированным состоянием длины n_1; то есть заменим $|0\rangle$ на $|\bar{0}\rangle$, а $|1\rangle$ на $|\bar{1}\rangle$ кода C_1. Результатом является код с длиной $n = n_1 n_2$, $k = 1$ и с расстоянием не меньшим, чем $d = d_1 d_2$. Будем называть код C_2 «внешним», а C_1 — «внутренним».

Собственно, мы уже обсуждали один пример такой конструкции: 9-кубитовый код Шора. В этом случае внутренним кодом является трехкубитовый код повторения с генераторами стабилизатора

$$\mathbf{ZZ1}, \qquad \mathbf{1ZZ}, \tag{7.201}$$

а внешним — трехкубитовый «фазовый код» с генераторами стабилизатора

$$\mathbf{XX1}, \qquad \mathbf{1XX} \tag{7.202}$$

(код повторения, повернутый преобразованием Адамара). Стабилизатор каскадного кода строится следующим образом. Прежде всего, в него включаются генераторы внутреннего кода. В рассматриваемом примере это пары генераторов, которые действуют на каждый из трех кубитов, содержащихся в данном блоке внешнего кода, то есть $\mathbf{Z}_1\mathbf{Z}_2$, $\mathbf{Z}_2\mathbf{Z}_3$ и так далее — всего шесть генераторов. Затем добавляются генераторы внешнего кода. В рассматриваемом случае это

$$\bar{\mathbf{X}}\bar{\mathbf{X}}\bar{\mathbf{1}}, \qquad \bar{\mathbf{1}}\bar{\mathbf{X}}\bar{\mathbf{X}}, \tag{7.203}$$

где $\bar{\mathbf{1}} = 111$, а $\bar{\mathbf{X}} = \mathbf{XXX}$, то есть пара генераторов повернутого преобразованием Адамара кода повторения (7.202), но с операторами $\mathbf{1}$ и $\mathbf{X}$, замененными на закодированные операторы внутреннего кода. Вы легко распознаете в них восемь генераторов стабилизатора обсуждавшегося ранее кода Шора. В этом случае внутренний и внешний коды имеют единичное расстояние (например, $\mathbf{Z11}$ коммутирует со стабилизатором внутреннего кода), однако каскадный код имеет расстояние $3 > d = d_1 d_2 = 1$. Это случилось потому, что код был так искусно сконструирован, что закодированные операции внутреннего кода с весами 1 и 2 не коммутируют со стабилизатором внешнего кода. (Все было бы иначе, если бы мы состыковали код повторения с самим собой, а не с фазовым кодом!)

Каскадное соединение кода $[[5, 1, 3]]$ с самим собой дает код с расстоянием $d = 9$ (способный исправить четыре ошибки); $n = 25$ является минимальной длиной любого известного кода при $k = 1$ и $d = 9$. (Код $[[n, 1, 9]]$ при $n = 23, 24$ согласовывался бы с границей Рейнса, но неизвестно, существует ли на самом деле такой код).

Стабилизатор каскадного кода $[[25, 1, 9]]$ имеет 24 генератора. Двадцать из них получаются как четыре генератора $\mathbf{M}_{1,2,3,4}$, действующие на каждый из пяти субблоков внешнего кода, а оставшиеся четыре — это закодированные операторы $\bar{\mathbf{M}}_{1,2,3,4}$ внешнего кода. Отметим, что стабилизатор содержит элементы с весом четыре (элементы стабилизатора, действующие на каждый из пяти внутренних кодов); следовательно, код вырожденный. Это типичный пример каскадного кода.

Нет причин ограничиваться двухуровневым каскадированием кодов. Из L ККО с параметрами $[[n_1, 1, d_1]], \ldots, [[n_L, 1, d_L]]$ можно построить иерархический код всего с L уровнями кодов внутри кодов; он имеет длину

$$n = n_1 n_2 \ldots n_L, \tag{7.204}$$

и расстояние

$$d \geqslant d_1 d_2 \ldots d_L, \tag{7.205}$$

В частности, путем L-кратного каскадирования кода $[[5, 1, 3]]$ можно построить код с параметрами

$$[[5^L, 1, 3^L]]. \tag{7.206}$$

Строго говоря, это семейство кодов не может защитить от количества ошибок, пропорционального их длине. Скорее, отношение количества ошибок t, которые могут быть исправлены, к длине n равно

$$\frac{t}{n} \sim \frac{1}{2}\left(\frac{3}{5}\right)^L, \tag{7.207}$$

что стремится к нулю при большом L. Но расстояние d может оказаться обманчивой мерой того, как хорошо работает код. Вполне достаточно, чтобы восстановление, отказывая лишь для некоторых способов выбора $t \ll pn$ ошибок, оставалось успешным для *типичных* способов выбора pn ошибочных кубитов. Действительно, при большом n и $p > 0$ каскадные коды могут исправить pn типичных ошибок.

Фактически, тот способ, которым обычно используются каскадные коды, не в полной мере реализует их возможности исправления ошибок. Чтобы быть конкретнее, рассмотрим код $[[5, 1, 3]]$ в случае, когда каждый из пяти кубитов независимо подвергается действию деполяризующего канала с вероятностью ошибки p (то есть каждая из ошибок $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ возникает с вероятностью $p/3$). Несомненно, восстановление будет успешным, если в блоке возникнет меньше двух ошибок. Следовательно, как в разделе 7.4.2, вероятность сбоя можно ограничить неравенством

$$p_{\text{fail}} \equiv p^{(1)} \leqslant \binom{5}{2} p^2 = 10p^2. \tag{7.208}$$

Теперь рассмотрим работу каскадного кода $[[25, 1, 9]]$. Чтобы облегчить себе жизнь, выполним восстановление простым (но неоптимальным) способом. Сначала произведем восстановление в каждом из пяти субблоков, измеряя $\mathbf{M}_{1,2,3,4}$, чтобы получить синдромы содержащихся в них ошибок. После этого измерим генераторы стабилизатора $\bar{\mathbf{M}}_{1,2,3,4}$ внешнего кода, чтобы получить его синдром и, если он обнаруживает ошибку, к одному из субблоков применим один из закодированных операторов $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$ или $\bar{\mathbf{Z}}$.

Для внешнего кода восстановление будет успешным, если повреждено не более одного из субблоков, а вероятность $p^{(1)}$ повреждения субблока ограничена неравенством (7.208); таким образом, для кода $[[25, 1, 9]]$ вероятность отказа процедуры восстановления ограничена сверху

$$p^{(2)} \leqslant 10(p^{(1)})^2 \leqslant 10(10p^2)^2 = 1000p^4. \tag{7.209}$$

Очевидно, что это не самая лучшая процедура, поскольку причиной сбоя могут стать четыре ошибки, если они окажутся по две в двух разных субблоках. Так как код имеет расстояние $d = 9$, существует лучшая процедура, которая всегда будет успешно исправлять четыре ошибки, так что $p^{(2)}$ будет иметь порядок p^5, а не p^4. Тем не менее, эта неоптимальная процедура имеет то преимущество, что она очень легко обобщается (и анализируется) в случае многоуровнего соединения.

Действительно, при наличии L уровней каскадного соединения восстановление начинается на самом глубоком внутреннем уровне и прокла-

дывает путь наружу. Решая рекуррентную систему неравенств

$$p^{(\ell)} \leqslant C\left[p^{(\ell-1)}\right]^2 \tag{7.210}$$

с начальным условием $p^{(0)} = p$, мы приходим к выводу, что

$$p^{(L)} \leqslant \frac{1}{C}(Cp)^{2^L} \tag{7.211}$$

(здесь $C = 10$). Нетрудно видеть, что до тех пор, пока $p < 1/10$, вероятность сбоя можно сделать сколь угодно малой, добавляя к коду достаточное количество уровней.

Мы можем записать

$$p^{(L)} \leqslant p_o \left(\frac{p}{p_o}\right)^{2^L}, \tag{7.212}$$

где $p_o = 1/10$ — оценка *пороговой* (то есть допустимой) вероятности ошибки (ниже мы получим лучшие коды и лучшие оценки этого порога). Отметим: чтобы получить

$$p^{(L)} < \varepsilon \tag{7.213}$$

мы можем выбрать размер блока $n = 5^L$; так что

$$n \leqslant \left\lceil \frac{\log(p_o/\varepsilon)}{\log(p_o/p)} \right\rceil^{\log_2 5}. \tag{7.214}$$

В принципе, каскадный код может дать сбой на высоком уровне при гораздо меньшем, чем $n/10$, количестве ошибок, но они должны быть распределены весьма специальным образом, что совсем не характерно для большого n.

Каскадное кодирование неизвестного квантового состояния может выполняться уровень за уровнем. Например, чтобы закодировать $a|0\rangle + b|1\rangle$ в блоке $[[25, 1, 9]]$, мы можем сначала приготовить состояние $a|\bar{0}\rangle + b|\bar{1}\rangle$ в пятикубитовом блоке, используя описанную ранее схему кодирования, а также приготовить четыре пятикубитовых блока в состоянии $|\bar{0}\rangle$. Состояние $a|\bar{0}\rangle + b|\bar{1}\rangle$ можно закодировать на следующем уровне, снова выполняя эту же схему, но теперь со всеми вентилями, замененными на закодированные, действующие на пятикубитовые блоки. Обсуждение конструкции этих закодированных вентилей можно найти в приложении.

7.15.2. Торические коды

Торические коды представляют собой еще одно семейство кодов, которые, как и каскадные, работают гораздо лучше, чем этого можно было бы ожидать, учитывая значения их расстояний. Их описывает профессор Китаев (который их и открыл).[1]

7.15.3. Коды Рида–Маллера

Другим способом построения кодов, которые могут исправить множество ошибок, является осуществление КШС-конструкции. Вспомним, например, частный случай этой конструкции, в котором используется классический код C, содержащийся в дуальном ему коде (в таком случае говорят, что C является «слабо самодуальным» кодом). В КШС-конструкции с каждым смежным классом C в $C^{\perp}$ ассоциируется кодовое слово. Таким образом, мы получаем квантовый код $[[n, k, d]]$, где n – длина кода C, d – (по крайней мере) расстояние кода $C^{\perp}$, а $k = \dim C^{\perp} - \dim C$. Следовательно, для построения КШС-кодов, исправляющих множество ошибок, необходимы слабо самодуальные классические коды с большим минимальным расстоянием.

Один из классов слабо самодуальных классических кодов представляют собой коды Рида–Маллера. Хотя эти коды не особенно эффективны, они очень удобны, поскольку достаточно легко кодируются. Несложно понять также принцип их работы и математическую структуру.[2]

Чтобы подготовиться к построению кодов Рида–Маллера, рассмотрим булевы функции на m битах

$$f : \{0,1\}^m \to \{0,1\}. \tag{7.215}$$

Существует 2^{2^m} таких функций, образующих множество, которое можно рассматривать как двоичное векторное пространство размерности 2^m. Полезно ввести базис в этом пространстве. Вспомним (см. раздел 6.1), что любая булева функция имеет дизъюнктивную нормальную форму. Так как NOT бита x равно $1 - x$, а OR двух битов x и y можно записать как

$$x \vee y = x + y - xy, \tag{7.216}$$

[1] Детальное описание торических кодов можно найти в статье: A. Kitaev, *Fault-tolerant quantum computation by anyons*, Ann. Phys. (NY) **303**(1) pp. 2–30 (2003), quant-ph/9707021. См. также лекции: A. Kitaev, C. Laumann, *Topological phases and quantum computation*, cond-mat/0904.2771. – *Прим. ред.*

[2] См., например, E. J. MacWilliams, N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland Publishing Company, Amsterdam, New York, Oxford (1977), chapter 13; перевод: Ф. Дж. Мак-Вильямс, Н. Дж. А. Слоэн, *Теория кодов, исправляющих ошибки*. – М.: Связь, 1979, гл. 13.

то любая булева функция может быть представлена в виде полинома от m двоичных переменных $x_{m-1}, x_{m-2}, \ldots, x_1, x_0$. Базис векторного пространства полиномов состоит из 2^m функций

$$1, \quad x_i, \quad x_i x_j, \quad x_i x_j x_k, \ldots, \tag{7.217}$$

(где каждый моном представляет собой произведение различных сомножителей, так как $x^2 = x$). Каждую такую функцию f можно представить двоичной строкой длины 2^m, значение которой в позиции, обозначенной двоичной строкой $x_{m-1}, x_{m-2}, \ldots, x_1, x_0$, равно $f(x_{m-1}, x_{m-2}, \ldots, x_1, x_0)$. Например, для $m = 3$

$$\begin{aligned} 1 &= (11111111), \\ x_0 &= (10101010), \\ x_1 &= (11001100), \\ x_2 &= (11110000), \\ x_0 x_1 &= (10001000), \\ x_0 x_2 &= (10100000), \\ x_1 x_2 &= (11000000), \\ x_0 x_1 x_2 &= (10000000). \end{aligned} \tag{7.218}$$

Подпространство этого векторного пространства получается, если ограничить степень полинома до r или менее. Это подпространство является кодом Рида–Маллера (или РМ-кодом) и обозначается $R(r, m)$. Его длина равна $n = 2^m$, а его размерность

$$k = 1 + \binom{m}{1} + \binom{m}{2} + \ldots + \binom{m}{r}. \tag{7.219}$$

Некоторые частные случаи, представляющие интерес:

- $R(0, m)$ — код повторения длины 2^m.
- $R(m-1, m)$ — код, дуальный коду повторения, пространство всех строк четного веса длины 2^m.
- $R(1, 3)$ — код, натянутый на $1, x_0, x_1, x_2$ с параметрами $n = 8$, $k = 4$; фактически, это уже обсуждавшийся расширенный код Хэмминга $[8, 4, 4]$.
- В более общем случае $R(m-2, m)$ при любом $m \geqslant 3$ представляет собой расширенный код Хэмминга с $d = 4$. Выкалывая его (убирая последний бит из каждого кодового слова), мы получим совершенный код Хэмминга $[n = 2^m - 1, k = n - m, d = 3]$.

- $R(1,m)$ имеет $d = 2^{m-1} = \frac{1}{2}n$ и $k = m$. Он дуален расширенному коду Хэмминга и известен как «код Рида–Маллера первого порядка». Он сам по себе представляет значительный практический интерес, благодаря его большому расстоянию и особой простоте процедуры декодирования.

Применяя метод индукции по m, можно вычислить расстояние кода $R(r,m)$. Сначала мы должны определить, как $R(r,m+1)$ связан с $R(r,m)$. Функцию от $x_m,\ldots,x_0$ можно записать как

$$f(x_m,\ldots,x_0) = g(x_{m-1},\ldots,x_0) + x_m h(x_{m-1},\ldots,x_0), \tag{7.220}$$

если f имеет степень r, тогда g должна иметь степень r, а h — степень $r-1$. Рассматривая f как вектор длины 2^{m+1}, имеем[1]

$$f = (g|g) + (h|0), \tag{7.221}$$

где g,h — векторы длиной 2^m. Рассмотрим расстояние между f и

$$f' = (g'|g') + (h'|0). \tag{7.222}$$

При $h = h'$ и $f \neq f'$ это расстояние равно $\mathrm{wt}(f-f') = 2\cdot\mathrm{wt}(g-g') \geqslant 2\cdot\mathrm{dist}(R(r,m))$; при $h \neq h'$ оно по крайней мере $\mathrm{wt}(h-h')) \geqslant \mathrm{dist}(R(r-1,m))$. Если $d(r,m)$ обозначает расстояние кода $R(r,m)$, то можно видеть, что

$$d(r,m+1) = \min[2d(r,m); d(r-1,m)]. \tag{7.223}$$

Теперь с помощью индукции по m можно показать, что $d(r,m) = 2^{m-r}$. Во-первых, проверим, что $d(r,m=1) = 2^{1-r}$ при $r = 0,1$; $R(1,1)$ представляет собой пространство всех строк длины два, а $R(0,1)$ — код повторения длины два. Предположим теперь, что при всех $m \leqslant M$ и $0 \leqslant r \leqslant m$ расстояние $d = 2^{m-r}$. Тогда мы делаем вывод, что при всех $1 \leqslant r \leqslant m$

$$d(r,m+1) = \min(2^{m-r+1}; 2^{m-r+1}) = 2^{m-r+1}. \tag{7.224}$$

[1]Здесь символ $(f|g)$ обозначает следующую конструкцию: если $f = (f_1, f_2 \cdots, f_n)$ — вектор-строка длины n, а $g = (g_1, g_2 \cdots, g_m)$ — вектор-строка длины m, то $(f|g) = (f_1, f_2 \cdots, f_n, g_1, g_2 \cdots, g_m)$ — вектор-строка длины $n+m$. Вывод выражения (7.221) для вектора-строки длины 2^{m+1}, изображающей булеву функцию (7.220), а также подробности вычисления расстояния $d(r,m)$ кода Рида–Маллера $R(r,m)$ можно найти в книге: E. J. MacWilliams, N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland Publishing Company, Amsterdam, New York, Oxford (1977), chapter 13; перевод: Ф. Дж. Мак-Вильямс, Н. Дж. А. Слоэн, *Теория кодов, исправляющих ошибки*. — М.: Связь, 1979, гл. 13. — *Прим. ред.*

Также ясно, что $d(m+1, m+1) = 1$, так как $R(m+1, m+1)$ представляет собой пространство всех двоичных строк длины 2^{m+1}, и что $d(0, m+1) = 2^{m+1}$, так как $R(0, m+1)$ является кодом повторения длины 2^{m+1}. Это завершает индуктивный шаг и доказывает, что $d(r, m) = 2^{m-r}$.

Отсюда, в частности, следует, что $R(m-1, m)$ имеет расстояние два, поэтому дуальным коду $R(r, m)$ является $R(m-r-1, m)$. Прежде всего, заметим, что сумма биномиальных коэффициентов $\binom{m}{j}$ $(0 \leqslant j \leqslant m)$ равна 2^m, так что $R(m-r-1, m)$ имеет правильную размерность, чтобы иметь смысл $R^{\perp}(r, m)$. Тогда этого достаточно, чтобы показать, что $R(m-r-1, m)$ содержится в $R(r, m)$. Но если $f \in R(r, m)$, а $g \in R(m-r-1, m)$, то их произведение является полиномом степени не выше $m-1$ и, следовательно, принадлежит $R(m-1, m)$. Каждый вектор в $R(m-1, m)$ имеет четный вес, поэтому внутреннее произведение $f \cdot g$ обращается в нуль; следовательно, g принадлежит дуальному пространству $R^{\perp}(r, m)$. Это показывает, что

$$R^{\perp}(r, m) = R(m-r-1, m). \tag{7.225}$$

Именно благодаря этому замечательному свойству дуальности коды Рида–Маллера хорошо подходят для КШС-конструкции квантовых кодов.

В частности, код Рида–Маллера является слабо самодуальным при $r \leqslant m-r-1$, или $2r \leqslant m-1$, и самодуальным при $2r = m-1$. В последнем случае его расстояние равно

$$d = 2^{m-r} = 2^{\frac{1}{2}(m+1)} = \sqrt{2n}, \tag{7.226}$$

а количество закодированных битов —

$$k = \frac{1}{2}n = 2^{m-1}. \tag{7.227}$$

При $m = 3, 5, 7$ эти самодуальные коды имеют параметры

$$[8, 4, 4], \quad [32, 16, 8], \quad [128, 64, 16]. \tag{7.228}$$

(Как уже отмечалось, код [8,4,4] является расширенным кодом Хэмминга.) С ними связаны квантовые коды с параметрами $(k = 0)$

$$[[8, 0, 4]], \quad [[32, 0, 8]], \quad [[128, 0, 16]] \tag{7.229}$$

и так далее.

Для того чтобы получить квантовый код с $k = 1$, достаточно *выколоть* самодуальный код Рида–Маллера, то есть удалить из него один из $n = 2^m$

битов (какой бит удалить, значения не имеет). Результатом этой операции является классический код с параметрами $n = 2^m - 1$, $d = 2^{(m+1)/2} - 1 = \sqrt{2(n+1)} - 1$ и $k = (n+1)/2$. Более того, дуальным этому выколотому коду является его четный субкод. (Четный субкод состоит из тех РМ-кодовых слов, для которых удаляемый при выкалывании бит равен нулю, а из самодуальности РМ-кода следует, что они ортогональны всем словам (с четным и нечетным весами) выколотого кода.) Из этих выколотых кодов с помощью КШС-конструкции мы получаем квантовые коды с параметрами $(k = 1)$

$$[[7, 1, 3]], \quad [[31, 1, 7]], \quad [[127, 1, 15]] \tag{7.230}$$

и так далее. Код Хэмминга $[7, 4, 3]$ получается при выкалывании РМ-кода $[8, 4, 4]$, а ККК0, соответствующий коду $[7, 1, 3]$, естественно, является кодом Стина. Эти ККК0 имеют расстояние, возрастающее как квадратный корень их длины.

Эти коды с $k = 1$ не самые эффективные из известных ККК0. Тем не менее, они представляют особый интерес, так как их свойства особенно подходят для применения помехоустойчивых квантовых вентилей на закодированную информацию (см. приложение). В частности, одним полезным свойством самодуальных РМ-кодов является их «двойная четность» — все кодовые слова имеют кратный четырем вес.

Конечно, применяя КШС-конструкцию к РМ-кодам, мы также можем построить квантовые коды с $k > 1$. Например, $R(3, 6)$ с параметрами

$$\begin{aligned} n &= 2^m = 64, \\ d &= 2^{m-r} = 8, \\ k &= 1 + \binom{6}{1} + \binom{6}{2} + \binom{6}{3} = 1 + 6 + 15 + 20 = 42 \end{aligned} \tag{7.231}$$

дуален по отношению к $R(2, 6)$ с параметрами

$$\begin{aligned} n &= 2^m = 64, \\ d &= 2^{m-r} = 16, \\ k &= 1 + \binom{6}{1} + \binom{6}{2} = 1 + 6 + 15 = 22, \end{aligned} \tag{7.232}$$

и, следовательно, КШС-конструкция дает ККК0 с параметрами

$$[[64, 20, 8]]. \tag{7.233}$$

Известны многие другие слабо самодуальные коды, которые можно использовать таким же образом.

7.15.4. Код Голея

С точки зрения чистой математики, самым интересным из когда-либо открытых корректирующих ошибки кодов (классических или квантовых) является код Голея, который был еще и одним из первых, описанных в открытой печати. Здесь мы кратко опишем его, поскольку с помощью КШС-конструкции этот код тоже может трансформироваться в хороший ККК0. (Возможно, этот ККК0 на самом деле не настолько важен, чтобы посвящать ему раздел в этой главе; и все же он достаточно занятный, так что я включил его сюда.)

Код Голея (расширенный) представляет собой самодуальный классический код $[24, 12, 8]$. Если мы выколем его (удалим любой из его 24-х битов), то получим код Голея $[23, 12, 7]$, который может исправить три ошибки. Этот код на самом деле является совершенным, так как он насыщает границу упаковки сфер:

$$1 + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} = 2^{11} = 2^{23-12}. \tag{7.234}$$

На самом деле, совершенные коды, исправляющие больше одной ошибки, — невероятная редкость. Можно показать,[1] что способными исправить больше одной ошибки совершенными кодами (линейными или нелинейными) над *любым* конечным полем являются *всего лишь два*: код $[23, 12, 7]$ и еще один открытый Голеем двоичный код с параметрами $[11, 6, 5]$.

Код Голея $[24, 12, 8]$ имеет очень сложную симметрию. Она характеризуется своей группой автоморфизмов — группой перестановок 24-х битов, преобразующих одни кодовые слова в другие. Это группа Матье M_{24}, открытая в XIX веке спорадическая простая группа порядка 244 823 040.

$2^{12} = 4096$ кодовых слов имеют распределение весов (в очевидном обозначении)

$$0^1 8^{759} 12^{2576} 16^{759} 24^1. \tag{7.235}$$

Отметим, в частности, что каждый вес кратен четырем (код имеет двойную четность). Каков смысл числа $759 (= 3 \cdot 11 \cdot 23)$? На самом деле оно равно

$$\binom{24}{5} \Big/ \binom{8}{5} = 759 \tag{7.236}$$

[1] См. § 6.10 в книге E. J. MacWilliams, N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North Holland Publishing Company, Amsterdam, New York, Oxford (1977); перевод: Ф. Дж. МакВильямс, Н. Дж. Слоэн, *Теория кодов, исправляющих ошибки*. — М.: Связь, 1979.

и возникает по комбинаторной причине: каждое кодовое слово с весом восемь характеризуется своим носителем — 8-элементным множеством («октадой»). Последние выбираются таким образом, чтобы каждое 5-элементное подмножество 24-х битов содержалось (целиком) в одной и только одной такой октаде (отражение высокой симметрии кода).

Что придает коду Голея математическую значимость? Его открытие в 1949 году привело в движение последовательность событий, которые примерно к 1980 году завершились полной классификацией конечных простых групп. Эта классификация является одним из величайших достижений математики XX века.

(Группа является простой, если она не содержит ни одной нетривиальной нормальной подгруппы. Конечные простые группы можно рассматривать как строительные блоки всех конечных групп в том смысле, что для любой конечной группы G существует однозначное разложение вида

$$G \equiv G_0 \supseteq G_1 \supseteq G_2 \supseteq \ldots \supseteq G_n, \tag{7.237}$$

где каждая G_{j+1} представляет собой нормальную подгруппу G_j, а каждая фактор-группа G_j/G_{j+1} является простой. Конечные простые группы можно систематизировать в разные бесконечные семейства, плюс 26 дополнительных не поддающихся классификации «спорадических» простых групп.)

В 1964 году код Голея привел Лича к открытию чрезвычайно плотной укладки шаров в 24-х измерениях, известной как *решетка Лича* Λ. Узлы решетки (центры сфер) представляют собой 24-компонентные целочисленные векторы со следующими свойствами: чтобы определить, содержится ли $\vec{x} = (x_1, x_2, \ldots x_{24})$ в Λ, запишем каждую компоненту x_j в двоичном представлении

$$x_j = \ldots x_{j^3} x_{j^2} x_{j^1} x_{j^0}. \tag{7.238}$$

Тогда $\vec{x} \in \Lambda$, если[1]

(i) все x_{j^0} либо нули, либо единицы;

(ii) x_{j^2} представляют собой четную 24-битовую строку, если все x_{j^0} равны нулю, и — нечетную 24-битовую строку, если все x_{j^0} равны единице;

[1]Некоторые альтернативные определения решетки Лича можно найти в книге J. H. Conway, N. J. A. Sloane, *Sphere Packing, Lattices and Groups*, Springer Verlag, NY, Berlin, et al. (1988), Chapter 4, § 11; перевод: Дж. Конвей, Н. Слоэн, *Упаковки шаров, решетки и группы*. — М.: Мир, 1990, глава 4, § 11.— *Прим. ред.*

(iii) x_{j^1} представляют собой 24-битовую строку, содержащуюся в коде Голея.

При употреблении этих правил отрицательное число представляется его двоичным дополнением, например,

$$\begin{aligned} -1 &= \ldots 1111, \\ -2 &= \ldots 1110, \\ -3 &= \ldots 1101, \\ &\text{и так далее.} \end{aligned} \tag{7.239}$$

Нетрудно проверить, что Λ является решеткой: она замкнута относительно сложения. (На остальные биты, кроме битов последних трех разрядов двоичного разложения x_j, никаких ограничений не накладывается.)

Подсчитаем число ближайших к началу координат[1] соседей (или количество сфер, касающихся любой данной сферы). Все эти точки находятся на расстоянии $(\text{distance})^2 = 32$ от начала координат

$$\begin{aligned} (\pm 2)^8(0)^{16} &: 2^7 \cdot 759, \\ (\pm 3)(\mp 1)^{23} &: 2^{12} \cdot 24, \\ (\pm 4)^2(0)^{22} &: 2^2 \cdot \binom{24}{2}. \end{aligned} \tag{7.240}$$

Таким образом существует $2^7 \cdot 759$ ближайших соседей, радиус-векторы которых имеют восемь отличных от нуля компонент, равных ± 2 (среди них количество отрицательных — четное) и заполняющих позиции, принадлежащие носителям 759 кодовых слов Голея с весом восемь. Далее, существует $2^{12} \cdot 24$ ближайших соседей, радиус-векторы которых имеют одну компоненту, равную ± 3 (она может находиться в любой из 24-х позиций), а остальные 23 компоненты равны ∓ 1. При этом верхние знаки приписываются компонентам, которые заполняют позиции, принадлежащие носителям кодовых слов Голея произвольного веса. Если, например, выбрана компонента $+3$, то занимаемая ею позиция вместе с позициями всех отрицательных компонент (то есть -1) образует носитель одного из 2^{11} оставшихся (из 2^{12}) кодовых слов Голея. Наконец, имеется $2^2 \cdot \binom{24}{2}$ ближайших соседей, радиус-векторы которых имеют только две отличные от нуля компоненты ± 4, положение и знак которых ничем не ограничены. В сумме координационное число решетки равно 196 560.

[1] Символ $(a)^k(b)^l$ обозначает вектор решетки $\vec{x} \in \Lambda$, имеющий k и l компонент, равных a и, соответственно, b. — *Прим. ред.*

Решетка Лича имеет замечательную группу автоморфизмов, открытую Конвэем в 1968 году. Это сохраняющая решетку конечная подгруппа группы вращений $\mathrm{SO}(24)$ пространства размерности 24. Порядок этой конечной группы (известной как $\cdot 0$, или «точка нуль») равен

$$2^{22} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23 = 8\,315\,553\,613\,086\,720\,000 \simeq 8.3 \times 10^{18}. \quad (7.241)$$

Если ее двухэлементный центр выкидывается, получается спорадическая простая группа $\cdot 1$. К моменту ее открытия, $\cdot 1$ была самой большой из построенных спорадических простых групп.

Решетка Лича и ее группа автоморфизмов, в конечном счете, (путем, который не будет здесь описан) в 1982 году привели Грисса к построению наиболее удивительной из всех спорадической простой группы (на существование которой ранее указывали Фишер и Грисс). Это конечная подгруппа группы вращений в пространстве размерности 196 884, порядок которой приблизительно равен 8.08×10^{53}. Это чудище, известное как F_1, получило прозвище «монстр» (хотя Грисс предпочитает называть его «дружелюбным гигантом»). Открытая последней, она является самой большой спорадической простой группой.

Таким образом, классификация конечных простых групп многим обязана (классической) теории кодирования и, в частности, коду Голея. Возможно, теория ККК0 также завещает математике что-нибудь значительное и очень интересное!

Во всяком случае, поскольку (расширенный) код Голея $[24, 12, 8]$ является самодуальным, то получаемый при выкалывании код $[23, 12, 7]$ — слабо самодуальный; дуальный ему код $[23, 11, 8]$ является его собственным субкодом. Отсюда с помощью метода КШС можно построить ККК0 $[[23, 1, 7]]$. Это не самый эффективный квантовый код, который может исправить три ошибки (существует код $[[17, 1, 7]]$, насыщающий границу Рейнса), но он обладает особенно тонкими свойствами, благоприятствующими помехоустойчивым квантовым вычислениям (см. приложение).

7.16. Пропускная способность квантового канала

Как это до сих пор формулировалось, целью построения ККК0 является достижение максимального значения расстояния кода d, при заданных длине n и количестве закодированных кубитов k. Большее расстояние обеспечивает лучшую защиту от ошибок, так как код с расстоянием d может исправить $d - 1$ стираний или $(d - 1)/2$ ошибок в неизвестных позициях. Мы видели, что можно построить хорошие коды, поддерживающие конеч-

ную скорость воспроизведения k/n при большом n и корректирующие pn ошибок, количество которых пропорционально n.

Теперь мы обратимся к другому, но достаточно близкому вопросу об асимптотическом исполнении ККО. Пусть \$ — супероператор, действующий на операторы плотности в гильбертовом пространстве $\mathcal{H}$. Будем рассматривать супероператоры \$, действующие независимо в каждой копии $\mathcal{H}$, содержащейся в n-кратном тензорном произведении

$$\mathcal{H}^{(n)} = \mathcal{H} \otimes \ldots \otimes \mathcal{H}. \tag{7.242}$$

Мы бы хотели выбрать такое кодовое подпространство $\mathcal{H}_{\text{code}}^{(n)}$ пространства $\mathcal{H}^{(n)}$, чтобы содержащаяся в $\mathcal{H}_{\text{code}}^{(n)}$ квантовая информация подвергалась действию супероператора

$$\$^{(n)} = \$ \otimes \ldots \otimes \$ \tag{7.243}$$

и, тем не менее, могла быть декодирована с высокой точностью воспроизведения.

Скорость воспроизведения кода определяется как

$$R = \frac{\log \mathcal{H}_{\text{code}}^{(n)}}{\log \mathcal{H}^{(n)}}; \tag{7.244}$$

это количество кубитов, предназначенных для переноса одного кубита закодированной информации. *Пропускная способность квантового канала* $Q(\$)$ супероператора \$ представляет собой максимум асимптотической скорости воспроизведения, при которой квантовую информацию можно послать по каналу со сколь угодно высокой точностью воспроизведения. Другими словами, $Q(\$)$ является таким наибольшим числом, что для любого $R < Q(\$)$ и любого $\varepsilon > 0$ существует такой код $\mathcal{H}_{\text{code}}^{(n)}$ со скоростью воспроизведения, по крайней мере равной R, что для любого $|\psi\rangle \in \mathcal{H}_{\text{code}}^{(n)}$ состояние $\boldsymbol{\rho}$, *восстановленное после того, как* $|\psi\rangle$ подвергалось действию $\$^{(n)}$, имеет точность воспроизведения

$$F = \langle\psi|\boldsymbol{\rho}|\psi\rangle > 1 - \varepsilon. \tag{7.245}$$

Таким образом, $Q(\$)$ представляет собой квантовую версию определенной Шенноном пропускной способности классического канала с шумом. Как мы уже видели в пятой главе, это не единственный вид пропускной способности, которую можно связать с квантовым каналом. Также большой интерес представляет $C(\$)$ — максимальная скорость, с которой *классическую* информацию можно передавать по квантовому каналу

со сколь угодно малой вероятностью ошибки. Формальный ответ на этот вопрос был сформулирован в разделе 5.4, но только для ограниченного класса возможных схем кодирования; общий ответ до сих пор неизвестен. Пропускная способность квантового канала $Q(\$)$ даже еще менее понятна, чем классическая пропускная способность $C(\$)$ квантового канала. Отметим, что $Q(\$)$ и максимальная асимптотическая скорость воспроизведения k/n, которая может быть достигнута хорошими $[[n, k, d]]$ ККК0 с положительным d/n, суть не одно и то же. В случае пропускной способности квантового канала мы не должны требовать, чтобы код корректировал *любое* возможное распределение pn ошибок, при условии, что ошибки, которые невозможно исправить, становятся в высшей степени атипичными при большом n.

Здесь мы в основном ограничимся обсуждением двух интересных примеров квантовых каналов, действующих на одиночный кубит — квантового стирающего канала (для которого точное значение Q известно) и деполяризующего канала (для которого Q не известна до сих пор, но для нее можно установить полезные верхнюю и нижнюю границы).

Что это за каналы? В случае квантового стирающего канала, переданный кубит либо приходит неповрежденным, либо (с вероятностью p) теряется и его никогда не получают. Мы можем найти унитарное представление этого канала, погружая кубит в трехмерное гильбертово пространство с ортонормированным базисом $\{|0\rangle, |1\rangle, |2\rangle\}$. Канал действует согласно правилу

$$\begin{aligned} |0\rangle \otimes |0\rangle_E &\rightarrow \sqrt{1-p}|0\rangle \otimes |0\rangle_E + \sqrt{p}|2\rangle \otimes |1\rangle_E, \\ |1\rangle \otimes |0\rangle_E &\rightarrow \sqrt{1-p}|1\rangle \otimes |0\rangle_E + \sqrt{p}|2\rangle \otimes |2\rangle_E, \end{aligned} \tag{7.246}$$

где $\{|0\rangle_E, |1\rangle_E, |2\rangle_E\}$ — взаимно ортогональные состояния окружения. Получатель может измерить наблюдаемую $|2\rangle\langle 2|$, чтобы определить, остался ли кубит неповрежденным или был «стерт».

Деполяризующий канал (с вероятностью ошибки p) детально обсуждался в разделе 3.4.1. Мы видим, что при $p \leqslant 3/4$ судьбу переданного по каналу кубита можно описать следующим образом: с вероятностью $1 - q$ (где $q = 4p/3$) кубит доходит неповрежденным, а с вероятностью q — разрушается; в последнем случае его состояние описывается случайной матрицей плотности $\frac{1}{2}1$.

И стирающий, и деполяризующий каналы разрушают кубит с определенной вероятностью. Их главное различие состоит в том, что в случае стирающего канала получатель знает, какие кубиты были разрушены; в случае деполяризующего канала поврежденные кубиты не несут никаких способствующих восстановлению отличительных признаков. Конечно, в обоих

случаях, отправитель не может заранее знать, какие кубиты будут уничтожены.

7.16.1. Стирающий канал

Пропускную способность квантового стирающего канала можно точно определить. Сначала мы установим верхнюю границу для Q, а затем покажем, что существуют коды, достигающие высокой точности воспроизведения и сколь угодно близкой к верхней границе скорости воспроизведения. На первом этапе вывода верхней границы пропускной способности покажем, что $Q = 0$ при $p > 1/2$.

Стирающий канал может быть реализован, если Алиса посылает кубит Бобу, а третья сторона в лице Чарли решает случайным образом, украсть кубит (с вероятностью p) или позволить кубиту дойти до Боба неповрежденным (с вероятностью $1-p$). Если Алиса посылает большое количество кубитов n, то примерно $(1-p)n$ кубитов доходят до Боба, а pn — перехватываются Чарли. Следовательно, при $p > 1/2$ Чарли обладает бóльшим количеством кубитов, чем Боб, и если Боб может восстановить закодированную Алисой квантовую информацию, то вне всякого сомнения это может сделать и Чарли. Следовательно, если $Q(p) > 0$ при $p > 1/2$, Боб и Чарли могут клонировать отправленные Алисой неизвестные закодированные квантовые состояния, что невозможно. (Строго говоря, они могут клонировать с точностью воспроизведения $F = 1 - \varepsilon$, для любого $\varepsilon > 0$.) Таким образом, при $p > 1/2$ пропускная способность квантового канала $Q(p) = 0$.

Чтобы найти границу для $Q(p)$ в случае $p < 1/2$, мы обратимся к следующей лемме. Предположим, Алиса и Боб связаны посредством идеального канала и канала с шумом с пропускной способностью $Q > 0$. Допустим также, что Алиса посылает m кубитов по идеальному каналу и n кубитов по каналу с шумом. Тогда количество закодированных кубитов r, которые Боб может восстановить со сколь угодно высокой точностью воспроизведения, должно удовлетворять неравенству

$$r \leqslant m + Qn. \tag{7.247}$$

Мы получим его, заметив, что Алиса и Боб могут имитировать m отправленных по идеальному каналу кубитов, посылая m/Q кубитов по каналу с шумом и таким образом достигая скорости воспроизведения

$$R = \frac{r}{m/Q + n} = \left(\frac{r}{m + Qn}\right) Q. \tag{7.248}$$

Если бы r превысило $m + Qn$, эта скорость R превысила бы пропускную способность канала с шумом Q, что невозможно. Следовательно, справедливо неравенство (7.247).

Теперь рассмотрим стирающий канал с вероятностью ошибки p_1 и предположим, что $Q(p_1) > 0$. Тогда для $p_2 \leqslant p_1$ мы можем ограничить $Q(p_2)$ неравенством

$$Q(p_2) \leqslant 1 - \frac{p_2}{p_1} + \frac{p_2}{p_1} Q(p_1). \tag{7.249}$$

(Другими словами, если мы строим график $Q(p)$ на плоскости (p, Q) и проводим секущую линию из любой точки (p_1, Q_1) до точки $(p = 0, Q = 1)$, то в интервале $0 \leqslant p \leqslant p_1$ кривая $Q(p)$ не может лежать выше секущей; если $Q(p)$ дважды дифференцируема, то ее вторая производная не может быть отрицательной.) Чтобы получить эту границу, представим, что Алиса посылает Бобу n кубитов, заранее зная, какие $n(1-p_2/p_1)$ из них прибудут неповрежденными. Оставшиеся $n(p_2/p_1)$ кубитов стираются с вероятностью p_1. Следовательно, Алиса и Боб используют как идеальный канал, так и канал с шумом с вероятностью стирания p_1; неравенство (7.247) верно, а скорость воспроизведения R, которой они могут достичь, ограничена неравенством

$$R \leqslant 1 - \frac{p_2}{p_1} + \frac{p_2}{p_1} Q(p_1). \tag{7.250}$$

С другой стороны, при больших n стирается всего около np_2 кубитов, а $(1-p_2)n$ кубитов доходят неповрежденными. Таким образом, Алиса и Боб имеют стирающий канал с вероятностью стирания p_2, но с тем дополнительным преимуществом, что они заранее знают, что некоторые из отправленных Алисой кубитов неуязвимы для стирания. Располагая этой информацией, они оказываются в менее затруднительном положении, чем без нее; отсюда следует (7.249). Эта же граница применима и для деполяризующего канала.

Теперь результат $Q(p) = 0$ при $p > 1/2$ можно скомбинировать с неравенством (7.249). Мы делаем вывод, что кривая $Q(p)$ не может находиться выше прямой линии, соединяющей точки $(p=0, Q=1)$ и $(p=1/2, Q=0)$, или

$$Q(p) \leqslant 1 - 2p, \qquad 0 \leqslant p \leqslant \frac{1}{2}. \tag{7.251}$$

Действительно, существуют стабилизирующие коды, которые фактически достигают скорости воспроизведения $1 - 2p$ при $0 \leqslant p \leqslant 1/2$. Это можно увидеть, позаимствовав идею Клода Шеннона и усреднив по случайным стабилизирующим кодам. Представим выбор (подряд) всех $n - k$

генераторов стабилизатора. Каждый выбирается среди 4^n операторов Паули, имеющих одинаковую априорную вероятность, за исключением того, что каждый новый генератор должен коммутировать со всеми выбранными в предыдущих раундах генераторами.

Теперь Алиса использует этот стабилизирующий код, чтобы закодировать произвольное квантовое состояние в 2^k-мерном кодовом подпространстве, и посылает Бобу n кубитов по стирающему каналу с вероятностью стирания p. Сможет ли Боб восстановить отправленное Алисой состояние?

Боб заменяет каждый стертый кубит кубитом в состоянии $|0\rangle$, а затем приступает к измерению всех $n-k$ генераторов стабилизатора. Из этого измерения синдромов он надеется извлечь оператор Паули $\mathbf{E}$, действующий на замещенные кубиты. Коль скоро $\mathbf{E}$ известен, он может применить $\mathbf{E}^\dagger$, чтобы восстановить идеальную копию отправленного Алисой состояния. При большом n количество кубитов, которые Боб должен заменить, примерно равно pn, и он успешно их восстановит, если существует единственный оператор Паули $\mathbf{E}$, производящий искомый синдром. Если один и тот же синдром имеет более одного оператора Паули, действующего на замещенные кубиты, то восстановление может не удаться.

Какова вероятность сбоя? Так как мы имеем около pn замещенных кубитов, существует около 4^{pn} операторов Паули с носителем на этих кубитах. Более того, для любого конкретного оператора Паули $\mathbf{E}$ случайный стабилизирующий код генерирует случайный синдром — каждый генератор *стабилизатора с вероятность* $1/2$ коммутирует с $\mathbf{E}$ и с такой же вероятностью антикоммутирует. Следовательно, вероятность того, что два оператора Паули имеют одинаковый синдром, равна $(1/2)^{n-k}$.

Существует по крайней мере один действующий на замещенные кубиты особый оператор Паули, который имеет искомый Бобом синдром. Но вероятность того, что другой оператор Паули имеет такой же синдром (а следовательно, вероятность сбоя), не более, чем

$$P_{\text{fail}} \leqslant 4^{pn}\left(\frac{1}{2}\right)^{n-k} = 2^{-n(1-2p-R)}. \tag{7.252}$$

где $R = k/n$ — скорость воспроизведения. Неравенство (7.252) ограничивает вероятность сбоя, если мы *усредняем* по всем стабилизирующим кодам со скоростью R; отсюда следует, что должен существовать по крайней мере один стабилизирующий код, вероятность сбоя которого также удовлетворяет этому неравенству.

Для этого конкретного кода P_{fail} становится сколь угодно малой при $n \to \infty$, при любой скорости воспроизведения $R = 1 - 2p - \delta$, строго

меньшей $1-2p$. Следовательно, $R=1-2p$ асимптотически достижима; объединяя этот результат с неравенством (7.251), мы получаем пропускную способность квантового стирающего канала

$$Q(p)=1-2p, \qquad 0\leqslant p\leqslant \frac{1}{2}. \tag{7.253}$$

Если бы мы хотели гарантировать, что каждому способу повреждения pn стертых кубитов можно сопоставить определенный синдром, тогда нам понадобился бы квантовый код $[[n,k,d]]$ с расстоянием $d>pn$. Граница Гилберта–Варшамова из раздела 7.14 гарантирует существование такого кода при

$$R<1-H_2(p)-p\log_2 3. \tag{7.254}$$

Эта скорость может быть достигнута кодом, который защищает от всех возможных способов стирания до pn кубитов. При $p>0$ она лежит строго ниже пропускной способности, потому что для достижения высокой средней точности воспроизведения достаточно быть способным исправлять *типичные* стирания, а не все возможные ошибки.

7.16.2. Деполяризующий канал

Пропускная способность канала деполяризации до сих пор точно не известна, но мы можем получить для нее некоторые интересные верхнюю и нижнюю границы.

Как и в случае стирающего канала, мы можем найти верхнюю границу для пропускной способности, прибегая к теореме о невозможности клонирования. Вспомним, что для деполяризующего канала с вероятностью ошибки $p<3/4$ каждый кубит с вероятностью $1-4p/3$ проходит неповрежденным, либо с вероятностью $q=4p/3$ рандомизируется (заменяется максимально смешанным $\boldsymbol{\rho}=\frac{1}{2}\mathbf{1}$). Тогда подслушивающий Чарли может имитировать канал, с вероятностью q перехватывая кубиты и замещая каждый украденный кубит максимально смешанным кубитом. При $q>1/2$ Чарли перехватывает больше половины кубитов и находится в более выгодном, чем Боб, положении для декодирования отправленного Алисой состояния. Следовательно, чтобы не позволить клонирование, скорость, с которой Алиса посылает Бобу квантовую информацию, должна быть строго нулевой при $q>1/2$ или $p>3/8$:

$$Q(p)=0, \qquad p>\frac{3}{8}. \tag{7.255}$$

На самом деле можно получить более строгую границу, заметив, что Чарли может избрать лучшую стратегию подслушивания — применить оптимальный *приближенный* клонер, который вы изучили в домашней задаче. Это устройство применяется к каждому отправленному Алисой кубиту и заменяет его двумя кубитами, так что каждый приближается к оригиналу с точностью воспроизведения $F = 5/6$, или

$$|\psi\rangle\langle\psi| \to \left[(1-q)|\psi\rangle\langle\psi| + q\frac{1}{2}\mathbf{1}\right]^{\otimes 2}, \tag{7.256}$$

где $F = 5/6 = 1 - q/2$. Управляя клонером, Чарли и Боб могут получить состояние Алисы, переданное по деполяризующему каналу с $q = 1/3$. Следовательно, достижимая скорость воспроизведения должна стремиться к нулю; иначе, объединяя приближенный клонер и квантовую коррекцию ошибок, Боб и Чарли смогли бы точно клонировать неизвестное состояние Алисы. Таким образом, уже при $q > 1/3$ или $p > 1/4$ пропускная способность должна обращаться в нуль:

$$Q(p) = 0, \qquad p > \frac{1}{4}. \tag{7.257}$$

Учитывая границу (7.249), мы приходим к выводу, что

$$Q(p) \leqslant 1 - 4p, \qquad 0 \leqslant p \leqslant \frac{1}{4}. \tag{7.258}$$

Этот результат фактически совпадает с найденной в разделе 7.8 границей для скорости воспроизведения кодов $[[n, k, d]]$ при $k \geqslant 1$ и $d \geqslant 2pn + 1$. Предел для пропускной способности и граница для допустимой вероятности ошибки кода $[[n, k, d]]$ (а в последнем случае граница Рейнса является более строгой) — это разные понятия. Тем не менее, сходство между ними не так уж и удивительно, поскольку обе эти границы выводятся из теоремы о невозможности клонирования.

Мы можем получить нижнюю границу для пропускной способности, приблизительно подсчитав скорость воспроизведения, которая, как и в случае стирающего канала, может быть достигнута с помощью случайного стабилизирующего кодирования. Теперь, когда Боб измеряет $n - k$ (выбранных случайным образом, коммутирующих) генераторов стабилизатора, он надеется получить синдром, указывающий на единственный из типичных паулиевских операторов ошибок, возникающих с конечной вероятностью, когда деполяризующий канал действует на n отправленных Алисой кубитов. Для любых $\delta, \varepsilon > 0$ и достаточно большого n, количество N_{typ} типичных операторов Паули с полной вероятностью $1 - \varepsilon$ можно ограничить

следующим образом:

$$N_{\text{typ}} \leqslant 2^{n(H_2(p)+p\log_2 3+\delta)}. \tag{7.259}$$

Попытка восстановления может оказаться неудачной, если среди этих типичных операторов Паули существует еще хотя бы один, имеющий тот же синдром, что и фактический оператор ошибки. Поскольку случайный код сопоставляет случайный $(n-k)$-битовый синдром для каждого оператора Паули, вероятность сбоя можно ограничить неравенством

$$P_{\text{fail}} \leqslant 2^{n(H_2(p)+p\log_2 3+\delta)}2^{k-n} + \varepsilon. \tag{7.260}$$

Здесь второй член ограничивает вероятность атипичной ошибки, а первый — вероятность неоднозначного синдрома в случае типичной ошибки. Мы видим, что усредненная по случайным стабилизирующим кодам вероятность сбоя становится сколь угодно малой при больших n, любых $\delta' < 0$ и такой скорости воспроизведения R, что

$$R \equiv \frac{k}{n} < 1 - H_2(p) - p\log_2 3 - \delta'. \tag{7.261}$$

Если усредненная по кодам вероятность сбоя мала, то существует особый код с малой вероятностью сбоя и, следовательно, скорость воспроизведения R достижима; пропускная способность канала деполяризации ограничена снизу неравенством

$$Q(p) \geqslant 1 - H_2(p) - p\log_2 3. \tag{7.262}$$

Не случайно, что достижимая случайным кодированием скорость воспроизведения согласуется с асимптотической формой квантовой верхней границы Хэмминга для скорости воспроизведения невырожденных кодов $[[n,k,d]]$ при $d > 2pn$; к обоим результатам мы приходим, приписывая свой синдром каждой типичной ошибке. Конечно, нижняя граница Гилберта–Варшамова для скорости воспроизведения кодов $[[n,k,d]]$ лежит ниже $Q(p)$, поскольку она получена при условии, что код может исправлять не только типичные, но и *все* ошибки с весом, не превышающим pn.

Это доказательство методом случайного кодирования можно также применить к несколько более общему каналу, в котором возможны ошибки $\mathbf{X}$, $\mathbf{Y}$ и $\mathbf{Z}$, возникающие с различными частотами. (Назовем его «каналом Паули».) Если ошибка $\mathbf{X}$ возникает с вероятностью p_X, ошибка $\mathbf{Y}$ — с вероятностью p_Y, ошибка $\mathbf{Z}$ — с вероятностью p_Z, а с вероятностью $p_I = 1 - p_X - p_Y - p_Z$ не возникает никакой ошибки, то количество типичных

ошибок в n кубитах равно

$$\frac{n!}{(p_X n)!(p_Y n)!(p_Z n)!(p_I n)!} \sim 2^{nH(p_I,p_X,p_Y,p_Z)}, \tag{7.263}$$

где

$$H \equiv H(p_I,p_X,p_Y,p_Z) = \\ = -p_I \log_2 p_I - p_X \log_2 p_X - p_Y \log_2 p_Y - p_Z \log_2 p_Z \tag{7.264}$$

— энтропия Шеннона распределения вероятностей $\{p_I,p_X,p_Y,p_Z\}$. Теперь мы находим

$$Q(p_I,p_X,p_Y,p_Z) \geqslant 1 - H(p_I,p_X,p_Y,p_Z); \tag{7.265}$$

если скорость воспроизведения R удовлетворяет неравенству $R < 1 - H$, тогда снова крайне маловероятно, что отдельный синдром случайного стабилизирующего кода укажет более, чем на один оператор типичной ошибки.

7.16.3. Вырождение и пропускная способность

Наш вывод нижней границы пропускной способности деполяризующего канала имеет близкое сходство с приведенным в разделе 5.1.3 выводом нижней границы пропускной способности классического двоичного симметричного канала. В классическом случае существует согласованная верхняя граница. Если бы скорость воспроизведения была больше, тогда не было бы достаточного количества синдромов, присваиваемых всем типичным ошибкам.

В квантовом случае это рассуждение не проходит, поскольку квантовые коды могут быть вырожденными. Мы не можем требовать существования своего синдрома у каждой типичной ошибки, так как действие некоторых из них в кодовом пространстве может быть тривиальным. Справедливость теряет не только сам вывод; верхняя самосогласованная граница действительно не существует, то есть в квантовом случае *достижимы* скорости воспроизведения, превышающие $1 - H_2(p) - p \log_2 3$.[1]

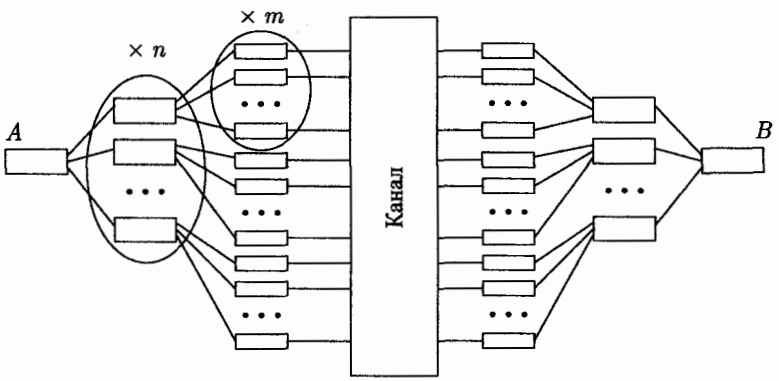
Шор и Смолин исследовали скорость воспроизведения, которая может быть достигнута с помощью каскадного кода, состоящего из случайного

[1] P. M. Shor and J. A. Stmolin, *Quantum Error-Correcting Codes Need Not Completely Reveal the Error Syndrome*, quant-ph/9604006; D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, *Quantum Channel Capacity of Very Noisy Channels*, Phys. Rev. **A57**, pp. 830–839 (1998); quant-ph/9706061.

стабилизирующего кода в качестве внешнего и вырожденного кода с относительно малым размером блока в качестве внутреннего. Согласно их идее, вырождение внутреннего кода позволяет достаточному количеству ошибок действовать тривиально в кодовом пространстве, благодаря чему может быть превышена скорость воспроизведения, достигаемая посредством одного лишь случайного кодирования.

Чтобы изучить эту схему, представим, что как кодирование, так и декодирование выполняются в два этапа. На первом этапе Алиса кодирует выбранное ей состояние в большом n-кубитовом блоке, используя согласованный с Бобом (случайный) внешний код. На втором этапе Алиса кодирует каждый из этих n кубитов в блоке из m кубитов, используя внутренний код. Подобным образом, когда Боб получает nm кубитов, он сначала декодирует каждый внутренний блок из m, а затем — блок из n кубитов.

Очевидно, эту процедуру можно описать на альтернативном языке: Алиса и Боб используют лишь внешний код, но кубиты передаются по составному каналу:



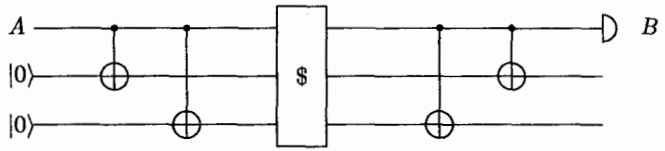
Этот модифицированный канал включает в себя (как показано на рисунке): во-первых, внутреннее кодирование, во-вторых, распространение по исходному каналу с помехами и, наконец, внутреннее декодирование и внутреннее восстановление. Скорость воспроизведения, которая в исходном канале может быть достигнута с помощью каскадного кодирования, такая же, как и скорость, которая может быть достигнута с помощью случайного кодирования в модифицированном канале.

В частности, предположим, что внутренним кодом является m-кубитовый код повторения со стабилизатором

$$\mathbf{Z}_1\mathbf{Z}_2, \quad \mathbf{Z}_1\mathbf{Z}_3, \quad \mathbf{Z}_1\mathbf{Z}_4, \ \ldots, \ \mathbf{Z}_1\mathbf{Z}_m. \tag{7.266}$$

Это далеко не лучший квантовый код; он имеет единичное расстояние, так как нечувствителен к фазовым ошибкам — каждый оператор $\mathbf{Z}_j$ коммутирует со стабилизатором. Но в данном случае для нас важнее высокая высокая степень его вырождения, все ошибки $\mathbf{Z}_i$ эквивалентны.

Кодирующая (и декодирующая) схема для кода повторения состоит лишь из $m-1$ вентилей CNOT, так что наш составной канал выглядит следующим образом (в случае $m=3$):



(Здесь не изображен заключительный восстановительный этап декодирования; например, если оба измеренных кубита показывают 1, то следует инвертировать информационный кубит. В действительности, чтобы упростить исследование составного канала, мы пренебрегаем этим шагом.)

Поскольку CNOT распространяет инвертирование вперед (от управляющего кубита к цели), а обращение фазы назад (от цели к управляющему кубиту), нетрудно понять, что для каждого возможного результата измерения вспомогательных кубитов составной канал является каналом Паули. Представим, что это измерение $m-1$ кубитов внутреннего блока выполняется для каждого из n кубитов внешнего блока. Тогда на каждый из n кубитов действует независимый канал Паули, характеризуемый своим набором параметров (вероятностей ошибок $p_I^{(i)}, p_X^{(i)}, p_Y^{(i)}, p_Z^{(i)}$ для i-го кубита). Таким образом, количество действующих на n кубитов операторов типичных ошибок равно

$$2^{\sum_{i=1}^{n} H_i}, \tag{7.267}$$

где

$$H_i = H(p_I^{(i)}, p_X^{(i)}, p_Y^{(i)}, p_Z^{(i)}) \tag{7.268}$$

— энтропия Шеннона канала Паули, действующего на i-й кубит. Согласно закону больших чисел, для большого n мы получим

$$\sum_{i=1}^{n} H_i = n\langle H\rangle, \tag{7.269}$$

где $\langle H\rangle$ — энтропия Шеннона, усредненная по 2^{m-1} возможным классическим результатам измерения дополнительных кубитов внутреннего кода.

Следовательно, скорость воспроизведения, которая может быть достигнута с помощью случайного внешнего кода, равна

$$R = \frac{1 - \langle H \rangle}{m} \tag{7.270}$$

(мы делим ее на m, потому что каскадный код имеет длину, в m раз большую, чем случайный код).

Шор и Смолин обнаружили, что существуют (m-кратные) коды повторения, для которых (в подходящем диапазоне p) $1 - \langle H \rangle$ является положительной величиной, тогда как $1 - H_2(p) - p \log_2 3$ — отрицательной. Тогда, в этом диапазоне, пропускная способность $Q(p)$ ненулевая, следовательно, нижняя граница (7.262) не является строгой.

Асимптотически неисчезающая скорость воспроизведения достижима посредством случайного кодирования при $1 - H_2(p) - p \log_2 3 > 0$, или $p < p_{\max} \simeq 0{,}18929$. Если случайный внешний код каскадируется с 5-кубитовым внутренним кодом повторения ($m = 5$ оказывается оптимальным выбором), тогда $1 - \langle H \rangle > 0$ при $p < p'_{\max} \simeq 0{,}19036$; максимальная вероятность ошибки, для которой достижима ненулевая скорость воспроизведения, возрастает примерно на 0,6%. То, что в этом диапазоне вероятностей ошибки каскадный код должен превзойти случайный, не является очевидным, хотя, как мы отметили, этого можно было ожидать, из-за (фазового) вырождения кода повторения. Не очевидно также и то, что $m = 5$ должно быть лучшим выбором, но это можно проверить явным вычислением $\langle H \rangle$.[1]

Деполяризующий канал является одним из наиболее простых квантовых каналов. Но даже для этого случая проблема характеристики и вычисления пропускной способности во многом не решена. Этот пример показывает, что из-за возможности вырожденного кодирования проблема пропускной способности для квантовых каналов оказывается куда более острой, чем для классических каналов.

Мы видели, что (если ошибки хорошо описываются деполяризующим каналом) квантовую информацию можно извлечь из квантовой памяти со сколь угодно высокой точностью воспроизведения, пока вероятность ошибки на один кубит меньше 19%. Это является улучшением относительно 10%-ой частоты появления ошибок, которой, как мы обнаружили, можно пользоваться при каскадном соединении кода $[[5, 1, 3]]$. В действительности, коды $[[n, k, d]]$, которые могут исправить вплоть до np ошибок любо-

[1] На самом деле можно достичь дальнейшего очень незначительного улучшения путем каскадного соединения случайного кода с описанным в упражнениях 25-кубитовым обобщенным кодом Шора — тогда ненулевая скорость достигается при $p < p'_{\max} \simeq 0{,}19056$ (еще на 0.1% лучше максимально допустимой вероятности ошибки в коде повторения).

го распределения, согласно границе Рейнса не существуют при $p > 1/6$. Ненулевая пропускная способность возможна для частот появления ошибок в интервале от 16.7% до 19%, потому что для ККК0 достаточно быть способным исправлять типичные ошибки, а не все возможные.

Однако утверждение о том, что восстановление возможно, даже если 19% кубитов подвергаются разрушению, весьма обманчиво в одном важном отношении. Этот результат применим, если кодирование, декодирование и восстановление могут быть выполнены безупречно. Но эти операции на самом деле представляют собой очень сложные квантовые вычисления, которые на практике, конечно, будут чувствительны к ошибкам. Мы не сможем полностью понять, насколько хорошо кодирование может защитить квантовую информацию от повреждений, пока не научимся составлять протокол исправления ошибок, надежный, даже если исполнение самого протокола неидеально. Такие помехоустойчивые протоколы обсуждаются в приложении.

7.17. Итоги

Квантовые коды коррекции ошибок. Коррекция квантовых ошибок может защитить квантовую информацию от декогерентизации и «унитарных ошибок», возникающих вследствие неидеальной реализации квантовых вентилей. В (двоичном) *квантовом коде коррекции ошибок* (ККК0) 2^k-мерное гильбертово пространство k закодированных кубитов $\mathcal{H}_{\text{code}}$ вложено в 2^n-мерное гильбертово пространство n кубитов. Действующие на n кубитов ошибки обратимы при условии, что $\langle\psi|\mathbf{M}_\nu^\dagger\mathbf{M}_\mu|\psi\rangle/\langle\psi|\psi\rangle$ не зависит от $|\psi\rangle$ для любого $|\psi\rangle \in \mathcal{H}_{\text{code}}$ и любых двух операторов Крауса $\mathbf{M}_{\mu,\nu}$, возникающих в разложении супероператора ошибки. Супероператор восстановления преобразует запутывание окружения с кодовым блоком в запутывание окружения со служебным кубитом, который затем может быть выброшен.

Квантовые стабилизирующие коды. Большинство ККК0, которые могут быть построены, представляют собой *стабилизирующие коды*. Двоичный стабилизирующий код характеризуется своим стабилизатором S и абелевой подгруппой n-кубитовой *группы Паули* $G_n = \{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}^{\otimes n}$ (где $\mathbf{X}$, $\mathbf{Y}$, $\mathbf{Z}$ — однокубитовые операторы Паули). Кодовое подпространство представляет собой пространство состояний, одновременно являющихся собственными векторами всех элементов S, соответствующими единичному собственному значению; если S имеет $n-k$ независимых генераторов, тогда существует k закодированных кубитов. Стабилизирующий код может исправить каждую ошибку из подмножества $\mathcal{E}$ группы G_n, если

для каждого $\mathbf{E}_a, \mathbf{E}_b \in \mathcal{E}$ оператор $\mathbf{E}_a^\dagger \mathbf{E}_b$ или принадлежит стабилизатору S, или не принадлежит нормализатору стабилизатора $S^\perp$. Если для $\mathbf{E}_{a,b} \in \mathcal{E}$ некоторый оператор $\mathbf{E}_a^\dagger \mathbf{E}_b$ принадлежит S, то код *вырожден*; в противном случае — *невырожден*. Операторы из $S^\perp \setminus S$ представляют собой «логические» операторы, действующие на закодированную квантовую информацию. Стабилизатор S может быть связан с аддитивным кодом над конечным полем $\mathrm{GF}(4)$, самоортогональным относительно симплектического внутреннего произведения. *Весом* оператора Паули является количество кубитов, на которые он действует нетривиально, а *расстоянием* d стабилизирующего кода — минимальный вес элемента из $S^\perp \setminus S$. Код, имеющий длину n, k закодированных кубитов и расстояние d, называется квантовым кодом $[[n, k, d]]$. Если код осуществляет восстановление от любого супероператора ошибки с носителем на операторах Паули с весами, не превышающими t, то мы говорим, что код «может исправить t ошибок». Код с расстоянием d может исправить $(d-1)/2$ ошибок в неизвестных позициях и $d-1$ ошибок в известных позициях. Можно построить «хорошие» семейства стабилизирующих кодов, в которых d/n и k/n остаются отличными от нуля при $n \to \infty$.

Примеры. Квантовый код $[[5, 1, 3]]$, связанный с классическим кодом Хэмминга над $\mathrm{GF}(4)$, представляет собой код минимальной длины, способный исправить одну ошибку. По заданному классическому линейному коду C_1 и его субкоду $C_2 \subseteq C_1$ можно построить квантовый код Колдербенка–Шора–Стина (КШС-код) с $k = \dim(C_1) - \dim(C_2)$ закодированными кубитами. Расстояние d КШС-кода удовлетворяет неравенству $d \geqslant \min(d_1, d_2^\perp)$, где d_1 — расстояние C_1, а $d_2^\perp$ — расстояние $C_2^\perp$, дуального коду C_2. Простейшим КШС-кодом является квантовый код $[[7, 1, 3]]$, построенный из классического кода Хэмминга $[7, 4, 3]$ и его четного субкода. *Каскадное соединение* квантовых кодов $[[n_1, 1, d_1]]$ и $[[n_2, 1, d_2]]$ дает вырожденный код $[[n_1 n_2, 1, d]]$ с $d \geqslant d_1 d_2$.

Пропускная способность квантового канала. Пропускной способностью квантового канала (квантового канала с помехами) является максимальная скорость воспроизведения, с которой квантовая информация может быть передана по каналу и декодирована со сколь угодно высокой точностью воспроизведения. Пропускная способность двоичного стирающего канала с вероятностью стирания p равна $Q(p) = 1 - 2p$ при $0 \leqslant p \leqslant 1/2$. Пропускная способность двоичного деполяризующего канала до сих пор неизвестна. Ее вычисление представляет собой довольно тонкую проблему, поскольку оптимальный код может быть вырожденным; в частности, случайные коды не достигают асимптотически оптимальной скорости воспроизведения по квантовому каналу.

7.18. Упражнения

7.1. Код коррекции фазовых ошибок. a) Постройте генераторы стабилизатора для кода с $n = 3$, $k = 1$, который может исправить одно инвертирование бита; убедитесь в том, что восстановление возможно при любой ошибке из множества $\mathcal{E} = \{\mathbf{111}, \mathbf{X11}, \mathbf{1X1}, \mathbf{11X}\}$. Найдите ортонормированный базис для двумерного кодового подпространства.

b) Постройте генераторы стабилизатора для кода с $n = 3$, $k = 1$, который может исправить одну фазовую ошибку; убедитесь в том, что восстановление возможно при любой ошибке из множества $\mathcal{E} = \{\mathbf{111}, \mathbf{Z11}, \mathbf{1Z1}, \mathbf{11Z}\}$. Найдите ортонормированный базис для двумерного кодового подпространства.

7.2. Коды обнаружения ошибок. a) Постройте генераторы стабилизатора для квантового кода $[[n, k, d]] = [[3, 0, 2]]$. Используя этот код, мы можем обнаружить любую однокубитовую ошибку. Найдите закодированное состояние. (Не кажется ли оно вам знакомым?)

b) Два ККО C_1 и C_2 (с одинаковой длиной n) *эквивалентны*, если перестановка кубитов, в совокупности с однокубитовым унитарным преобразованием, превращает кодовое подпространство C_1 в подпространство C_2. Все ли стабилизирующие коды $[[3, 0, 2]]$ эквивалентны?

c) Существует ли стабилизирующий код $[[3, 1, 2]]$?

7.3. Максимальное запутывание. Рассмотрите квантовый код $[[5, 1, 3]]$, генераторы стабилизатора которого $\mathbf{M}_1 = \mathbf{XZZX1}$, а $\mathbf{M}_{2,3,4}$ получаются из $\mathbf{M}_1$ циклическими перестановками, и выберите в качестве закодированной операции $\bar{\mathbf{Z}} = \mathbf{ZZZZZ}$. Из закодированных состояний $|\bar{0}\rangle$ с $\bar{\mathbf{Z}}|\bar{0}\rangle = |\bar{0}\rangle$ и $|\bar{1}\rangle$ с $\bar{\mathbf{Z}}|\bar{1}\rangle = -|\bar{1}\rangle$ постройте код с $n = 6$, $k = 0$, закодированное состояние которого имеет вид

$$\frac{1}{\sqrt{2}}(|0\rangle \otimes |\bar{0}\rangle + |1\rangle \otimes |\bar{1}\rangle). \tag{7.271}$$

a) Постройте множество генераторов стабилизатора для этого кода с $n = 6$, $k = 0$.

b) Определите расстояние этого кода. (Вспомните, что для кода с $k = 0$ расстояние определяется как минимальный вес любого элемента стабилизатора.)

c) Найдите $\boldsymbol{\rho}^{(3)}$, матрицу плотности, которая получается, если выбираются три кубита, а по состояниям трех других кубитов вычисляется след.

7.4. Кодовые слова и нелокальность. Для кода $[[5,1,3]]$ с генераторами стабилизатора и логическими операторами из предыдущей задачи:

a) Выразите $\bar{\mathbf{Z}}$ как оператор Паули с весом три, через тензорное произведение операторов $\mathbf{1}$, $\mathbf{X}$ и $\mathbf{Z}$ (без $\mathbf{Y}$). Обратите внимание, что вследствие цикличности кода все циклические перестановки вашего выражения являются эквивалентными способами представления $\bar{\mathbf{Z}}$.

b) Используйте предположение об эйнштейновской локальности (скрытые локальные переменные), чтобы предсказать зависимость между пятью (связанными циклически) найденными в **(a)** наблюдаемыми и наблюдаемой $\mathbf{ZZZZZ}$. Выполняется ли эта связь между наблюдаемыми в состоянии $|\bar{0}\rangle$?

c) Что сказал бы об этом Эйнштейн?

7.5. Обобщенный код Шора. Для целого числа $m \geqslant 2$, рассмотрите обобщение 9-кубитового кода Шора с параметрами $n = m^2$, $k = 1$ и кодовым подпространством, натянутым на два состояния:

$$\begin{aligned} |\bar{0}\rangle &= (|000\ldots0\rangle + |111\ldots1\rangle)^{\otimes m}, \\ |\bar{1}\rangle &= (|000\ldots0\rangle - |111\ldots1\rangle)^{\otimes m}. \end{aligned} \tag{7.272}$$

a) Постройте генераторы стабилизатора для этого кода, а так же логические операции $\bar{\mathbf{Z}}$ и $\bar{\mathbf{X}}$, такие что

$$\begin{aligned} \bar{\mathbf{Z}}|\bar{0}\rangle &= |\bar{0}\rangle, & \bar{\mathbf{X}}|\bar{0}\rangle &= |\bar{1}\rangle, \\ \bar{\mathbf{Z}}|\bar{1}\rangle &= -|\bar{1}\rangle, & \bar{\mathbf{X}}|\bar{1}\rangle &= |\bar{0}\rangle. \end{aligned} \tag{7.273}$$

b) Каково расстояние этого кода?

c) Предположите, что m — нечетное число и что каждый из $n = m^2$ кубитов подвергается действию деполяризующего канала с вероятностью ошибки p. Насколько хорошо этот код защищает закодированный кубит? В частности,

(i) в главном нетривиальном порядке по p, оцените вероятность логической ошибки инвертирования бита $|\bar{0}\rangle \leftrightarrow |\bar{1}\rangle$,

(ii) в главном нетривиальном порядке по p, оцените вероятность логической фазовой ошибки $|\bar{0}\rangle \to |\bar{0}\rangle$, $|\bar{1}\rangle \to -|\bar{1}\rangle$.

d) Рассмотрите асимптотическое поведение вашего результата в **(c)** при большом m. Какому условию должно удовлетворять p, чтобы код обеспечил хорошую защиту: (i) от инвертирования битов и (ii) от фазовых ошибок в пределе $n \to \infty$?

7.6. Кодирующие схемы. Для квантового кода $[[n, k, d]]$ кодирующим преобразованием служит унитарное преобразование $\mathbf{U}$, действующее как

$$\mathbf{U} : |\psi\rangle \otimes |0\rangle^{\otimes(n-k)} \rightarrow |\bar{\psi}\rangle, \tag{7.274}$$

где $|\psi\rangle$ — произвольное k-кубитовое состояние, а $|\bar{\psi}\rangle$ — соответствующее закодированное состояние. Разработайте квантовую схему, осуществляющую кодирующее преобразование для

a) кода Шора $[[9, 1, 3]]$;

b) кода Стина $[[7, 1, 3]]$.

7.7. Укорачивание квантового кода. a) Рассмотрите двоичный стабилизирующий код $[[n, k, d]]$. Покажите, что можно выбрать $n - k$ генераторов стабилизатора так, чтобы на последний кубит нетривиально действовали не более двух (то есть оставшиеся $n - k - 2$ генераторов применяют к последнему кубиту 1).

b) Эти $n - k - 2$ генераторов стабилизатора, применяющих 1 к последнему кубиту, по-прежнему будут коммутирующими и независимыми, если мы выбросим последний кубит. Следовательно, они представляют собой генераторы для кода с длиной $n-1$ и $k+1$ закодированными кубитами. Покажите, что если исходный код невырожден, то расстояние укороченного кода равно по крайней мере $d - 1$. (**Указание:** Сначала покажите, что если существует элемент $(n - 1)$-кубитовой группы Паули с весом t, коммутирующий со стабилизатором укороченного кода, то существует элемент n-кубитовой группы Паули с весом, не превышающим $t + 1$, коммутирующий со стабилизатором исходного кода.)

c) Примените процедуру укорачивания **(a)** и **(b)** для ККК $[[5, 1, 3]]$. Узнаёте ли вы полученный код? (**Указание:** Может оказаться полезным воспользоваться свободой выбора базиса для некоторых из кубитов.)

7.8. Коды для кудитов. *Кудит* представляет собой d-мерную квантовую систему. Действующие на кубиты операторы Паули $\mathbf{I}$, $\mathbf{X}$ и $\mathbf{Z}$ можно обобщить на кудиты следующим образом. Пусть $\{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$ обозначает ортонормированный базис в гильбертовом пространстве одного кудита. Определите операторы:

$$\begin{aligned} \mathbf{X} &: |j\rangle \rightarrow |j + 1 \,(\mathrm{mod}\, d)\rangle, \\ \mathbf{Z} &: |j\rangle \rightarrow \omega^j |j\rangle, \end{aligned} \tag{7.275}$$

где $\omega = \exp(2\pi i/d)$. Тогда $d \times d$-операторы Паули $\mathbf{E}_{r,s}$ равны

$$\mathbf{E}_{r,s} = \mathbf{X}^r \mathbf{Z}^s, \quad r, s = 0, 1, \ldots, d-1. \tag{7.276}$$

a) Образуют ли $\mathbf{E}_{r,s}$ базис в пространстве операторов, действующих на кубит? Унитарны ли они? Вычислите $\mathrm{tr}(\mathbf{E}_{r,s}^\dagger \mathbf{E}_{t,u})$.

b) Операторы Паули удовлетворяют условиям

$$\mathbf{E}_{r,s}\mathbf{E}_{t,u} = \eta_{r,s;t,u}\mathbf{E}_{t,u}\mathbf{E}_{t,u}, \tag{7.277}$$

где $\eta_{r,s;t,u}$ — фазовый множитель. Вычислите этот фазовый множитель. n-кратные тензорные произведения этих действующих на кудит операторов Паули образуют группу $G_n^{(d)}$ порядка d^{2n+1} (и если мы удалим его d-элементный центр, то получим группу $\bar{G}_n^{(d)}$ порядка d^{2n}). Чтобы построить стабилизирующий код для кудитов, мы выбираем абелеву подгруппу группы $G_n^{(d)}$ с $n-k$ генераторами; такой код является общим собственным состоянием этих генераторов с собственным значением единица. Если d — простое число, то кодовое подпространство имеет размерность d^k: k логических кудитов закодированы в блоке из n кудитов.

c) Объясните, насколько иной может быть размерность, если d не является простым числом. (**Указание:** Рассмотрите случай $d = 4$ и $n = 1$.)

7.9. Измерение синдрома для кудитов. Ошибки в кудитах выявляются при измерении генераторов стабилизатора. С этой целью мы можем осуществить двухкудитовый вентиль SUM (который обобщает вентиль CNOT), действующий как

$$\mathrm{SUM} : |j\rangle \otimes |k\rangle \rightarrow |j\rangle \otimes |k + j \,(\mathrm{mod}\ d)\rangle. \tag{7.278}$$

a) Опишите содержащую вентили SUM квантовую схему, которую можно осуществить для измерения n-кудитовой наблюдаемой вида

$$\bigotimes_a \mathbf{Z}_a^{s_a}. \tag{7.279}$$

Если d — простое число, то для каждого $r, s = 0, 1, 2, \ldots, d-1$ существует такой однокудитовый унитарный оператор $\mathbf{U}_{r,s}$, что

$$\mathbf{U}_{r,s}\mathbf{E}_{r,s}\mathbf{U}_{r,s}^\dagger = \mathbf{Z}. \tag{7.280}$$

b) Опишите содержащую вентили SUM и $\mathbf{U}_{r,s}$ квантовую схему, которую можно осуществить для измерения произвольного элемента $G_n^{(d)}$ вида

$$\bigotimes_a \mathbf{E}_{r_a,s_a}. \tag{7.281}$$

7.10. Коды обнаружения ошибок для кудитов. Кудит с $d = 3$ называется *кутритом*. Рассмотрите кутритовый стабилизирующий код с длиной $n = 3$ и с одним ($k = 1$) закодированным кутритом, определяемый двумя генераторами стабилизатора

$$\mathbf{ZZZ}, \quad \mathbf{XXX}. \tag{7.282}$$

a) Коммутируют ли генераторы?

b) Определите расстояние кода.

c) Найдите явное выражение ортонормированного базиса для трехмерного кодового подпространства через ортонормированный базис $\{|0\rangle, |1\rangle, |2\rangle\}$ для кутрита.

d) Постройте генераторы стабилизатора для $n = 3m$ кутритового кода (где m — произвольное положительное целое число) с $k = n - 2$, который может обнаружить одну ошибку.

e) Постройте генераторы стабилизатора для выявляющего одну ошибку кудитового кода с параметрами $n = d$, $k = d - 2$.

7.11. Коды коррекции ошибок для кудитов. Рассмотрите кудитовый стабилизирующий код при $n = 5$, $k = 1$ с генераторами стабилизатора

$$\begin{array}{lccccc} \mathbf{M}_1 = & \mathbf{X} & \mathbf{Z} & \mathbf{Z}^{-1} & \mathbf{X}^{-1} & \mathbf{1} \\ \mathbf{M}_2 = & \mathbf{1} & \mathbf{X} & \mathbf{Z} & \mathbf{Z}^{-1} & \mathbf{X}^{-1} \\ \mathbf{M}_3 = & \mathbf{X}^{-1} & \mathbf{1} & \mathbf{X} & \mathbf{Z} & \mathbf{Z}^{-1} \\ \mathbf{M}_4 = & \mathbf{Z}^{-1} & \mathbf{X}^{-1} & \mathbf{1} & \mathbf{X} & \mathbf{Z} \end{array} \tag{7.283}$$

(второй, третий и четвертый генераторы получены из первого при помощи циклических перестановок кудитов).

a) Определите порядок каждого генератора. Действительно ли генераторы независимы? Коммутируют ли они? Является ли пятая циклическая перестановка $\mathbf{ZZ}^{-1}\mathbf{X}^{-1}\mathbf{1X}$ независимой от остальных?

b) Определите расстояние этого кода. Является ли код невырожденным?

c) Постройте закодированные операции $\bar{\mathbf{X}}$ и $\bar{\mathbf{Z}}$, выразив их как операторы с весом три. (Убедитесь в том, что для любого значения d эти операторы подчиняются правильным коммутационным соотношениям).

Решения упражнений к главе 7[1]

7.1. Коды коррекции фазовых ошибок

a) Квантовый код коррекции инвертирования бита является классическим кодом повторения. Он имеет генераторы стабилизатора

$$\mathbf{M}_1 = \mathbf{ZZ1},$$
$$\mathbf{M}_2 = \mathbf{1ZZ}.$$

Выбор представления закодированных операторов в виде

$$\bar{\mathbf{X}} = \mathbf{XXX},$$
$$\bar{\mathbf{Z}} = \mathbf{ZZZ}$$

отбирает базис кодовых слов

$$|\bar{0}\rangle = |000\rangle,$$
$$|\bar{1}\rangle = |111\rangle.$$

Другое представление закодированных операторов привело бы к другому базису. Данный выбор делает этот код наиболее похожим на его классический аналог.

b) Квантовый код коррекции обращения фазы также представляет собой классический код повторения, хотя и в другом базисе. Он имеет генераторы стабилизатора

$$\mathbf{M}_1 = \mathbf{XX1},$$
$$\mathbf{M}_2 = \mathbf{1XX}.$$

Выбор такого же, как и выше, представления закодированных операторов

$$\bar{\mathbf{X}} = \mathbf{XXX},$$
$$\bar{\mathbf{Z}} = \mathbf{ZZZ}$$

отбирает базис кодовых слов

$$|\bar{0}\rangle = \frac{1}{\sqrt{8}}(|0\rangle + |1\rangle)^{\otimes 3},$$

$$|\bar{1}\rangle = \frac{1}{\sqrt{8}}(|0\rangle - |1\rangle)^{\otimes 3}.$$

[1] Решения выполнены Эндрю Лэндалом.

Заметим, что если мы локально совершим адамаровский поворот базиса каждого кубита [$\mathbf{H}|0\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $\mathbf{H}|1\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$] и операторов стабилизатора [$\mathbf{HZH}^{-1} = \mathbf{X}$, $\mathbf{HXH}^{-1} = \mathbf{Z}$], то получим тот же самый код, что и в части **(a)**. По этой причине можно сказать, что квантовые коды инвертирования бита и обращения фазы *эквивалентны с точностью до локальных унитарных преобразований* [ср. задачу **7.2(b)**].

7.2. Коды детектирования ошибок

a) Код $[[3, 0, 2]]$ можно построить, дополняя один из кодов из задачи **7.1** закодированной операцией (коммутирующей со стабилизатором). Чтобы сохранить расстояние кода, эту закодированную операцию следует выбрать с минимальным весом не ниже двух. [Имея в виду то, как в задаче **7.3(b)** определяется расстояние кода с $k = 0$.] Одним из таких выборов является

$$\begin{aligned}\mathbf{M}_1 &= \mathbf{ZZ1},\\ \mathbf{M}_2 &= \mathbf{1ZZ},\\ \mathbf{M}_3 &= \mathbf{XXX}.\end{aligned}$$

Закодированным состоянием [то есть $\bar{\mathbf{X}}|\bar{0}\rangle$ из **7.1(a)**] является знакомое трехкубитовое кот-состояние, также известное как состояние Гринбергера–Горна–Цайлингера (ГГЦ):

$$\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle).$$

b) Да, в этом смысле все стабилизирующие коды $[[3, 0, 2]]$ эквивалентны. Покажем, что это так, различными способами конструируя их генераторы стабилизаторов. Такой подход ведет к ответу, который больше похож на решение «логической загадки».

Начнем с того, что стабилизатор содержит всего $n - k = 3$ генератора, которые в самом общем случае имеют вид

$$\begin{aligned}\mathbf{M}_1 &= \pm\mathbf{ABC},\\ \mathbf{M}_2 &= \pm\mathbf{DEF},\\ \mathbf{M}_3 &= \pm\mathbf{GHJ}.\end{aligned}$$

С помощью локальных унитарных преобразований (ЛУП) всегда можно избавиться от общих фаз генераторов.

Код должен антикоммутировать со всеми ошибками единичного веса (чтобы детектировать их!), то есть $\mathbf{A}$, $\mathbf{D}$ и $\mathbf{G}$ не могут быть все одинаковыми. Следовательно, без потери общности можно считать $\mathbf{A} \neq \mathbf{D}$, более того, используя ЛУП, преобразовать $\mathbf{A} \to \mathbf{X}$, $\mathbf{D} \to \mathbf{Z}$. Тогда, действуя на $\mathbf{M}_3$ операторами $\mathbf{M}_1$ или $\mathbf{M}_2$ (или обоими) до тех пор пока $\mathbf{G}$ не обратится в единичный $\mathbf{1}$, получим

$$\begin{aligned}\mathbf{M}_1 &= \mathbf{XBC},\\ \mathbf{M}_2 &= \mathbf{ZEF},\\ \mathbf{M}_3 &= \mathbf{1HJ}.\end{aligned}$$

Все генераторы должны коммутировать, следовательно либо $\{\mathbf{B},\mathbf{E}\} = 0$ либо $\{\mathbf{C},\mathbf{F}\} = 0$. Вновь без потери общности (при необходимости поменяв местами второй и третий кубиты) можно положить $\{\mathbf{B},\mathbf{E}\} = 0$ и, применяя ЛУП, выбрать $\mathbf{B} = \mathbf{X}$ и $\mathbf{E} = \mathbf{Z}$. Теперь генераторы выглядят как

$$\begin{aligned}\mathbf{M}_1 &= \mathbf{XXC},\\ \mathbf{M}_2 &= \mathbf{ZZF},\\ \mathbf{M}_3 &= \mathbf{1HJ}.\end{aligned}$$

Так как ошибка с единичным весом $\mathbf{11J}$ не должна коммутировать со стабилизатором, операторы $\mathbf{C}$ и $\mathbf{F}$ не могут одновременно быть равными единичному $\mathbf{1}$. Как мы увидим в задаче **7.3(b)** код с расстоянием два и $k=0$ не может иметь элементов стабилизатора с меньшим, чем два, весом. (В этом смысле все коды с $k = 0$ невырождены.) Следовательно, ошибка $\mathbf{11J}$ не может принадлежать стабилизатору, то есть $\mathbf{H} \neq \mathbf{1}$. Но поскольку $[\mathbf{M}_1, \mathbf{M}_2] = 0$, то $[\mathbf{C},\mathbf{F}] = 0$. Следовательно, или $\mathbf{C} = \mathbf{F} \neq \mathbf{1}$, или только один из них равен единичному оператору. Так как мы всегда можем применить ЛУП $\mathbf{X} \leftrightarrow \mathbf{Z}$ одновременно к первым двум кубитам, то без потери общности можно выбрать оператор $\mathbf{C}$ не равным единичному. Подействовав другим ЛУПом на третий кубит, можно положить $\mathbf{C} = \mathbf{X}$. Следовательно, оператор $\mathbf{F}$ равен или $\mathbf{X}$, или $\mathbf{1}$. Если $\mathbf{F} = \mathbf{X}$, то, выполняя отображение $\mathbf{M}_2 \to \mathbf{M}_1\mathbf{M}_2 = \mathbf{YY1}$, а затем применяя к первым двум кубитам ЛУП $\mathbf{Y} \leftrightarrow \mathbf{Z}$, второй генератор $\mathbf{M}_2$ можно преобразовать в $\mathbf{ZZ1}$. Следовательно, не теряя общности, можно выбрать $\mathbf{F} = \mathbf{1}$. Теперь генераторы выглядят как

$$\begin{aligned}\mathbf{M}_1 &= \mathbf{XXX},\\ \mathbf{M}_2 &= \mathbf{ZZ1},\\ \mathbf{M}_3 &= \mathbf{1HJ}.\end{aligned}$$

Наконец, из равенств $[\mathbf{M}_2, \mathbf{M}_3] = 0$ и $[\mathbf{M}_1, \mathbf{M}_3] = 0$ следует, что $\mathbf{H} = \mathbf{Z}$, а $\mathbf{J}$ равен или $\mathbf{Z}$, или $\mathbf{1}$. Но $\mathbf{J}$ не может быть равным $\mathbf{1}$, поскольку тогда ошибка с единичным весом $\mathbf{1Z1}$ коммутировала бы с $\mathbf{M}_3$ (фактически, в этом случае $\mathbf{1Z1}$ *совпадает* с $\mathbf{M}_3$). Следовательно, $\mathbf{J} = \mathbf{Z}$ и, значит, наиболее общий стабилизирующий код $[[3, 0, 2]]$ имеет генераторы

$$\mathbf{M}_1 = \mathbf{XXX},$$
$$\mathbf{M}_2 = \mathbf{ZZ1},$$
$$\mathbf{M}_3 = \mathbf{1ZZ}.$$

c) Нет, стабилизирующий код $[[3, 1, 2]]$ не существует. Для того чтобы такой код детектировал все возможные однокубитовые ошибки, каждый столбец генераторов стабилизатора должен содержать как оператор $\mathbf{X}$, так и $\mathbf{Z}$ (или $\mathbf{Y}$). Однако код $[[3, 1, 2]]$ имеет только два генератора стабилизатора, и нет возможности сделать коммутирующими произведение трех пар антикоммутирующих операторов.

Замечание. То, что код $[[3, 1, 2]]$ не существует, может показаться удивительным, поскольку, как мы видели в задаче **7.1(a)**, коды с $n = 3$, $k = 1$, детектирующие ошибку инвертирования бита и обращения фазы, *существуют*. Дело в том, что не существует кода с $n = 3$, $k = 1$, способного детектировать *все* возможные ошибки; такие коды могут только корректировать ошибки, возникающие в некотором *базисе*. Поскольку кубиты весьма дороги, то, стремясь максимизировать эффективность квантового кода коррекции ошибок с $n = 3$, важно знать, в каком базисе предпочитает действовать окружение.

7.3. Максимальное запутывание

a) Пусть $|\psi\rangle$ обозначает закодированное состояние. Мы должны найти $n - k = 6$ линейно независимых фиксирующих $|\psi\rangle$ операторов. Данное в задаче разложение типа Шмидта для $|\psi\rangle$ делает ясным, что это состояние фиксируется всеми операторами вида $\mathbf{1M}$, где $\mathbf{M}$ принадлежит стабилизатору кода $[[5, 1, 3]]$. Более того, поскольку $|\psi\rangle$ имеет ту же природу, что и кот-состояние, то оно фиксируется и операторами $\mathbf{X\bar{X}}$ и $\mathbf{Z\bar{Z}}$. Поэтому полный список генераторов стабилизатора этого кода выглядит следующим образом:

$$\mathbf{M}_1 = \mathbf{1XZZX1},$$

$$\mathbf{M}_2 = \mathbf{11XZZX},$$
$$\mathbf{M}_3 = \mathbf{1X1XZZ},$$
$$\mathbf{M}_4 = \mathbf{1ZX1XZ},$$
$$\mathbf{M}_5 = \mathbf{XXXXXX},$$
$$\mathbf{M}_6 = \mathbf{ZZZZZZ}.$$

b) Расстояние кода равно четырем, и он невырожден. Эти выводы зависят от того, как интерпретируется расстояние и вырождение квантового кода с $k = 0$.

Замечания о расстоянии при $k = 0$

Для кодов с $k = 0$ стандартное понятие расстояния определено недостаточно четко. Обычно мы говорим, что расстоянием кода является наибольший вес представлений минимального веса для закодированных операторов $\bar{\mathbf{X}}$ и $\bar{\mathbf{Z}}$. Но для кодов с $k = 0$ не существует закодированных операторов! Следовательно, мы должны вернуться к самым основам, чтобы понять, что означает расстояние в этом случае.

1) **Расстояние на языке корректирования**

   Один способ состоит в определении расстояния кода с $k = 0$ с помощью его свойств, корректирующих ошибки. Однако это ведет к выводу, что расстояние всегда равно n, так как закодированное состояние *известно*. «Восстановлением» является просто приготовление состояния заново! Эта интерпретация не выглядит достаточно ясной и, фактически, является не лучшим способом думать о том, что здесь происходит.

2) **Расстояние на языке детектирования**

   Физически более мотивированным понятием расстояния является то, которое количественно определяет, сколько из n физических кубитов взаимодействовало с окружением если было приготовлено k логических кубитов. Эта интерпретация фокусируется скорее на *детектировании*, чем на коррекции. Тогда расстояние может быть разумно определено как число, превышающее на единицу максимальное количество детектируемых взаимодействий кубитов ($d = t+1$). Мы можем использовать это определение расстояния, чтобы поставить разумный вопрос о декогерентизирующей силе окружения и экспериментально ответить на него, используя коды с $k = 0$.

   Напомним, что ошибка является детектируемой, если и только если она антикоммутирует с некоторым генератором стабилизатора. Одна-

ко для кодов с $k = 0$ стабилизатору принадлежат только элементы, коммутирующие со *всеми* его же элементами. По этой причине максимальное количество детектируемых ошибок равно минимальному из весов нетривиальных элементов стабилизатора. Это позволяет найти расстояние кода (во второй из приведенных трактовок), лишь посмотрев на его стабилизатор. Воспользуемся этим результатом, чтобы найти расстояние кода, предложенного в части **(а)** этой задачи.

Будем использовать определение (2), чтобы найти расстояние кода $[[6, 0, 4]]$. Поскольку его стабилизатор достаточно мал, мы можем обследовать его явно, чтобы найти элемент с минимальным весом. Его стабилизатором является

$$\begin{array}{rr}
\mathbf{1XZZX1}, & \mathbf{ZY11YZ}, \\
-\mathbf{1YXXY1}, & -\mathbf{ZXYYXZ}, \\
-\mathbf{1ZYYZ1}, & -\mathbf{Z1XX1Z}, \\
\mathbf{111111}, & \mathbf{ZZZZZZ}, \\
 & \\
\mathbf{X1YY1X}, & \mathbf{YZXXZY}, \\
-\mathbf{XZ11ZX}, & -\mathbf{YX11XY}, \\
-\mathbf{XYZZYX}, & -\mathbf{Y1ZZ1Y}, \\
\mathbf{XXXXXX}, & \mathbf{YYYYYY},
\end{array}$$

плюс перестановки.

Непосредственная проверка показывает, что минимальный вес элементов стабилизатора равен четырем.

Замечания о вырождении

В применении к коду с $k = 0$ вырождение также является не четко определенным понятием. Обычно мы говорим, что код вырожден, если минимальный вес элемента стабилизатора меньше d. Но как мы видели выше, для кодов с $k = 0$ расстояние d *определяется* как минимальный вес элементов стабилизатора. В этом смысле все коды с $k = 0$ суть невырожденные.

Более физическим способом определения вырождения кода с $k = 0$ является описание его как вырожденного, если существует две различные детектируемые ошибки, которые одинаково влияют на закодированное состояние. (Следовательно, не существует измерения синдрома, способного

различить эти две ошибки, даже несмотря на то, что можно детектировать, когда происходит одна из них.) Согласно этому определению, код $[[6,0,4]]$ наследует невырожденность своего родительского кода $[[5,1,3]]$.

Используя обе эти интерпретации, мы находим, что квантовый код $[[6,0,4]]$ *невырожден*.

c) В разделе 7.3.4 было в достаточно общем виде доказано, что вычисление следа по $d-1$ кубитам невырожденного кода с расстоянием d дает матрицу плотности, пропорциональную единице. Следовательно,

$$\boldsymbol{\rho}^{(3)} = \frac{1}{8}\mathbf{1}.$$

7.4. Кодовые слова и нелокальность

a) С помощью преобразования $\bar{\mathbf{Z}} \to \bar{\mathbf{Z}}\mathbf{M}_1\mathbf{M}_3$ (возможны и другие преобразования) мы можем перейти от представления $\bar{\mathbf{Z}} = \mathbf{ZZZZZ}$ к представлению с весом три:

$$\begin{aligned}
\bar{\mathbf{Z}} &= \mathbf{ZZZZZ},\\
\mathbf{M}_1 &= \mathbf{XZZX1},\\
\mathbf{M}_3 &= \mathbf{X1XZZ},\\
\bar{\mathbf{Z}}' &= -\mathbf{Z1XX1}.
\end{aligned}$$

Следует быть внимательным при перемножении в четвертом столбце: правильный общий знак наверняка получится при разбиении умножения на два шага $\mathbf{ZX} = \mathbf{Y}$, $\mathbf{YZ} = -\mathbf{X}$.

b) Теоретик, занимающийся скрытыми переменными, хотел бы знать результат измерения $\mathbf{ZZZZZ}$ без его фактического выполнения. Скорее, он (или она) хотел бы *сделать вывод* о значении этой наблюдаемой, используя измерение некоторой из ее подсистем и знание некоторого глобального свойства состояния, то есть примерно в том же духе, как и в мысленном ЭПР-эксперименте, в начале которого известно, что две частицы имеют суммарный спин, равный нулю, а затем предпринимается попытка сделать вывод о значении $\boldsymbol{\sigma}_x^{(1)}$ (или $\boldsymbol{\sigma}_x^{(2)}$) по результату измерения $\boldsymbol{\sigma}_x^{(2)}$ (или $\boldsymbol{\sigma}_x^{(1)}$).

Рассмотрим систему, первоначально приготовленную как общее собственное пространство найденных в части **(a)** пяти циклически связанных

наблюдаемых

$$\begin{aligned}\bar{\mathbf{Z}}_0 &= -\mathbf{Z1XX1},\\ \bar{\mathbf{Z}}_1 &= -\mathbf{1Z1XX},\\ \bar{\mathbf{Z}}_2 &= -\mathbf{X1Z1X},\\ \bar{\mathbf{Z}}_3 &= -\mathbf{XX1Z1},\\ \bar{\mathbf{Z}}_4 &= -\mathbf{1XX1Z}.\end{aligned}$$

Теоретик, занимающийся скрытыми переменными, замечает, что о собственном значении $\mathbf{Z}$ на i-ом кубите можно сделать вывод, измеряя $\mathbf{X}$ на кубитах $(i+2) \bmod 5$ и $(i+3) \bmod 5$ и зная (глобальное) собственное число m_i наблюдаемой $\bar{\mathbf{Z}}_i$. Он (или она) доказывает, что, измеряя $\mathbf{XXXXX}$, можно сделать вывод о собственном значении $\mathbf{ZZZZZ}$, фактически не измеряя его. Следовательно, предсказание состоит в том, что собственное значение $\bar{z}$ наблюдаемой $\mathbf{ZZZZZ}$ связано с собственными значениями x_i наблюдаемой $\mathbf{X}_{(i)}$ и с m_i соотношением

$$\begin{aligned}\bar{z} &= \left(\frac{-m_0}{x_2x_3}\right)\left(\frac{-m_1}{x_3x_4}\right)\left(\frac{-m_2}{x_4x_0}\right)\left(\frac{-m_3}{x_0x_1}\right)\left(\frac{-m_4}{x_1x_2}\right) =\\ &= -(m_0m_1m_2m_3m_4)\left(x_0^2x_1^2x_2^2x_3^2x_4^2\right)^{-1} =\\ &= -(m_0m_1m_2m_3m_4).\end{aligned}$$

Для состояния $|\bar{0}\rangle$ $m_0 = m_1 = m_2 = m_3 = m_4 = +1$ так, что предсказывается $\bar{z} = -1$. Однако это находится в прямом противоречии с квантовомеханическим результатом, который предсказывает, что $\mathbf{ZZZZZ}|\bar{0}\rangle = +1 \cdot |\bar{0}\rangle$.

c) Эйнштейн сказал бы, что приведенное выше доказательство устанавливает экспериментально проверяемое различие между двумя эпистемиологическими точками зрения, поддерживаемыми теорией скрытых переменных и квантовой механикой соответственно. Он мог бы добавить, что его совместный с Розеном и Подольским первоначальный пример демонстрирует то же самое различие, но гораздо проще для понимания.

7.5. Обобщенный код Шора

a) Задача лучше решается на словах, чем в громоздкой записи. Концептуально, стабилизирующими операторами являются $m-1$ операторов $\mathbf{ZZ}$,

действующих на ближайшие соседние кубиты внутри каждого из m блоков, плюс $m-1$ операторов $\mathbf{XX}\ldots\mathbf{X}$, действующих на каждую пару ближайших соседних блоков. Закодированные операции представляют собой $\bar{\mathbf{X}}$, который инвертирует закодированный бит, обращая фазы каждого из блоков (используя один оператор $\mathbf{Z}$ на каждый блок), и $\bar{\mathbf{Z}}$, который обращает закодированную фазу, инвертируя все биты внутри блока (используя $\mathbf{X}^{\otimes m}$). В громоздких обозначениях мы можем использовать двойной индекс кубитов: первый индекс помечает, в каком блоке находится кубит, а второй — позицию кубита в этом блоке. Принимая эти обозначения, мы имеем

$$\begin{aligned}
\mathbf{M}_{i,j}^{(z)} &= \mathbf{Z}_{i,j}\mathbf{Z}_{i,j+1}, \quad i=1,\ldots,m, \quad j=1,\ldots,m-1,\\
\mathbf{M}_{i}^{(x)} &= \mathbf{X}_{i,1}\cdots\mathbf{X}_{i,m}\mathbf{X}_{i+1,1}\cdots\mathbf{X}_{i+1,m}, \quad i=1,\ldots,m-1,\\
\bar{\mathbf{X}} &= \mathbf{Z}_{1,1}\cdots\mathbf{Z}_{m,1},\\
\bar{\mathbf{Z}} &= \mathbf{X}_{1,1}\cdots\mathbf{X}_{1,m}.
\end{aligned}$$

Заметим, что, как и ожидалось, существует $m(m-1)+m-1=m^2-1=$ $=n-k$ генераторов стабилизатора.

b) Закодированный оператор минимального веса ($\bar{\mathbf{Z}}$) имеет вес m, так что расстояние кода равно m. Непосредственной проверкой можно убедиться в том, что не может быть других операций с более низким весом.

c) 1) Ошибка $\bar{\mathbf{X}}$ требует, чтобы более чем в половине блоков произошло обращение фазы (с помощью оператора $\mathbf{Z}$). Событие, которое выполняет это в главном порядке по p, состоит в том, что в отдельных блоках появляется *точно* $(m+1)/2$ ошибок $\mathbf{Z}$. Вероятность этого события равна

$$\begin{aligned}
P_x &= \left(\begin{array}{c} m \\ \frac{m+1}{2} \end{array}\right)^{\text{блоки}} \left(\begin{array}{c} m \\ 1 \end{array}\right)^{\text{кубиты/блоки}} {}^{(m+1)/2\ \text{блоки}}\\
&\quad \times \left(\frac{p}{3}[\text{для ошибки } \mathbf{Z}] + \frac{p}{3}[\text{для ошибки } \mathbf{X}]\right)^{(m+1)/2\ \text{ошибки}}\\
&= \left(\begin{array}{c} m \\ \frac{m+1}{2} \end{array}\right)\left(\begin{array}{c} m \\ 1 \end{array}\right)^{(m+1)/2}\left(\frac{2p}{3}\right)^{(m+1)/2}.
\end{aligned}$$

2) Ошибка $\bar{\mathbf{Z}}$ требует, чтобы более чем в половине в нечетном количестве блоков произошло инвертирование их битов (с помощью операторов $\mathbf{X}$). Событие, которое выполняет это в главном порядке по p, состоит в том, что в одном блоке появляется *точно* $(m+1)/2$ ошибок $\mathbf{X}$. Вероятность этого события равна

$$P_z = \underbrace{\binom{m}{1}}_{\text{блоки}} \underbrace{\binom{m}{\frac{m+1}{2}}}_{\text{кубиты/блоки}}^{\text{1 блок}} \times$$

$$\times \underbrace{\left(\frac{p}{3}[\text{для ошибки } \mathbf{Z}] + \frac{p}{3}[\text{для ошибки } \mathbf{X}]\right)^{(m+1)/2}}_{\text{ошибки}} =$$

$$= \binom{m}{1}\binom{m}{\frac{m+1}{2}}\left(\frac{2p}{3}\right)^{(m+1)/2}.$$

Заметим, что $P_x = m^{(m-1)/2} P_z$.

d) При больших m мы можем воспользоваться формулой Стирлинга

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

чтобы упростить выражения, найденные в части **(c)**. В этом приближении мы имеем

$$\binom{m}{\frac{m+1}{2}} \approx \binom{m}{m/2} =$$

$$= \frac{m!}{\frac{m}{2}!\frac{m}{2}!} \approx$$

$$\approx \frac{\sqrt{2\pi m}\left(\frac{m}{e}\right)^m}{2\pi\frac{m}{2}\left(\frac{m}{2e}\right)^{m/2}\left(\frac{m}{2e}\right)^{m/2}} =$$

$$= \sqrt{\frac{2}{\pi m}}2^m.$$

Следовательно, вероятность P_z приближенно равна

$$P_z \approx \sqrt{\frac{2m}{\pi}} 2^m \left(\frac{2p}{3}\right)^{(m+1)/2} =$$
$$= \sqrt{\frac{m}{2\pi}} 4^{(m+1)/2} \left(\frac{2p}{3}\right)^{(m+1)/2} =$$
$$= \sqrt{\frac{m}{2\pi}} \left(\frac{8p}{3}\right)^{(m+1)/2},$$

аналогично, вероятность P_x приближенно равна

$$P_x \approx \sqrt{\frac{m}{2\pi}} m^{(m-1)/2} \left(\frac{8p}{3}\right)^{(m+1)/2} =$$
$$= \sqrt{\frac{1}{2\pi m}} \left(\frac{8mp}{3}\right)^{(m+1)/2}.$$

Чтобы обеспечить хорошую защиту против ошибок обращения фазы при $m \to \infty$, нам необходимо, чтобы $p < 3/8$, так как тогда $P_z \to 0$. Чтобы обеспечить хорошую защиту против ошибок инвертирования бита при $m \to \infty$, нам необходимо, чтобы $p < 3/(8m)$, так как тогда $P_x \to 0$. Однако при $m \to \infty$ это требование эквивалентно требованию $p = 0$. Следовательно, это асимптотическое семейство кодов не может обеспечить надежную защиту против ошибок инвертирования бита, хотя может защитить от ошибок обращения фазы, если $p < 3/8$.

7.6. Кодирующие схемы

На лекциях обсуждение вопросов, касающихся кодирующих схем, было довольно кратким, поэтому здесь я его расширю. Прежде чем объяснить, как они работают, я представлю универсальный алгоритм конструирования кодирующей схемы данного стабилизирующего кода. Алгоритм естественным образом делится на три фазы.

Фаза I

1) Выразим стабилизатор на языке двоичного векторного пространства как $(H_x|H_z)$. Выполним процедуру исключения Гаусса–Жордана, так чтобы H_x начинался с единичной матрицы ранга r.

2) Для каждой закодированной операции $\bar{\mathbf{X}}_i$ найдем представление $\bar{\mathbf{X}}_i = \tilde{\mathbf{Z}}^{\otimes r}\tilde{\mathbf{X}}^{\otimes(n-k+r)}\mathbf{X}_{(n-k+i)}$, где $\tilde{\mathbf{Z}}$ равно $\mathbf{Z}$ или $\mathbf{1}$, а $\tilde{\mathbf{X}}$ равно $\mathbf{X}$ или $\mathbf{1}$. Такое представление всегда может быть найдено.[1] В литературе это представление известно как «стандартная форма» операторов $\bar{\mathbf{X}}_i$.

3) Проведем n горизонтальных контрольных линий схемы; начиная с верхней, снабдим их метками $1, \ldots, n$. Пусть k кодируемых кубитов занимают k нижних контрольных линий. Пусть $|0\rangle$ занимают остальные контрольные линии.

4) Изобразим каждый из операторов $\bar{\mathbf{X}}_i$ из пункта 2), последовательно действующих на кубиты, но заменим каждый сомножитель $\tilde{\mathbf{Z}}$ единицей, а $\mathbf{X}_{(n-k+i)}$ — контрольным узлом, управляющим остальной частью оператора $\bar{\mathbf{X}}_i$.

Фаза II

Изобразим поворот Адамара на каждой из первых r контрольных линий.

Фаза III

Изобразим первые r генераторов стабилизатора $\mathbf{M}_1, \ldots, \mathbf{M}_r$ [а именно, первые r строк матрицы $(H_x|H_z)$, приведенной с помощью выполненной в пункте 1 фазы I процедуры Гаусса–Жордана], последовательно действующих на кубиты, но в каждом $\mathbf{M}_i$ заменим множитель $\mathbf{X}_{(i)}$ контрольным узлом, управляющим остальной частью генератора $\mathbf{M}_i$. Более того, заменим в $\mathbf{M}_i$ на $\mathbf{1}$ каждый сомножитель $\mathbf{Z}_j$ с номером $j > i$. Если в $\mathbf{M}_i$ присутствует $\mathbf{Y}_{(i)}$, а не $\mathbf{X}_{(i)}$, то перед применением оператора контролируемое $\mathbf{M}_i$ предварительно умножим контрольный узел на оператор $\mathbf{Z}$.

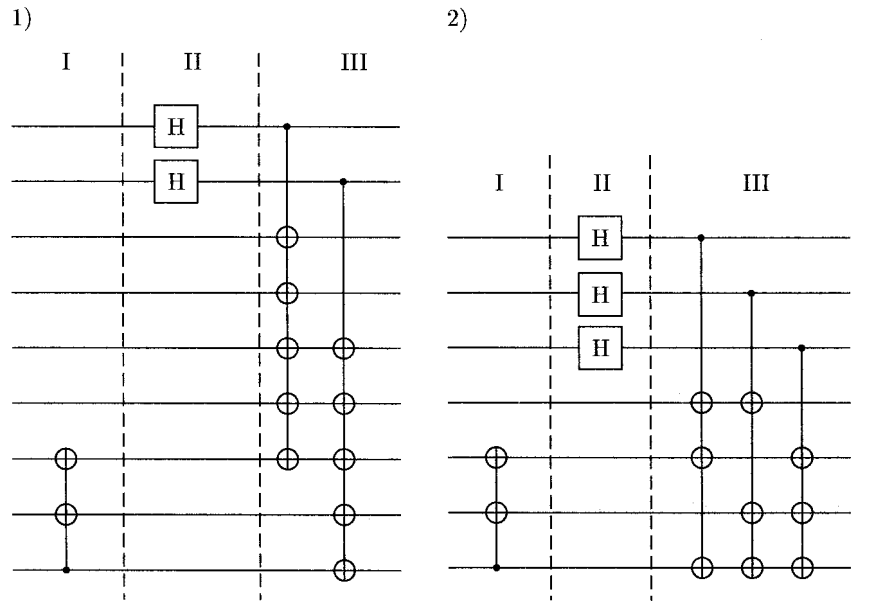
Фазы I и II всегда могут выполняться параллельно, несмотря на то, что концептуально они различны. Заметим, что фаза II полностью исчезает при $n - k = r$, как это имеет место в рассмотренном на лекции случае кода $[[5, 1, 3]]$. Однако этот этап нельзя забывать в случае кодирующих схем для кодов $[[9, 1, 3]]$ и $[[7, 1, 3]]$.

[1] Подробности можно найти в диссертации Д. Готтесмана *Stabilizier Codes and Quantum Error Correction*, quant-ph/9705052. Я не хочу здесь доказывать это утверждение.

Анализ алгоритма кодирования

Анализ того, почему это работает, увел бы нас далеко в сторону и сделал бы еще длиннее это уже и без того длинное решение. Интересующегося читателя я отсылаю к диссертации Даниэля Готтесмана *Stabilizier Codes and Quantum Error Correction*, доступной на домашней странице этого курса. Однако некоторые диаграммы в диссертации неправильны. В частности, в них не учитывается тот факт, что всякий раз, когда $j > i$, в $\mathbf{M}_i$ необходимо заменять $\mathbf{Z}_j$ на $\mathbf{1}$. Для предлагаемых здесь задач этот нюанс не важен, так как вся контролирующая логика осуществляется только с помощью операторов $\mathbf{X}$.

Применяя этот алгоритм к кодам Шора и Стина, мы генерируем кодировщики:



7.7. Укорачивание квантового кода

а) Допустим, что на последний кубит нетривиально действует более двух генераторов. Выберем один из них, назовем его для определенности $\mathbf{M}_1$ и выполним на последнем кубите локальное унитарное преобразование так, чтобы этот генератор действовал здесь как $\mathbf{X}$. Затем умножим на $\mathbf{M}_1$ каж-

дый другой генератор, действующий на последний кубит как **X** или **Y**. Если после этого шага не осталось других генераторов с нетривиальным носителем на последнем кубите, то мы выполнили свою работу. В противном случае, в соответствии с предыдущим действием $\mathbf{M}_1$, все такие оставшиеся генераторы должны действовать на последний кубит как **Z**. Выберем один из этих генераторов и назовем его $\mathbf{M}_2$. Умножим на $\mathbf{M}_2$ каждый другой генератор, действующий на последний кубит как **Z**. Получающийся в результате стабилизатор имеет самое большее два (а именно, $\mathbf{M}_1$ и $\mathbf{M}_2$) генератора, нетривиально действующих на последний кубит.

b) Это утверждение трудно доказать, поскольку оно неверно. В качестве контрпримера рассмотрим «выколотый»[1] код Шора $[[9,1,3]]$. Он не является кодом $[[8,2,2]]$, как это утверждается в условии. В действительности, он представляет собой код $[[8,2,1]]$, так как минимальный вес выкалываемой закодированной операции равен единице:

$$
\begin{array}{l}
\mathbf{M}_1 = \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1\ 1\ 1\ 1 \\
\mathbf{M}_2 = 1\ \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1\ 1\ 1\ 1 \\
\mathbf{M}_3 = 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1\ 1 \\
\mathbf{M}_4 = 1\ 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1 \\
\mathbf{M}_5 = 1\ 1\ 1\ 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z}\ 1 \\
\mathbf{M}_6 = 1\ 1\ 1\ 1\ 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z} \\
\mathbf{M}_7 = \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ 1\ 1\ 1 \\
\mathbf{M}_8 = 1\ 1\ 1\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X} \\
\bar{\mathbf{X}} = \mathbf{Z}\ 1\ 1\ \mathbf{Z}\ 1\ 1\ \mathbf{Z}\ 1\ 1 \\
\bar{\mathbf{Z}} = \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ 1\ 1\ 1\ 1\ 1\ 1
\end{array}
\quad \Rightarrow \quad
\begin{array}{l}
\mathbf{M}'_1 = \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1\ 1\ 1\ 1 \\
\mathbf{M}'_2 = 1\ \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1\ 1\ 1 \\
\mathbf{M}'_3 = 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z}\ 1\ 1\ 1 \\
\mathbf{M}'_4 = 1\ 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z}\ 1\ 1 \\
\mathbf{M}'_5 = 1\ 1\ 1\ 1\ 1\ 1\ \mathbf{Z}\ \mathbf{Z} \\
\mathbf{M}'_6 = \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ 1\ 1 \\
\bar{\mathbf{X}}'_1 = \mathbf{Z}\ 1\ 1\ \mathbf{Z}\ 1\ 1\ \mathbf{Z}\ 1 \\
\bar{\mathbf{Z}}'_1 = \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ 1\ 1\ 1\ 1\ 1 \\
\bar{\mathbf{X}}'_2 = 1\ 1\ 1\ 1\ 1\ 1\ 1\ \mathbf{Z} \\
\bar{\mathbf{Z}}'_2 = 1\ 1\ 1\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}
\end{array}
$$

Результат правилен, если ограничиться выкалыванием невырожденных кодов. Я докажу это от противного:

Утверждение: Выколотый код $[[n,k,d]]$ представляет собой код $[[n-1,k+1,d^*]]$, где $d^* \geqslant d-1$.

Доказательство: Пусть C является кодом $[[n,k,d]]$. Согласно результатам части **(a)** мы знаем, что выкалывание C дает в результате код C' $[[n-1,k+1,d^*]]$. Пусть d^* является расстоянием кода C'; предположим, что $d^* < d-1$. Покажем, что это предположение ведет к противоречию (и, следовательно, докажем утверждение).

Согласно предположению, существует оператор $\mathbf{E}'$ с весом d^*, коммутирующий со всеми генераторами стабилизатора кода C'. Оператор $\mathbf{E}'$

[1] Эта терминология позаимствована из теории классических кодов коррекции ошибок. Эту задачу лучше было озаглавить «Выкалывание квантового кода». Процесс укорачивания кода представляет собой другую операцию.

можно расширить до $\mathbf{E}$, который действует на C и имеет вес, не превышающий $d^* + 1$. Чтобы выполнить это, заметим сначала, что результаты части **(a)** говорят нам о том, что в самом общем виде генераторы C могут быть записаны как

$$\begin{aligned}\mathbf{M}_i &= \mathbf{M}'_i \otimes \mathbf{1},\\ \mathbf{M}_{n-k-1} &= \bar{\mathbf{X}}'_{k+1} \otimes \mathbf{X},\\ \mathbf{M}_{n-k} &= \bar{\mathbf{Z}}'_{k+1} \otimes \mathbf{Z},\end{aligned}$$

где $\mathbf{M}'_i$ $(i = 1, \ldots, n-k-2)$ — генераторы стабилизатора C', а $\bar{\mathbf{X}}'_{k+1}$ и $\bar{\mathbf{Z}}'_{k+1}$ — закодированные $\mathbf{X}$ и $\mathbf{Z}$ операторы, действующие на логический кубит $k+1$.

Построим $\mathbf{E}$, определяя его в зависимости от того, как оператор $\mathbf{E}' \otimes \mathbf{1}$ коммутирует с генераторами $\mathbf{M}_{n-k-1}$ и $\mathbf{M}_{n-k}$, следующим способом[1] (см. таблицу).

$[\mathbf{E}' \otimes \mathbf{1},\ \mathbf{M}_{n-k-1}]$	$[\mathbf{E}' \otimes \mathbf{1},\ \mathbf{M}_{n-k}]$	$\mathbf{E}$	Вес
$+1$	$+1$	$\mathbf{E}' \otimes \mathbf{1}$	d^*
$+1$	-1	$\mathbf{E}' \otimes \mathbf{X}$	$d^* + 1$
-1	$+1$	$\mathbf{E}' \otimes \mathbf{Z}$	$d^* + 1$
-1	-1	$\mathbf{E}' \otimes \mathbf{Y}$	$d^* + 1$

Эта конструкция гарантирует, что $\mathbf{E}$ коммутирует со всеми генераторами C' и имеет вес, меньший или равный $d^* + 1$. Но подождите! Так как C является кодом с расстоянием d, все ошибки с весом, не превосходящим $d-1$, или антикоммутируют с некоторым генератором стабилизатора C, или сами содержатся в C. Однако мы только что показали, что

$$\begin{aligned}\mathrm{wt}(\mathbf{E}) &\leqslant d^* - 1 <\\ &< d - 2 \not\geqslant\\ &\not\geqslant d - 1,\end{aligned}$$

и тем не менее $\mathbf{E}$ коммутирует со всеми генераторами стабилизатора C! Тогда $\mathbf{E}$ на самом деле *должен принадлежать* стабилизатору C, что имело бы место, если бы C был вырожденным. Но мы взяли C невырожденным, то есть все элементы его стабилизатора имеют веса, превосходящие d. Следовательно, полученное выше неравенство представляет искомое нами противоречие. □

c) Выкалывание кода $[[5, 1, 3]]$ дает код $[[4, 2, 2]]$, о чем говорилось на лекциях. Сначала мы перегруппируем элементы нормализатора кода $[[5, 1, 3]]$,

[1] Значения $+1(-1)$ в таблице означают, что операторы коммутируют (антикоммутируют).

при необходимости умножая генераторы на $\mathbf{M}_1$ или на $\mathbf{M}_2$,

$$\begin{array}{l} \mathbf{M}_1 = \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X}\ 1 \\ \mathbf{M}_2 = 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X} \\ \mathbf{M}_3 = \mathbf{X}\ 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z} \\ \mathbf{M}_4 = \mathbf{Z}\ \mathbf{X}\ 1\ \mathbf{X}\ \mathbf{Z} \\ \bar{\mathbf{X}} = \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X} \\ \bar{\mathbf{Z}} = \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z} \end{array} \quad \Rightarrow \quad \begin{array}{l} \mathbf{M}_1 = \ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X}\ 1 \\ \mathbf{M}_2 = -\mathbf{Y}\ \mathbf{X}\ \mathbf{X}\ \mathbf{Y}\ 1 \\ \mathbf{M}_3 = \ 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X} \\ \mathbf{M}_4 = \ \mathbf{Z}\ \mathbf{X}\ 1\ \mathbf{X}\ \mathbf{Z} \\ \bar{\mathbf{X}} = \ \mathbf{X}\ 1\ \mathbf{Y}\ \mathbf{Y}\ 1 \\ \bar{\mathbf{Z}} = \ 1\ \mathbf{Y}\ \mathbf{Z}\ \mathbf{Y}\ 1 \end{array}$$

Затем мы обрезаем последний бит и для удобства устанавливаем общую для всех генераторов фазу $+1$ (результирующий код изоморфен исходному с точностью до произвольной общей фазы). Генераторы $\mathbf{M}_3$ и $\mathbf{M}_4$ превращаются в закодированные операторы $\bar{\mathbf{Z}}$ и $\bar{\mathbf{X}}$ для дополнительного закодированного кубита:

$$\begin{array}{l} \mathbf{M}_1 = \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X}\ 1 \\ \mathbf{M}_2 = \mathbf{Y}\ \mathbf{X}\ \mathbf{X}\ \mathbf{Y}\ 1 \\ \mathbf{M}_3 = 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X} \\ \mathbf{M}_4 = \mathbf{Z}\ \mathbf{X}\ 1\ \mathbf{X}\ \mathbf{Z} \\ \bar{\mathbf{X}} = \mathbf{X}\ 1\ \mathbf{Y}\ \mathbf{Y}\ 1 \\ \bar{\mathbf{Z}} = 1\ \mathbf{Y}\ \mathbf{Z}\ \mathbf{Y}\ 1 \end{array} \quad \text{выкалывание} \quad \longrightarrow \quad \begin{array}{l} \mathbf{M}_1 = \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X} \\ \mathbf{M}_2 = \mathbf{Y}\ \mathbf{X}\ \mathbf{X}\ \mathbf{Y} \\ \bar{\mathbf{X}}_1 = 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z} \\ \bar{\mathbf{Z}}_1 = \mathbf{Z}\ \mathbf{X}\ 1\ \mathbf{X} \\ \bar{\mathbf{X}}_2 = \mathbf{X}\ 1\ \mathbf{Y}\ \mathbf{Y} \\ \bar{\mathbf{Z}}_2 = 1\ \mathbf{Y}\ \mathbf{Z}\ \mathbf{Y} \end{array}$$

Мы вправе изменить базисы некоторых кубитов, то есть применить одновременные ЛУПы $\mathbf{X} \to \mathbf{Z} \to \mathbf{Y} \to \mathbf{X}$ к первому и последнему кубитам, чтобы получить эквивалентный стабилизатор:

$$\begin{array}{l} \mathbf{M}_1 = \mathbf{X}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{X} \\ \mathbf{M}_2 = \mathbf{Y}\ \mathbf{X}\ \mathbf{X}\ \mathbf{Y} \\ \bar{\mathbf{X}}_1 = 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Z} \\ \bar{\mathbf{Z}}_1 = \mathbf{Z}\ \mathbf{X}\ 1\ \mathbf{X} \\ \bar{\mathbf{X}}_2 = \mathbf{X}\ 1\ \mathbf{Y}\ \mathbf{Y} \\ \bar{\mathbf{Z}}_2 = 1\ \mathbf{Y}\ \mathbf{Z}\ \mathbf{Y} \end{array} \quad \text{ЛУПы} \quad \Rightarrow \quad \begin{array}{l} \mathbf{M}_1 = \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z} \\ \mathbf{M}_2 = \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X} \\ \bar{\mathbf{X}}_1 = 1\ \mathbf{X}\ \mathbf{Z}\ \mathbf{Y} \\ \bar{\mathbf{Z}}_1 = \mathbf{Y}\ \mathbf{X}\ 1\ \mathbf{Z} \\ \bar{\mathbf{X}}_2 = \mathbf{Z}\ 1\ \mathbf{Y}\ \mathbf{X} \\ \bar{\mathbf{Z}}_2 = 1\ \mathbf{Y}\ \mathbf{Z}\ \mathbf{X} \end{array}$$

Заменяя закодированные операторы их двойниками с весом два, мы можем сделать более очевидным, что расстояние проколотого кода равно двум. Одним из способов выполнения этого является замена:

$$\begin{array}{lll} \bar{\mathbf{X}}_1 & \to & \bar{\mathbf{X}}_1\bar{\mathbf{Z}}_2\bar{\mathbf{X}}_2\mathbf{M}_1 \quad = -11\mathbf{XX} \\ \bar{\mathbf{Z}}_1 & \to & \bar{\mathbf{X}}_1\bar{\mathbf{Z}}_2 \quad = -1\mathbf{Z}1\mathbf{Z} \\ \bar{\mathbf{X}}_2 & \to & \bar{\mathbf{X}}_2\bar{\mathbf{Z}}_1\bar{\mathbf{X}}_1 \quad = -\mathbf{X}1\mathbf{X}1 \\ \bar{\mathbf{Z}}_2 & \to & \bar{\mathbf{X}}_2\bar{\mathbf{Z}}_1\mathbf{M}_2 \quad = -11\mathbf{ZZ} \end{array}$$

Заметим, что эти преобразования сохраняют коммутационные соотношения между логическими операциями. В качестве заключительного шага мы

можем удалить общие фазы в нормализаторе:

$$\begin{aligned}
\mathbf{M}_1 &= \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z}\ \mathbf{Z} \\
\mathbf{M}_2 &= \mathbf{X}\ \mathbf{X}\ \mathbf{X}\ \mathbf{X} \\
\bar{\mathbf{X}}_1 &= \mathbf{1}\ \mathbf{1}\ \mathbf{X}\ \mathbf{X} \\
\bar{\mathbf{Z}}_1 &= \mathbf{1}\ \mathbf{Z}\ \mathbf{1}\ \mathbf{Z} \\
\bar{\mathbf{X}}_2 &= \mathbf{X}\ \mathbf{1}\ \mathbf{X}\ \mathbf{1} \\
\bar{\mathbf{Z}}_2 &= \mathbf{1}\ \mathbf{1}\ \mathbf{Z}\ \mathbf{Z}
\end{aligned}$$

Используя это представление закодированных операций, чтобы выбрать базис для кодовых слов, мы находим, что кодовыми словами кода $[[4,2,2]]$ являются

$$|\bar{0}\bar{0}\rangle = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle),$$

$$|\bar{0}\bar{1}\rangle = \frac{1}{\sqrt{2}}(|1010\rangle + |0101\rangle),$$

$$|\bar{1}\bar{0}\rangle = \frac{1}{\sqrt{2}}(|0011\rangle + |1100\rangle),$$

$$|\bar{1}\bar{1}\rangle = \frac{1}{\sqrt{2}}(|1001\rangle + |0110\rangle).$$

Может быть, непосредственно этот код и не знаком, но его структура — да. Каждое кодовое слово является суперпозицией строк четного веса и длины четыре. Из классического кодирования Рида–Маллера нам известно, что $R(m-1,m)$ является пространством всех строк четного веса и длины 2^m. Следовательно, вышеприведенные квантовые кодовые слова выглядят, как смежные классы пространства $R(1,2)$ и другого, входящего в него, кода. Это подозрение подтверждается, если мы применим конструкцию КШС к $R(1,2)$ и к дуальному к нему $R(0,2)$. (Вспомним, что $R^{\perp}(r,m) = R(m-r-1,m)$.) Поскольку $R(0,2) \subseteq R(1,2)$ [так как $R(r,m)$ представляет полиномы степени r над $\mathbb{F}_m$], конструкция КШС справедлива и дает квантовый код с параметрами $[[4,2,2]]$. Похоже, что этот код хорошо описывается как «квантовый код Рида–Маллера».

7.8. Коды для кудитов

а)

1) Да, $\{\mathbf{E}_{r,s}\}$ образуют базис для $\mathrm{U}(d)$. Чтобы доказать это, мы должны показать, что они являются линейной оболочкой $\mathrm{U}(d)$. Достаточно показать, что базис $\{|a\rangle\langle b|\}$ для $\mathrm{U}(d)$ может быть разложен по $\{\mathbf{E}_{r,s}\}$, так как оба множества операторов имеют размер $d^2 = \dim \mathrm{U}(d)$.

Заметим сначала, что $\mathbf{E}_{r,s}$ может быть записан в базисе $\{|a\rangle\langle b|\}$ как[1]

$$\mathbf{E}_{r,s} = \sum_{j=0}^{d-1} \omega^{js} |j+r\rangle\langle j|.$$

Базисный элемент $|a\rangle\langle b|$ имеет разложение

$$|a\rangle\langle b| = \sum_{r,s=0}^{d-1} c_{rs}\mathbf{E}_{r,s},$$

коэффициенты которого даются выражением

$$\begin{aligned} c_{rs} &= \frac{1}{d}\,\mathrm{tr}\left(\mathbf{E}_{r,s}^{\dagger}|a\rangle\langle b|\right) = \\ &= \frac{1}{d}\,\mathrm{tr}\left(\sum_{j=0}^{d-1}\omega^{-js}|j\rangle\langle j+r|a\rangle\langle b|\right) = \\ &= \frac{1}{d}\sum_{i,j=0}^{d-1}\omega^{-js}\langle i|j\rangle\langle j+r|a\rangle\langle b|i\rangle = \\ &= \frac{1}{d}\omega^{-bs}\delta_{a,b+r}. \end{aligned}$$

Чтобы убедиться в том, что это «разложение Фурье» правильно, заметим, что

$$\begin{aligned} \left\langle b\left|\sum_{r,s=0}^{d-1} c_{rs}\mathbf{E}_{r,s}\right|a\right\rangle &= \frac{1}{d}\sum_{r,s,j=0}^{d-1}\omega^{-bs+js}\delta_{a,b+r}\langle b|j+r\rangle\langle j|a\rangle = \\ &= \frac{1}{d}\sum_{r,s=0}^{d-1}\omega^{(a-b)s}\delta_{a,b+r}\delta_{b,a+r} = \\ &= \frac{1}{d}\sum_{r,s=0}^{d-1}\omega^{(a-b)s}\delta_{r,a-b}\delta_{r,b-a} = \\ &= \frac{1}{d}\sum_{s=0}^{d-1}(1)^s = \\ &= 1. \end{aligned}$$

[1] С этого момента все дополнительные сложения и вычитания в решениях интерпретируются по модулю d, где d — размерность рассматриваемого кудита.

2) Да, операторы $\{\mathbf{E}_{r,s}\}$ унитарны. Так как $\mathbf{X}$ и $\mathbf{Z}$ сохраняют внутреннее произведение, то таким же свойством обладает $\mathbf{E}_{r,s} = \mathbf{X}^r\mathbf{Z}^s$:

$$\begin{aligned}\langle i|\mathbf{X}^\dagger\mathbf{X}|j\rangle &= \langle i+1|j+1\rangle = \\ &= \delta_{ij}, \\ \langle i|\mathbf{Z}^\dagger\mathbf{Z}|j\rangle &= \langle i|\omega^{-is}\omega^{js}|j\rangle = \\ &= \omega^{(j-i)s}\langle i|j\rangle = \\ &= \delta_{ij}.\end{aligned}$$

3) След внутреннего произведения базисных элементов (которые явно использовались в пункте 1 этой задачи) равен

$$\begin{aligned}\mathrm{tr}(\mathbf{E}_{r,s}^\dagger\mathbf{E}_{t,u}) &= \mathrm{tr}(\mathbf{Z}^{-s}\mathbf{X}^{-r}\mathbf{X}^t\mathbf{Z}^u) = \\ &= \mathrm{tr}(\mathbf{X}^{t-r}\mathbf{Z}^{u-s}) = \\ &= \sum_{j=0}^{d-1}\langle j|\mathbf{X}^{t-r}\mathbf{Z}^{u-s}|j\rangle = \\ &= \sum_{j=0}^{d-1}\omega^{(u-s)j}\langle j|j+t-r\rangle = \\ &= \delta_{tr}\sum_{j=0}^{d-1}\omega^{(u-s)j} = \\ &= \begin{cases}\delta_{tr}\cdot d, & u=s, \\ \delta_{tr}\dfrac{1-\omega^{(u-s)d}}{1-\omega^{u-s}} = 0, & u\neq s\end{cases} = \\ &= d\cdot\delta_{tr}\delta_{us}.\end{aligned}$$

b) Вычисляя действие коммутатора[1] $[\mathbf{E}_{r,s},\mathbf{E}_{t,u}] = \mathbf{E}_{r,s}\mathbf{E}_{t,u}\mathbf{E}_{r,s}^{-1}\mathbf{E}_{t,u}^{-1}$ на пробное состояние $|j\rangle$, мы находим

$$\begin{aligned}[\mathbf{E}_{r,s},\mathbf{E}_{t,u}]|j\rangle &= \mathbf{X}^r\mathbf{Z}^s\mathbf{X}^t\mathbf{Z}^u\mathbf{Z}^{-s}\mathbf{X}^{-r}\mathbf{Z}^{-u}\mathbf{X}^{-t}|j\rangle = \\ &= \omega^{-u(j-t)-s(j-t-r)+u(j-t-r)+s(j-r)}|j\rangle = \\ &= \omega^{st-ur}|j\rangle.\end{aligned}$$

[1] Это алгебраическое определение коммутатора. Когда вычисляются синдромы, мы интересуемся именно *этим*, а не групповым коммутатором $[a,b] = ab - ba$.

Следовательно, $\eta_{r,s;t,u} = \omega^{st-ur} = \omega^{(t,u)*(r,s)}$, где $*$ обозначает определенное на лекциях симплектическое внутреннее произведение.

c) Всякий раз, когда мы добавляем генератор к стабилизирующему коду для кудитов, мы сокращаем размерность его кодового подпространства на множитель, равный порядку генератора. Если d не простое число, то возможно иметь генераторы порядка d_i, являющегося делителем d, но не самим d. После $n-k$ выборов генераторов размерность кодового подпространства становится

$$\begin{aligned} D &= \frac{d^n}{\prod_{i=1}^{n-k} d_i} = \\ &= \frac{d^n}{d^{n-k} \prod_{i=1}^{n-k} \frac{1}{r_i}} = \\ &= d^k \left(\prod_{i=1}^{n-k} r_i \right) \geqslant \\ &\geqslant d^k. \end{aligned}$$

Следовательно, такие коды могут кодировать более k кудитов! Это поднимает вопрос: что же тогда здесь закодировано? Чтобы ответить на него, исследуем два кода $d = 4$. Первый код, который мы изучим, имеет $n = 1$. Его стабилизатором является

$$\mathbf{M}_1 = \mathbf{X}^2$$

размерность кодового подпространства равна $d^n/|\mathbf{M}_1| = 4^1/2 = 2$. Этот код недостаточно велик, чтобы сохранять даже один кудит. Но он достаточно велик, чтобы вместить один кубит. Нормализатор кода содержит $\mathbf{X}$ и $\mathbf{Z}^2$, которые, мы подозреваем, имеют коммутационное соотношение

$$\begin{aligned} [\mathbf{X}, \mathbf{Z}^2] = \mathbf{X}\mathbf{Z}^2\mathbf{X}^{-1}\mathbf{Z}^{-2} &= \\ = \omega^{-2} &= \\ = -1. & \end{aligned}$$

Наше подозрение подтверждается — этот код $[[1,0,1]]_4$ кодирует один кубит:

$$|\bar{0}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |2\rangle),$$

$$|\bar{1}\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |3\rangle).$$

Теперь более интересный пример. Рассмотрим код $[[5,1,3]]_4$ из задачи **7.11**. Квадраты всех генераторов образуют новый код со стабилизатором

$$\begin{array}{ccccc} \mathbf{X}^2 & \mathbf{Z}^2 & \mathbf{Z}^{-2} & \mathbf{X}^{-2} & \mathbf{1} \\ \mathbf{1} & \mathbf{X}^2 & \mathbf{Z}^2 & \mathbf{Z}^{-2} & \mathbf{X}^{-2} \\ \mathbf{X}^{-2} & \mathbf{1} & \mathbf{X}^2 & \mathbf{Z}^2 & \mathbf{Z}^{-2} \\ \mathbf{Z}^{-2} & \mathbf{X}^{-2} & \mathbf{1} & \mathbf{X}^2 & \mathbf{Z}^2 \end{array}$$

Размерность кодового подпространства равна $d^n / \prod_i |\mathbf{M}_i| = 4^5/2^4 = 4^3 = 64$. Он достаточно велик, чтобы сохранять три «кукварта», или два кукварта и два кубита, или один кукварт и четыре кубита и так далее. Какое решение принять? Оказывается, что мы добавляем к коду по одному кубиту для каждого генератора возведенного в квадрат исходного кода куквартов. Таким образом, в первом примере, где стабилизатором был просто $\mathbf{X}$, мы имели $k = 0$ куквартов. Тогда возведение в квадрат этого генератора привело к одному закодированному кубиту. В этом примере после возведения в квадрат четырех генераторов мы имеем $k = 1$ кукварт плюс четыре кубита. Доказательство этого общего факта увело бы нас даже еще дальше в сторону, но достаточно сказать, что для этого примера закодированными операциями являются:

один кудит:

$$\bar{\mathbf{X}}_4 = \mathbf{XXXXX},$$
$$\bar{\mathbf{Z}}_4 = \mathbf{ZZZZZ},$$

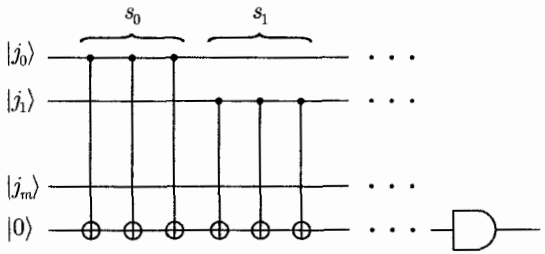
четыре кубита:

$$\bar{\mathbf{X}}_2^{(i)} = \mathbf{Z1XX1} \quad + \text{три перестановки},$$
$$\bar{\mathbf{Z}}_2^{(i)} = \mathbf{Z11ZX} \quad + \text{три перестановки}.$$

7.9. Измерение синдрома для кудита

а) Измерение наблюдаемой $\bigotimes_a \mathbf{Z}_a^{s_a}$ на $\bigotimes_a |j_a\rangle$ дает собственное значение $\omega^{\sum_a s_a j_a}$. Однако измерение $\mathbf{Z}$ на $|\sum_a s_a j_a\rangle$ также дает собственное значение $\omega^{\sum_a s_a j_a}$. Это подсказывает, что для того чтобы измерить $\bigotimes_a \mathbf{Z}_a^{s_a}$, мы должны применить s_a копий схемы SUM между каждым кубитом $|j_a\rangle$ и служебным кубитом, первоначально приготовленным в состоянии $|0\rangle$. Тогда мы можем измерить $\mathbf{Z}$ на служебном кубите, чтобы получить собствен-

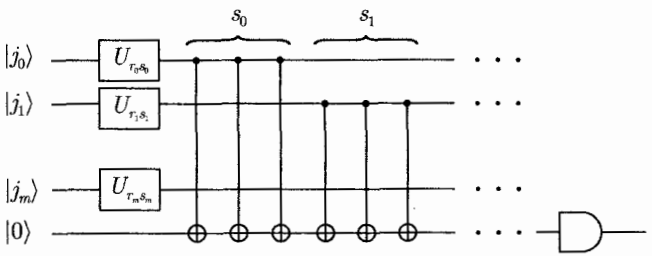
ное значение наблюдаемой $\bigotimes_a \mathbf{Z}_a^{s_a}$. Соответствующая схема имеет вид



b) Измерение наблюдаемой $\bigotimes_a \mathbf{E}_{r_a,s_a}$ на $\bigotimes_a |j_a\rangle$ дает такое же собственное значение, какое дает измерение $\bigotimes_a \mathbf{U}_{r_a,s_a} \mathbf{E}_{r_a,s_a} \mathbf{U}^\dagger_{r_a,s_a}$ на $\bigotimes_a \mathbf{U}_{r_a,s_a} |j_a\rangle$. Но поскольку (только для простых d!)

$$\bigotimes_a \mathbf{U}_{r_a,s_a} \mathbf{E}_{r_a,s_a} \mathbf{U}^\dagger_{r_a,s_a} = \mathbf{Z},$$

то все, что нам нужно сделать, — это модифицировать схему из части **(a)**, предварительно умножая каждый кубит $|j_a\rangle$ на $\mathbf{U}_{r_a,s_a}$:



7.10. Коды детектирования ошибок в кудитах

a) Да, генераторы коммутируют

$$\begin{aligned}[\mathbf{ZZZ}, \mathbf{XXX}] &= [\mathbf{Z}, \mathbf{X}]^3 = \\ &= \eta_{0,1;1,0}^3 = \\ &= 1.\end{aligned}$$

b) Расстояние этого кода равно двум. Оно равно минимальному весу приведенных ниже операторов нормализатора

$$\begin{aligned}\bar{\mathbf{X}} &= \mathbf{1}\ \mathbf{X}\ \mathbf{X}^2, \\ \bar{\mathbf{Z}} &= \mathbf{Z}^2\ \mathbf{Z}\ \mathbf{1}.\end{aligned}$$

Вместо того, чтобы доказывать, что эти представления имеют минимальный вес, равный двум, проще заметить, что расстояние кода должно быть больше единицы, так как все операторы с единичным весом антикоммутируют (то есть имеют нетривиальный синдром) по меньшей мере с одним из генераторов стабилизатора.

Достаточно составить список представленных выше операторов $\bar{\mathbf{X}}$ и $\bar{\mathbf{Z}}$, чтобы точно определить полный нормализатор, поскольку любой другой логический оператор $\bar{\mathbf{E}}_{r,s}$ может быть записан как $\bar{\mathbf{X}}^r\bar{\mathbf{Z}}^s$.

c) Ортонормированный базис для кодового подпространства можно получить, формируя сначала $|\bar{0}\rangle$, а затем необходимое количество раз применяя логические повышающие операторы $\bar{\mathbf{X}}$. Чтобы образовать $|\bar{0}\rangle$, обычно действуют на $|000\rangle$ суммой всех элементов стабилизатора. Но, поскольку для этого кода $\mathbf{ZZZ}$ действует на $|000\rangle$ как единица, нам на самом деле необходима только сумма по элементам стабилизатора, порождаемым оператором $\mathbf{XXX}$. Выполняя это, мы находим логические закодированные состояния

$$|\bar{0}\rangle = \frac{1}{\sqrt{3}}(|000\rangle + |111\rangle + |222\rangle),$$

$$|\bar{1}\rangle = \frac{1}{\sqrt{3}}(|012\rangle + |120\rangle + |201\rangle),$$

$$|\bar{2}\rangle = \frac{1}{\sqrt{3}}(|021\rangle + |102\rangle + |210\rangle).$$

Для быстрого запоминания заметим, что $|\bar{1}\rangle$ и $|\bar{2}\rangle$ представляют собой суперпозиции четных и нечетных перестановок в S_3 соответственно.

d) Мы можем обобщить предыдущие результаты, чтобы создать квантовый код $[[3m, 3m-2, 2]]_3$. Его стабилизатор генерируется операторами

$$\mathbf{M}_1 = \mathbf{X}^{\otimes 3m},$$
$$\mathbf{M}_2 = \mathbf{Z}^{\otimes 3m}.$$

Они коммутируют между собой, поскольку $\omega^{3m} = 1$. Я благодарен Дэвиду Бэкману, нашедшему нормализатор этого кода:

$$\begin{aligned}
\bar{\mathbf{X}}_1 &= 1\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_1 &= \mathbf{Z}^{-1}\mathbf{Z} \\
\bar{\mathbf{X}}_2 &= 11\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_2 &= \mathbf{Z}^{-2}\mathbf{Z}\mathbf{Z} \\
\bar{\mathbf{X}}_3 &= 111\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_3 &= \mathbf{Z}^{-3}\mathbf{Z}\mathbf{Z}\mathbf{Z} \\
&\vdots & &\vdots \\
\bar{\mathbf{X}}_{3m-2} &= 1^{\otimes(3m-2)}\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_{3m-2} &= \mathbf{Z}^{-(3m-2)}\mathbf{Z}^{\otimes(3m-2)}
\end{aligned}$$

е) Мы можем обобщить эти результаты еще дальше, рассматривая кудиты вместо кутритов. Стабилизатор для квантового кода $[[d, d-2, 2]]_d$ представляет собой

$$\begin{aligned}\mathbf{M}_1 &= \mathbf{X}^{\otimes d},\\ \mathbf{M}_2 &= \mathbf{Z}^{\otimes d},\end{aligned}$$

а его нормализатор генерируется операторами

$$\begin{aligned}
\bar{\mathbf{X}}_1 &= 1\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_1 &= \mathbf{Z}^{-1}\mathbf{Z}\\
\bar{\mathbf{X}}_2 &= 11\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_2 &= \mathbf{Z}^{-2}\mathbf{Z}\mathbf{Z}\\
\bar{\mathbf{X}}_3 &= 111\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_3 &= \mathbf{Z}^{-3}\mathbf{Z}\mathbf{Z}\mathbf{Z}\\
&\vdots & &\vdots\\
\bar{\mathbf{X}}_{d-2} &= 1^{\otimes(d-2)}\mathbf{X}\mathbf{X}^{-1} & \bar{\mathbf{Z}}_{d-2} &= \mathbf{Z}^{-(d-2)}\mathbf{Z}^{\otimes(d-2)}
\end{aligned}$$

7.11. Коды коррекции ошибок в кудитах

а) Порядок каждого генератора кода $[[5, 1, 3]]_d$ равен d. Это следует из того, что для каждого $\mathbf{E}_{r,s} \in \{\mathbf{X}, \mathbf{X}^{-1}, \mathbf{Z}, \mathbf{Z}^{-1}\}$ r и s являются взаимно простыми с d. Следовательно, каждый генератор $\mathbf{M}_i$ не может иметь порядок, который является делителем d, но не равен d. В качестве интересного добавления заметим, что поскольку *только* 1 и $d-1$ при всех d являются взаимно простыми с d, каждый генератор кода, который справедлив при любом d, в своем разложении на тензорные произведения должен содержать оператор из следующего множества:

$$\{1, \mathbf{X}, \mathbf{Z}, \mathbf{X}^{-1}, \mathbf{Z}^{-1}, \mathbf{X}\mathbf{Z}, \mathbf{X}\mathbf{Z}^{-1}, \mathbf{X}^{-1}\mathbf{Z}, \mathbf{X}^{-1}\mathbf{Z}^{-1}\}.$$

Все генераторы действительно независимы. Приравнивая произведения произвольных степеней генераторов единице, мы вынуждены положить равными нулю все показатели степеней (это доказательство является мультипликативным двойником обычного доказательства линейной независимости):

$$\begin{aligned}
1 &= \mathbf{M}_1^{c_1}\mathbf{M}_2^{c_2}\mathbf{M}_3^{c_3}\mathbf{M}_4^{c_4} =\\
&= (\mathbf{X}^{c_1-c_2}\mathbf{Z}^{-c_4}) \otimes (\mathbf{Z}^{c_1}\mathbf{X}^{c_2-c_4}) \otimes (\mathbf{Z}^{c_2-c_1}\mathbf{X}^{c_3}) \otimes\\
&\quad \otimes(\mathbf{X}^{-c_1}\mathbf{Z}^{c_3-c_2}\mathbf{X}^{c_4}) \otimes (\mathbf{X}^{c_2}\mathbf{Z}^{c_2-c_1}),\\
&\Rightarrow c_1 = c_2 = c_3 = c_4 = 0.
\end{aligned}$$

Все генераторы также коммутируют, в чем можно убедиться непосредственной проверкой: всякий раз, когда элемент $\mathbf{X}^a\mathbf{Z}^b$ выстраивается между двумя генераторами, еще один элемент $\mathbf{Z}^{-b}\mathbf{X}^{-a}$ также выстраивается таким образом, что в общем генераторы будут коммутировать.

Пятая циклическая перестановка *не является* независимой от остальных

$$\begin{aligned}\mathbf{Z}\mathbf{Z}^{-1}\mathbf{X}^{-1}\mathbf{1}\mathbf{X} &= (\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3\mathbf{M}_4)^{-1} = \\ &= \mathbf{M}_4^{-1}\mathbf{M}_3^{-1}\mathbf{M}_2^{-1}\mathbf{M}_1^{-1} = \\ &= \mathbf{M}_1^{-1}\mathbf{M}_2^{-1}\mathbf{M}_3^{-1}\mathbf{M}_4^{-1}.\end{aligned}$$

b) Код $[[5,1,3]]_d$ невырожден. Чтобы понять это, заметим сначала, что его расстояние не больше трех, так как закодированные операции в части **c)** имеют вес три. Затем рассмотрим общий оператор Паули с весом $\leqslant 2$. Вследствие циклической природы кода этот оператор можно записать как $\mathbf{E}_a = \mathbf{E}_{r,s} \otimes \mathbf{E}_{t,u} \otimes \mathbf{1} \otimes \mathbf{1} \otimes \mathbf{1}$ или $\mathbf{E}_b = \mathbf{E}_{r,s} \otimes \mathbf{1} \otimes \mathbf{E}_{t,u} \otimes \mathbf{1} \otimes \mathbf{1}$.

В первом случае, чтобы коммутировать со стабилизатором, необходимо

$$\begin{aligned}[\mathbf{M}_1, \mathbf{E}_a] &= \omega^{-s+t} = 1, \\ [\mathbf{M}_2, \mathbf{E}_a] &= \omega^{u} = 1, \\ [\mathbf{M}_3, \mathbf{E}_a] &= \omega^{r} = 1, \\ [\mathbf{M}_4, \mathbf{E}_a] &= \omega^{-r+u} = 1, \\ &\Rightarrow r = s = t = u = 0.\end{aligned}$$

Во втором случае, чтобы коммутировать со стабилизатором, необходимо

$$\begin{aligned}[\mathbf{M}_1, \mathbf{E}_b] &= \omega^{-s-t} = 1, \\ [\mathbf{M}_2, \mathbf{E}_b] &= \omega^{t} = 1, \\ [\mathbf{M}_3, \mathbf{E}_b] &= \omega^{r-u} = 1, \\ [\mathbf{M}_4, \mathbf{E}_b] &= \omega^{-r} = 1, \\ &\Rightarrow r = s = t = u = 0.\end{aligned}$$

В обоих случаях это возможно только для единичного оператора. Следовательно, все операторы Паули с весом единица и два «антикоммутируют» с некоторым элементом стабилизатора и, следовательно, расстояние этого кода равно трем. Более того, это показывает, что не существует элемента стабилизатора, который может иметь вес меньше трех (так как он не может

коммутировать со всеми другими элементами стабилизатора), то есть код также является невырожденным.

c) Одним возможным представлением закодированных операций является

$$\bar{\mathbf{X}} = \mathbf{Z}^{-1}\,\mathbf{X}^{-1}\,\mathbf{Z}^{-1}\;\mathbf{1}\;\mathbf{1},$$
$$\bar{\mathbf{Z}} = \mathbf{X}\;\mathbf{X}\;\mathbf{1}\;\mathbf{Z}^{-1}\;\mathbf{1}.$$

Можно непосредственно проверить, что они имеют правильные коммутационные соотношения и коммутируют со всеми генераторами стабилизатора. Так как все сомножители включают только $\mathbf{Z}$, $\mathbf{X}$ и обратные к ним операторы, эти закодированные операции справедливы при любом d.

Глава 8

Топологические квантовые вычисления

8.1. Анионы?

Одной из главных идей квантовой теории является концепция *неразличимости частиц* (часто называемых *тождественными частицами*). Например, все электроны во Вселенной в точности одинаковы. Следовательно, для многоэлектронной системы операция *перестановки* двух электронов (обмена их положениями) является симметрией — она оставляет неизменной физику. Эту симметрию представляет унитарное преобразование, действующее на многоэлектронную волновую функцию.

Для неразличимых частиц в трехмерном пространстве, о котором мы обычно говорим в физике, перестановка частиц представляется одним из двух различных способов. Если частицы являются бозонами (например, атомы ^{4}He в сверхтекучей жидкости), то перестановку двух частиц представляет тождественный оператор: волновая функция инвариантна, то есть частицы подчиняются статистике Бозе. Если частицы являются фермионами (как, например, электроны в металлах), то перестановка представляется умножением на (-1): волновая функция меняет знак, то есть частицы подчиняются статистике Ферми.

В одномерном пространстве концепция статистики тождественных частиц становится неоднозначной. Дело в том, что в этом случае, обмениваясь местами, две частицы должны пройти друг сквозь друга. Если при перестановке двух тождественных частиц волновая функция меняет знак, то можно сказать, что они являются невзаимодействующими фермионами, но с тем же успехом можно сказать, что эти частицы являются взаимодействующими бозонами, так что изменение знака обусловлено взаимодействием проходящих друг сквозь друга частиц. В общем случае перестановка может изменить волновую функцию на мультипликативную фазу $e^{i\theta}$, принимающую значения, отличные от $+1$ и -1. Но даже такое изменение фазы можно учесть, описывая частицы или как бозоны, или как фермионы.

Таким образом, в пространстве трех (и большего числа) измерений, а также в одном измерении, статистика тождественных частиц довольно

проста. Но между этими двумя скучными случаями, в двумерном пространстве, возможно замечательно богатое разнообразие типов статистик частиц, настолько богатое, что мы просто утонем в нем, прежде чем сможем дать полезную классификацию всех возможностей.

Неразличимые частицы в двумерии, не являющиеся ни бозонами, ни фермионами, называются *анионами*. Анионы — это удивительно красивая теоретическая конструкция, но имеют ли они хоть какое-нибудь отношение к физике реальных, изучаемых в лаборатории, систем? Ответ замечателен: «Да!» Даже в нашем трехмерном мире можно создать двумерный электронный газ, удерживая электроны в тонком слое между двумя полупроводниковыми пластинами, так что при низких энергиях их движение в перпендикулярном слою направлении будет заморожено. Если электроны в материале достаточно подвижны, то в достаточно сильном магнитном поле и при достаточно низкой температуре такой двумерный электронный газ переходит в исключительно запутанное основное состояние, отделенное от всех возбужденных состояний отличной от нуля энергетической щелью. Более того, низкоэнергетические возбуждения частиц в таких системах не описываются электронными квантовыми числами; скорее, они являются анионами, несущими электрический заряд, равный дробной части заряда электрона. Анионы оказывают впечатляющее влияние на транспортные свойства образца, проявляющееся как дробный квантовый эффект Холла (ДКЭХ).

Анионы будут предметом нашего дальнейшего обсуждения. Но почему? В самом деле, я уже довольно много сказал в подтверждение того, что анионы — это глубокий и зачаровывающий объект. Все это так, но наш курс посвящен квантовым вычислениям, а отнюдь не экзотическим свойствам необычных фаз, реализующихся в конденсированных материальных системах.

Однако, между этими темами имеется тесная связь, впервые по достоинству оцененная Алексеем Китаевым в 1997 году: анионы обеспечивают необычные, захватывающие и, возможно, многообещающие способы реализации отказоустойчивых вычислений.

Так что, похоже, мы должны этим заинтересоваться. Как-никак, я уже прочитал 12 лекций по теории коррекции квантовых ошибок и отказоустойчивых вычислений. Это замечательная теория; я наслаждался, рассказывая о ней, и надеюсь, что вам также доставило удовольствие знакомство с этой теорией. Но она также и обескураживает. Мы видели, что идеальная квантовая схема может быть надежно смоделирована цепью шумящих вентилей, при условии, что шумят они не *слишком* сильно. Мы также видели, что требуемый для успешного моделирования верхний предел размера и глубины схемы вполне приемлем. Эти наблюдения вселяют в нас уверен-

ность, что работоспособные большие квантовые компьютеры когда-нибудь действительно будут построены. И все же, чтобы отказоустойчивость была действительно надежной, квантовые вентили должны иметь достаточно высокую точность воспроизведения (по меркам современной экспериментальной физики), а достижение этого требует значительных накладных расходов. Даже если в принципе надежные квантовые вычисления на шумящих вентилях возможны, никогда не исчезнет стремление повышать точность воспроизведения наших вычислений, совершенствуя, скорее, аппаратное обеспечение, нежели компенсируя его недостатки хитроумными схемными решениями. Используя анионы, можно достичь отказоустойчивости путем разработки «железа» с внутренней сопротивляемостью декогерентизации и другим ошибкам, заметно снижая стремительный рост размеров и глубины моделей наших схем. В таком случае у нас, очевидно, достаточно причин для изучения анионов. Не говоря даже о том, что это будет просто интересно!

В некоторых кругах этот предмет имеет репутацию (на мой взгляд, не совсем заслуженную) трудного и непостижимого. Я намерен начать с основ и не загромождать обсуждение несущественными для наших главных целей деталями. Таким образом, я надеюсь дать ясное представление без доходящих до абсурда упрощений.

Каковы наши цели? Я *не буду* объяснять, как теория анионов связана с наблюдаемыми явлениями в ДКЭХ-системах. В частности, в большинстве этих приложений возникают *абелевы* анионы. С точки зрения квантовой информации, абелевы анионы подходят для надежного *хранения* квантовой информации (и мы уже встречались с первыми признаками такой связи при изучении торических квантовых кодов). Мы обсудим здесь абелевы анионы, но наш основной интерес будет относиться к *неабелевым* анионам, которые, как мы увидим, могут быть наделены неожиданными вычислительными возможностями.

Как было замечено Китаевым [3], система неабелевых анионов с соответствующими свойствами может эффективно моделировать квантовые схемы; его идея была доработана Огберном и мной [4,5] и обобщена Мошоном [6, 7]. В первоначальной схеме Китаева для моделирования некоторого количества квантовых вентилей были необходимы измерения. Фридман, Ларсен и Ванг [11] заметили, что если использовать подходящие для этого анионы, то все измерения можно отложить вплоть до считывания окончательного результата вычисления. Фридман, Китаев и Ванг [10] также показали, что система анионов может эффективно моделироваться квантовой схемой; таким образом, вычислительные возможности анионного квантового компьютера и модели квантовых схем эквивалентны. Цель этих лекций — объяснить эти важные результаты.

Мы сосредоточимся на применении анионов к квантовым вычислениям, оставляя в стороне не менее важную проблему возможной реализации на практике системы анионов с требуемыми свойствами.[1] Я предлагаю вам подумать над этим самостоятельно!

8.2. Композиты поток–заряд

Тем из нас, у кого абстрактные математические конструкции вызывают отвращение, полезно начать изучение теории анионов с размышлений над конкретной моделью. Итак, начнем с напоминания о более знакомом понятии — *эффекте Ааронова–Бома*. Представим электромагнетизм в двумерном мире, где «трубка потока» является локализованным «точечным» объектом (в трех измерениях вы можете представить себе плоскость, пересекаемую перпендикулярным ей магнитным соленоидом). Поток может быть окружен непроницаемой стенкой, так что находящийся снаружи объект не может попасть в область, где магнитное поле отлично от нуля. Но даже в этом случае оно оказывает измеримое влияние на заряженные частицы, находящиеся за пределами трубки потока. Если частица с электрическим зарядом q адиабатически обходит (против часовой стрелки) вокруг потока Φ, ее волновая функция приобретает *топологическую фазу* $e^{iq\Phi}$ (где мы используем единицы, в которых $\hbar = c = 1$). Слово «топологический» здесь означает, что фаза Ааронова–Бома инвариантна относительно непрерывных деформаций траектории заряженной частицы — существенно только «количество оборотов» заряженной частицы вокруг потока.

Концепция топологической инвариантности естественным образом возникает при изучении отказоустойчивости. Топологическими свойствами являются те, что остаются неизменными при непрерывной деформации системы; а отказоустойчивым квантовым вентилем считается тот, чье действие на защищенную информацию остается неизменным (или почти неизменным), когда реализация вентиля деформируется включением шума. Топологическая инвариантность эффекта Ааронова–Бома является важнейшим свойством, которое мы надеемся использовать при разработке внутренне устойчивых квантовых вентилей.

Обычно эффект Ааронова–Бома рассматривается как явление, возникающее в квантовой электродинамике, в которой фотон является строго без-

[1] Недавно в литературе обсуждались два интересных подхода к реализации неабелевых анионов: (1) использование массивов сверхпроводящих контактов и (2) использование холодных атомов, удерживаемых оптическими решетками. [См., например: Л. Б. Иоффе, М. В. Фейгельман, *Реализация топологически защищенных квантовых битов в решетке джозефсоновских контактов*, Успехи физ. наук, **173**, 784–790 (2003). — *Прим. ред.*]

массовым. Однако важно понимать, что явления Аронова – Бома могут случаться и в массивных теориях. Например, мы можем рассмотреть «сверхпроводящую» систему, состоящую из частиц с зарядом e, так что композитные объекты с зарядом ne формируют конденсат (где n — целое число). В этом сверхпроводнике существует квант потока $\Phi_0 = 2\pi/ne$, минимальный ненулевой поток, при обходе вокруг которого частица конденсата, имеющая заряд (ne), приобретает *тривиальную* фазу Аронова – Бома. Изолированная область, содержащая квант потока, представляет собой окруженный сверхпроводящим конденсатом островок нормального металла, защищенный от расползания, поскольку магнитный поток не может проникнуть в сверхпроводник. Это стабильная частица, называемая *флаксоном*. Когда одна из частиц с зарядом e обходит вокруг флаксона, ее волновая функция приобретает нетривиальную топологическую фазу $e^{ie\Phi_0} = e^{2\pi i/n}$. Но в сверхпроводнике фотон приобретает массу благодаря механизму Хиггса, то есть безмассовых частиц здесь нет. Тот факт, что топологические фазы совместимы с массивными теориями, имеет очень важное значение, поскольку легко возбуждаемые безмассовые частицы являются потенциально богатым источником декогерентизации.

Теперь представим, что в нашем двумерном мире поток и электрический заряд намертво связаны друг с другом (по какой-то причине). Флаксон можно представить как поток Φ, захваченный внутри непроницаемого кругового барьера, а электрический заряд q приклеен к *внешней* стороне барьера. Что представляет собой угловой момент этого композита поток–заряд? Предположим, что мы осторожно поворачиваем этот объект на угол 2π против часовой стрелки, возвращая его к первоначальной ориентации. Совершив это, мы перенесли заряд q вокруг потока Φ, породив топологическую фазу $e^{iq\Phi}$. Этот поворот на 2π представляется в гильбертовом пространстве унитарным преобразованием

$$\mathbf{U}(2\pi) = e^{-i2\pi\mathbf{J}} = e^{iq\Phi}, \tag{8.1}$$

где $\mathbf{J}$ — оператор углового момента. Тогда мы приходим к выводу, что возможными собственными значениями углового момента являются

$$J = m - \frac{q\Phi}{2\pi} \quad (m = \text{целое число}). \tag{8.2}$$

Этот спектр можно характеризовать угловой переменной $\theta \in [0, 2\pi)$, определенной как $\theta = q\Phi \pmod{2\pi}$, и говорить, что собственные значения углового момента сдвинуты относительно целых значений на $-\theta/2\pi$. Будем говорить о фазе $e^{i\theta}$, представляющей поворот против часовой стрелки на угол 2π, как о *топологическом спине* композитного объекта.

Но не будет ли поворот на угол 2π действовать на физическую систему тривиально (как если бы ничего не происходило)? Нет, нам это хорошо известно из опыта обращения со *спинорами* в трехмерном пространстве. Для системы, содержащей F фермионов, мы имеем

$$e^{-2\pi i\mathbf{J}} = (-1)^F; \tag{8.3}$$

если количество фермионов нечетно, то собственные значения $\mathbf{J}$ сдвигаются на $1/2$ относительно целых значений. Этот сдвиг физически допустим, поскольку существует *правило суперотбора* по $(-1)^F$: не существует оператора локальной наблюдаемой, способного изменить значение $(-1)^F$ (не существует физического процесса, который может создать или уничтожить изолированный фермион). Действие оператора $e^{-2\pi i\mathbf{J}}$ на когерентную суперпозицию состояний с различными значениями $(-1)^F$ состоит в следующем:

$$e^{-i2\pi\mathbf{J}}(a|\text{ even } F\rangle + b|\text{ odd } F\rangle) = a|\text{ even } F\rangle - b|\text{ odd } F\rangle. \tag{8.4}$$

Относительный знак в суперпозиции обращается, но это не проявляется в наблюдаемых физических эффектах, поскольку все наблюдаемые блочно диагональны в базисе $(-1)^F$.

Аналогично, в двумерии сдвиг спектра углового момента $e^{-2\pi i\mathbf{J}} = e^{i\theta}$ не влечет за собой физически неприемлемых следствий, если существует правило суперотбора по θ, гарантирующее, что относительная фаза в суперпозиции состояний с различными значениями θ физически ненаблюдаема (не только практически, но и в принципе). Как и в случае фермионов, не существует разрешенного физического процесса, способного создать или разрушить изолированный анион.

В трех измерениях допустимы только значения $\theta = 0,\ \pi$ ввиду (как вам, вероятно, известно) топологического свойства трехмерной группы вращений SO(3): замкнутый путь в SO(3), начинающийся из тождественного преобразования и заканчивающийся поворотом на угол 4π, может быть непрерывным образом стянут в тривиальный путь. Отсюда следует, что поворот на угол 4π действительно представляет собой тождественное преобразование и, следовательно, собственными значениями поворота на угол 2π являются $+1$ и -1. Однако группа двумерных вращений SO(2) таким топологическим свойством не обладает, так что, в принципе, возможно любое значение θ.

Заметим, что угловой момент $\mathbf{J}$ изменяет знак при обращении времени ($\mathbf{T}$), а также при отражении ($\mathbf{P}$). За исключением случая, когда θ равен 0 или π, спектр $\mathbf{J}$ асимметричен относительно нуля, и, следовательно, теория анионов обычно неинвариантна относительно $\mathbf{T}$ или $\mathbf{P}$. В модели

композита заряд–поток природа этого нарушения симметрии не составляет загадки — оно происходит от ненулевого магнитного поля. Но если анионы появляются в системе без внутреннего нарушения **T** и **P**, то эти симметрии должны быть нарушены спонтанно или же спектр частиц должен быть «дуальным», так чтобы для каждого аниона с обменной фазой $e^{i\theta}$ существовала идентичная в других отношениях частица с обменной фазой $e^{-i\theta}$.

8.3. Спин и статистика

Для тождественных частиц в трех измерениях существует хорошо известная связь между спином и статистикой: неразличимые частицы с целым спином являются бозонами, а с полуцелым спином — фермионами. В двумерии спин может быть любым вещественным числом. Как эта новая возможность «дробного спина» проявляется в статистике? Ответ в том, что статистика тоже может быть «дробной»!

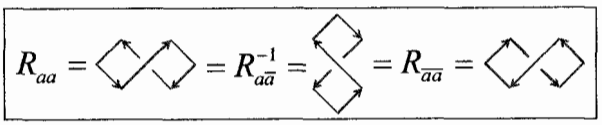
Что происходит, когда мы переставляем два композитных объекта поток–заряд против часовой стрелки? Каждый заряд q адиабатически проходит *половину пути* вокруг потока Ф другого объекта. Тогда можно предположить, что каждый заряд приобретает фазу Аронова–Бома, равную половине фазы, генерируемой при полном обороте заряда вокруг потока. Складывая возникающие от переноса обоих зарядов фазы, мы обнаружим, что перестановка двух композитов поток–заряд изменяет их волновую функцию на фазу

$$\exp\left[i\left(\frac{1}{2}q\Phi+\frac{1}{2}q\Phi\right)\right]=e^{iq\Phi}=e^{i\theta}=e^{-2\pi i\mathrm{J}}. \tag{8.5}$$

Фаза, генерируемая перестановкой двух объектов, совпадает с фазой, генерируемой поворотом одного из них на угол 2π. Таким образом, связь между спином и статистикой сохраняется в виде, который является естественным обобщением связи, применяемой к бозонам и фермионам.

В модели композита поток–заряд природа этой связи довольно прозрачна, но фактически она справедлива в значительно более общем случае. Почему? При чтении учебников по релятивистской квантовой теории поля может легко возникнуть впечатление, что связь между спином и статистикой основывается на лоренцевской инвариантности и имеет отношение к свойствам комплексной группы Лоренца. На самом деле это впечатление обманчиво. Все, что действительно важно для связи между спином и статистикой, — это *существование античастиц*. Специальная относительность — несущественный ингредиент.

Рассмотрим анион, характеризуемый фазой θ, и предположим, что эта частица имеет соответствующую античастицу. Это означает, что частица и ее античастица, комбинируясь, образуют объект с тривиальными квантовыми числами (в частности, с нулевым угловым моментом) и, следовательно, существуют физические процессы, в которых могут рождаться и уничтожаться пары частица–античастица. Проведем в пространстве-времени мировую линию, представляющую процесс, в котором рождаются две пары частица–античастица, одна пара слева, другая — справа (см. рис. ниже).[1] Частица из пары справа переставляется *против часовой стрелки* с частицей из пары слева, после чего обе пары аннигилируют. (Мировая линия имеет направление; если она направлена вперед во времени, то она представляет частицу, а если назад — античастицу.) Поворачивая нашу диаграмму на 90°, мы получаем изображение процесса, в котором рождается отдельная пара частица–античастица, затем частица и античастица переставляются *по часовой стрелке*, после чего пара аннигилирует. Повернем ее на 90° еще раз, тогда мы имеем процесс, в котором рождаются две пары, после чего, прежде чем аннигилировать, *античастица* из правой пары обменивается против часовой стрелки с античастицей из левой пары.



Какой вывод следует из этих манипуляций? Обозначим через R_{ab} унитарный оператор, представляющий перестановку против часовой стрелки двух частиц типа a и b (так что обратный оператор R^{-1}_{ab} представляет перестановку по часовой стрелке), и обозначим через $\bar{a}$ античастицу частицы a. Мы нашли, что

$$R_{aa} = R^{-1}_{a\bar{a}} = R_{\bar{a}\bar{a}}. \tag{8.6}$$

Если a — анион с обменной фазой $e^{i\theta}$, то его античастица $\bar{a}$ имеет *ту же* обменную фазу. Более того, если a и $\bar{a}$ обмениваются против часовой стрелки, то приобретаемая фаза равна $e^{-i\theta}$.

Эти выводы не удивительны, если их интерпретировать на основе модели аниона как композита поток–заряд. Античастица объекта с потоком Ф и зарядом q имеет поток $-\Phi$ и заряд $-q$. Следовательно, когда мы переставляем две античастицы, знаки минус взаимно уничтожаются и результат

[1] На всех трех диаграммах ось времени направлена снизу вверх, то есть изображаемые ими процессы начинаются в крайних нижних точках и заканчиваются в верхних. — *Прим. ред.*

будет тем же, как если бы переставлялись частицы. Но если мы переставляем частицу с античастицей, то относительный знак заряда и потока ведет к появлению обменной фазы $e^{-iq\Phi} = e^{-i\theta}$.

Но в чем состоит связь между этими наблюдениями и соотношением между спином и статистикой? Продолжая созерцать эту пространственно-временную диаграмму, обсудим вытекающие из нее выводы относительно *ориентации* частиц. Чтобы следить за ориентацией, удобно рассматривать мировую линию частицы не как нить, а как *ленту* в пространстве-времени. Я утверждаю, что наш процесс может быть непрерывным образом деформирован к процессу, в котором рождается пара частица–античастица, затем частица поворачивается против часовой стрелки на угол 2π, после чего пара аннигилирует. Удобный способ проверить это утверждение — снять свой ремень (или позаимствовать его у друга). Пряжка на одном конце задает ориентацию; направьте ваш большой палец навстречу пряжке и, прежде чем вновь застегнуть ремень, разверните ее на угол 2π, следуя правилу правой руки. Теперь вы должны быть в состоянии проверить, что можно ориентировать ремень так, чтобы сопоставить ему, а значит и процессу, в котором частица поворачивается на угол 2π, простанственно-временную диаграмму любого из описанных выше процессов перестановок.

Таким образом, в топологическом смысле поворот частицы на угол 2π против часовой стрелки фактически эквивалентен перестановке частиц против часовой стрелки (или перестановке частицы с античастицей по часовой стрелке), что предоставляет удовлетворительное объяснение общей связи между спином и статистикой.[1] Я еще раз подчеркиваю, что в этом рассуждении привлекаются процессы, в которых рождаются и уничтожаются пары частица–античастица, и, следовательно, существование античастиц является важнейшим предварительным условием наличия связи между спином и статистикой.

8.4. Объединение анионов

Мы знаем, что образованный из двух фермионов композитный объект является бозоном. Что произойдет, если построить композитный объект, комбинируя два аниона? Предположим, что a является анионом с обмен-

[1]На самом деле это обсуждение чересчур упрощено. Хотя оно адекватно для абелевых анионов, мы увидим, что для неабелевых анионов его следует усовершенствовать, поскольку в неабелевом случае R_{ab} имеет более одного собственного значения. Аналогично, обсуждение «комбинированных анионов» в следующем разделе также будет необходимо доработать, поскольку в неабелевом случае при соединении двух анионов можно получить более одного типа композитных анионов.

ной фазой $e^{i\theta}$ и что мы образуем «молекулу» из n таких анионов. Какая фаза накапливается при перестановке против часовой стрелки двух таких молекул?

В модели композита поток–заряд ответ очевиден. Каждый из n зарядов одной молекулы, проходя половину пути вокруг каждого из n потоков другой молекулы, приобретает фазу $e^{i\theta/2}$. Тогда всего порождается $2n^2$ фазовых множителей $e^{i\theta/2}$, давая в результате полную фазу

$$e^{i\theta_n} = e^{in^2\theta}. \tag{8.7}$$

Иначе говоря, фаза $e^{i\theta}$ появляется n^2 раз, поскольку на самом деле n анионов одной молекулы переставляются с n анионами другой молекулы. Если бы мы расщепили фермион (скажем) на две идентичные составляющие части, то, вопреки нашим наивным ожиданиям, эти составляющие имели бы обменные фазы $(e^{i\pi})^{1/4} = e^{i\pi/4}$, а не $\sqrt{-1} = i$.

Такое поведение совместимо со связью между спином и статистикой: угловой момент J n-анионной молекулы удовлетворяет условию

$$e^{-2\pi i J_n} = e^{-2\pi i n^2 J} = e^{in^2\theta}. \tag{8.8}$$

Рассмотрим, например, молекулу из двух анионов и представим, что она поворачивается против часовой стрелки на угол 2π. При этом не только каждый анион молекулы поворачивается на угол 2π; кроме этого, один из анионов обходит вокруг другого. Один оборот эквивалентен двум последовательным обменам, так что генерируемая при этом фаза равна $e^{i2\theta}$. Полным результатом двух поворотов и одного оборота является фаза

$$\exp[i(\theta + \theta + 2\theta)] = e^{i4\theta}. \tag{8.9}$$

Почему угловые моменты анионов комбинируются неаддитивно, можно понять и другим способом, заметив, что полный угловой момент молекулы состоит из двух частей — спинового углового момента S каждого из двух анионов (который *является* аддитивным) и *орбитального* углового момента L анионной пары. Поскольку перенос против часовой стрелки одного аниона вокруг другого порождает нетривиальную фазу $e^{i2\theta}$, зависимость двуханионной волновой функции ψ от относительного азимутального угла φ неоднозначна; вместо этого имеет место соотношение

$$\psi(\varphi + 2\pi) = e^{-i2\theta}\psi(\varphi). \tag{8.10}$$

Это означает, что спектр орбитального углового момента $\mathbf{L}$ сдвинут относительно целых значений:

$$e^{-i2\pi\mathbf{L}} = e^{2i\theta}, \tag{8.11}$$

и этот орбитальный угловой момент аддитивно комбинируется со спином S, в результате чего полный угловой момент становится равным

$$-2\pi J = -2\pi L - 2\pi S = 2\theta + 2\theta + 2\pi \cdot (\text{целое число}) = \\ = 4\theta + 2\pi \cdot (\text{целое число}). \quad (8.12)$$

Что если, с другой стороны, мы построим молекулу $\bar{a}a$ из аниона a и его античастицы $\bar{a}$? Тогда, как мы видели, спин S имеет то же самое значение, как и для молекулы aa. Но обменная фаза имеет противоположное значение, так что нецелая часть орбитального углового момента равна $-2\pi L = -2\theta$ вместо $-2\pi L = 2\theta$, а полный угловой момент $J = L + S$ является целым. Конечно, это свойство необходимо, для того чтобы пара $\bar{a}a$ могла аннигилировать, не оставляя после себя объекта, несущего нетривиальный угловой момент.

8.5. Унитарные представления группы «кос»

Мы уже отмечали, что спектр углового момента в двумерном пространстве обладает иными, нежели в трех измерениях, свойствами, поскольку SO(2) имеет другие по сравнению с SO(3) топологические свойства [(SO(3) имеет компактную односвязную накрывающую группу SU(2), тогда как SO(2) — нет]. Это наблюдение дает нам еще одну возможность понять, почему анионы возможны в двух измерениях и невозможны в трех. Также поучительно заметить, что *перестановка* частиц в двух и трех пространственных измерениях имеет разные топологические свойства.

Как мы обнаружили в нашем обсуждении связи между статистикой частиц и античастиц, перестановку частиц полезно рассматривать как процесс, протекающий в пространстве-времени. В частности, полезно представлять, что мы находим амплитуду квантового перехода для зависящего от времени n-частичного процесса, вычисляя сумму по историям частиц (хотя для наших целей вряд ли потребуется вычислять какие-либо интегралы по траекториям).

Рассмотрим систему n неразличимых точечных частиц, удерживаемых на двумерной поверхности (которую пока можно считать плоской), и предположим, что никакие две частицы не могут занимать одно и то же положение. Мы можем рассматривать конфигурацию частиц в фиксированный момент времени как плоскость с n «проколами» в указанных положениях, то есть с каждой частицей на поверхности связывается дырка с бесконечно малым радиусом. Условие, что частицам запрещено находиться в совпадающих позициях, обеспечивается тем, что в каждый момент времени

в плоскости существует точно n проколов. Более того, коль скоро частицы неразличимы, каждый прокол в точности такой же, как и любой другой. Таким образом, если мы произвели перестановку n проколов, это не повлечет за собой никаких физических эффектов; все проколы одинаковы, так что «кто есть кто» здесь совершенно не важно. Все, что действительно имеет значение, это n различных положений частиц на плоскости.

Чтобы вычислить квантовую амплитуду эволюции n частиц из конфигурации, задаваемой начальными положениями в момент времени $t = 0$, в конфигурацию, задаваемую их конечными положениями в момент времени $t = T$, необходимо просуммировать по всем классическим историям этих n частиц между начальной и конечной конфигурациями, взвешенным фазами e^{iS}, где S — классическое действие истории. Если мы рассматриваем мировую линию каждой частицы как нить, то каждую историю n частиц можно представить в виде «косы», в которой каждая частица начального ($t = 0$) временного среза может быть связана нитью с любой из частиц конечного ($t = T$) временного среза. Более того, поскольку пересечение мировых линий частиц запрещено, множество кос распадается на различные топологические классы, которые невозможно непрерывным образом деформировать друг в друга, а интеграл по траекториям разлагается на сумму вкладов, вносимых историями различных топологических классов.

Нетривиальные операции перестановок, действующие на частицы конечного временного среза, изменяют топологический класс косы. Таким образом, мы видим, что элементы группы симметрии, генерируемой перестановками, находятся во взаимно однозначном соответствии с топологическими классами. Эта (бесконечная) группа B_n называется группой кос из n нитей; групповой композиционный закон соответствует сочленению кос (то есть соединению следующих друг за другом кос). В квантовой теории состояние n неразличимых частиц принадлежит гильбертовому пространству, которое преобразуется по унитарному представлению группы кос B_n.

Группа может быть представлена набором генераторов, подчиняющихся особым определяющим соотношениям. Чтобы понять определяющие соотношения, представим, что n частиц занимают n распределенных на линии упорядоченных позиций (помеченных как $1, 2, 3, \ldots, n$). Пусть σ_1 обозначает перестановку против часовой стрелки частиц, первоначально занимавших позиции 1 и 2, σ_2 обозначает перестановку против часовой стрелки частиц, первоначально занимавших позиции 2 и 3, и так далее. Любая коса может быть построена как последовательность перестановок соседних частиц; следовательно, $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ являются генераторами группы.

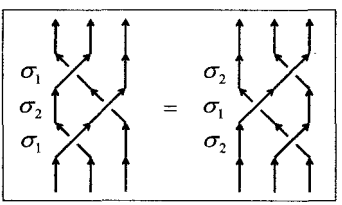
Имеется два типа определяющих соотношений, которым удовлетворяют эти генераторы. Соотношение первого типа

$$\sigma_j\sigma_k = \sigma_k\sigma_j, \quad |j-k| \geqslant 2, \tag{8.13}$$

утверждает лишь то, что перестановки несоседних пар частиц коммутируют. Соотношением второго, несколько более тонкого, типа является

$$\sigma_j\sigma_{j+1}\sigma_j = \sigma_{j+1}\sigma_j\sigma_{j+1}, \quad j = 1, 2, \ldots, n-2, \tag{8.14}$$

которое иногда называют *соотношением Янга–Бэкстера*. Вы можете проверить его, изобразив на листе бумаги две косы $\sigma_1\sigma_2\sigma_1$ и $\sigma_2\sigma_1\sigma_2$ и заметив, что обе они описывают процесс, в котором частицы, изначально находившиеся в позициях 1 и 3, переставляются против часовой стрелки вокруг частицы 2, которая остается неподвижной, то есть это топологически эквивалентные косы.



Поскольку группа кос бесконечна, она имеет бесконечное количество неприводимых унитарных представлений, а фактически существует бесконечное число *одномерных* представлений. Неразличимые частицы, преобразующиеся по одномерному представлению группы кос, называются *абелевыми анионами*. В одномерных представлениях каждый генератор σ_j группы B_n представлен фазой $\sigma_j = e^{i\theta_j}$. Более того, соотношение Янга–Бэкстера принимает вид $e^{i\theta_j}e^{i\theta_{j+1}}e^{i\theta_j} = e^{i\theta_{j+1}}e^{i\theta_j}e^{i\theta_{j+1}}$, откуда следует, что $e^{i\theta_j} = e^{i\theta_{j+1}} \equiv e^{i\theta}$ — все перестановки представляются *одной и той же* фазой. Конечно, смысл этого равенства в том, что обменные фазы не должны зависеть от того, какие пары переставляются, если частицы действительно неразличимы. При $\theta = 0$ мы получаем бозоны, а при $\theta = \pi$ — фермионы.

Группа кос также имеет множество неабелевых представлений, размерность которых больше единицы; неразличимые частицы, преобразующиеся по этим представлениям, называются *неабелевыми анионами* (или иногда *неабелионами*). Чтобы понять физические свойства неабелевых анионов, нам понадобится разобраться в математической структуре некоторых

из этих представлений. В ходе лекций я надеюсь передать некоторое интуитивное понимание неабелевых анионов, детально обсудив некоторые примеры.

А пока можно уже предугадать главную задачу, которую мы надеемся решить. Для неабелевых анионов реализованное n анионами неприводимое представление группы B_n действует на «топологическом векторном пространстве» V_n, размерность которого D_n экспоненциально растет вместе с n. Для анионов с подходящими свойствами образ представления может быть *плотным* в $\mathrm{SU}(D_n)$. Тогда сплетение анионов может моделировать квантовое вычисление, то есть подходящим выбором косы можно со сколь угодно высокой точностью воспроизведения реализовать любое (частное) унитарное преобразование, действующее в экспоненциально большом векторном пространстве V_n.

Таким образом, нас очень интересуют неабелевы представления группы кос. Но следует также подчеркнуть (а ниже мы обсудим это подробнее), что это больше относится к модели анионов, нежели просто к представлению группы кос. В модели трубки потока для абелевых анионов мы были в состоянии описывать не только результаты перестановок анионов, но также и типы частиц, которые можно получить, комбинируя два или больше анионов. Аналогично, общая анионная модель, описывающая различные типы анионов, включает в себя «правила композиции», которые определяют, какие типы анионов могут быть получены путем комбинации двух конкретных типов анионов. Нетривиальные условия согласования возникают вследствие ассоциативности соединения (слияние a с b с последующим соединением результата с c эквивалентно слиянию b с c с последующим соединением результата с a), а также вследствие того, что правила композиции (или слияния) должны согласовываться с правилами сплетения. Несмотря на то, что эти условия согласования накладывают жесткие ограничения, существует множество решений и, следовательно, множество в принципе реализуемых моделей неабелевых анионов.

8.6. Топологическое вырождение

Прежде чем перейти к неабелевым анионам, нам следует обсудить еще одно важное понятие, касающееся абелевых анионов. В любой модели анионов (в сущности, в любой локальной квантовой системе с массовой щелью) существует *основное*, или *вакуумное, состояние* — состояние, в котором отсутствуют частицы. Основное состояние на плоскости единственно, но в случае двумерной поверхности с нетривиальной топологией оно вырожденно, причем степень вырождения зависит от топологии. Мы уже

встречались с явлением «топологического вырождения» в модели абелевых анионов при изучении особого квантового кода коррекции ошибок, торического кода Китаева. Теперь мы увидим, что топологическое вырождение является общим свойством любой модели (абелевых) анионов.

К понятию топологического вырождения можно подойти, исследуя представления простой операторной алгебры. Рассмотрим случай тора, представляемого как квадрат с тождественными противоположными сторонами, и рассмотрим два фундаментальных 1-цикла на торе: C_1, который наматывается на квадрат в направлении x_1, и C_2, который наматывается в направлении x_2. Можно построить унитарный оператор $\mathbf{T}_1$, описывающий процесс, в котором рождается пара анион–антианион, анион распространяется вдоль C_1, после чего пара аннигилирует. Аналогично можно построить унитарный оператор $\mathbf{T}_2$, описывающий процесс, в котором рождается пара, и прежде чем она аннигилирует, анион распространяется вдоль цикла C_2. Каждый из операторов $\mathbf{T}_1$ и $\mathbf{T}_2$ сохраняет основное состояние системы (состояние в отсутствие частиц); действительно, каждый из них коммутирует с гамильтонианом $\mathbf{H}$ системы и, следовательно, может быть диагонализован одновременно с $\mathbf{H}$ ($\mathbf{T}_1$ и $\mathbf{T}_2$ являются симметриями).

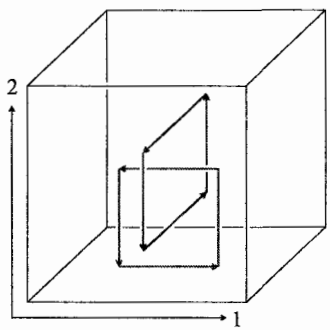
Однако $\mathbf{T}_1$ и $\mathbf{T}_2$ не коммутируют друг с другом. Если наш тор имеет бесконечный пространственный объем и существует массовая щель (так что взаимодействие между пространственно разделенными анионами возникает только благодаря эффекту Аронова–Бома), то коммутатор $\mathbf{T}_1$ и $\mathbf{T}_2$ равен[1]

$$\mathbf{T}_2^{-1}\mathbf{T}_1^{-1}\mathbf{T}_2\mathbf{T}_1 = e^{-i2\theta}\mathbf{1}, \tag{8.15}$$

где $e^{i\theta}$ — обменная фаза аниона. Нетривиальный коммутатор появляется, поскольку процесс (в котором (1) анион обходит вокруг C_1, (2) анион обходит вокруг C_2, (3) анион обходит вокруг C_1 в обратном направлении и (4) анион обходит вокруг C_2 в обратном направлении), топологически эквивалентен обходу по часовой стрелке одного аниона вокруг другого. Чтобы проверить это утверждение, рассмотрим действие $\mathbf{T}_2^{-1}\mathbf{T}_1^{-1}\mathbf{T}_2\mathbf{T}_1$ как процесс в пространстве-времени. Заметим сначала, что процесс, описываемый оператором $\mathbf{T}_1^{-1}\mathbf{T}_1$, в котором мировая линия аниона проходит сначала через C_1 и непосредственно после этого проходит C_1 в обратном направлении, может быть деформирован в процесс, в котором мировая линия аниона обходит топологически тривиальную петлю, которую можно непрерывным образом стянуть в точку (в соответствии с тем, что $\mathbf{T}_1^{-1}\mathbf{T}_1$ в действительности является единичным оператором). Подобным образом процесс, описы-

[1]Здесь используется алгебраическое определение коммутатора $[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{-1}$. — *Прим. ред.*

ваемый оператором $\mathbf{T}_2^{-1}\mathbf{T}_1^{-1}\mathbf{T}_2\mathbf{T}_1$, может быть деформирован к процессу, в котором мировые линии анионов обходят две замкнутые петли, но таким образом, что они один раз *сцепляются* друг с другом (см. рис. ниже); более того, одна петля протыкает поверхность, ограниченную другой петлей, в направлении, противоположном ее ориентации, определяемой правилом правой руки. Этот процесс можно непрерывным образом деформировать к процессу, в котором рождаются две пары, один анион обходит другой по часовой стрелке, а затем обе пары аннигилируют. Обход по часовой стрелке эквивалентен двум последовательным перестановкам по часовой стрелке, что в нашем одномерном представлении группы кос представляется фазой $e^{-i2\theta}$. Таким образом, $\mathbf{T}_1$ и $\mathbf{T}_2$ не коммутируют за исключением случаев $\theta = 0$ (бозоны) и $\theta = \pi$ (фермионы).



Поскольку $\mathbf{T}_1$ и $\mathbf{T}_2$ коммутируют с гамильтонианом $\mathbf{H}$, они оба сохраняют собственные подпространства $\mathbf{H}$, но поскольку $\mathbf{T}_1$ и $\mathbf{T}_2$ не коммутируют между собой, они не могут быть одновременно диагонализованы. Так как $\mathbf{T}_1$ – унитарный оператор, его собственные значения представляют собой фазы; используем угловую переменную $\alpha \in [0, 2\pi)$ для обозначения собственного состояния оператора $\mathbf{T}_1$ с собственным значением $e^{i\alpha}$:

$$\mathbf{T}_1|\alpha\rangle = e^{i\alpha}|\alpha\rangle. \tag{8.16}$$

Тогда действие оператора $\mathbf{T}_2$ на собственное состояние оператора $\mathbf{T}_1$ повышает значение α на 2θ:

$$\mathbf{T}_1(\mathbf{T}_2|\alpha\rangle) = e^{i2\theta}\mathbf{T}_2\mathbf{T}_1|\alpha\rangle = e^{i2\theta}e^{i\alpha}(\mathbf{T}_2|\alpha\rangle). \tag{8.17}$$

Предположим, что θ — *рациональное* кратное 2π, которое можно представить в виде

$$\theta = \pi p/q, \tag{8.18}$$

где q и p ($p < 2q$) — взаимно простые положительные целые числа. Тогда мы приходим к выводу, что $\mathbf{T}_1$ должен иметь по крайней мере q различных собственных значений; $\mathbf{T}_1$, действуя на $|\alpha\rangle$, генерирует орбиту с q различными значениями

$$\alpha + \left(\frac{2\pi p}{q}\right) k \pmod{2\pi}, \quad k = 0, 1, 2, \ldots, q-1. \tag{8.19}$$

Так как $\mathbf{T}_1$ коммутирует с $\mathbf{H}$, основное состояние нашей анионной системы на торе (по сути, собственное состояние с любой энергией) должно иметь вырождение, являющееся целым кратным числа q. Действительно, в общем случае (без учета дополнительных симметрий или случайных вырождений) кратность вырождения ожидается в точности равной q.

Для двумерной поверхности рода g (сферы с g «ручками») кратность топологического вырождения равна q^g, поскольку существуют аналогичные $\mathbf{T}_1$ и $\mathbf{T}_2$ операторы, связанные каждой из g ручек, и все операторы типа $\mathbf{T}_1$ могут быть одновременно диагонализованы. Более того, аналогичное рассуждение можно применить и к конечной плоской системе, если на ее *границах* могут рождаться и уничтожаться *отдельные* анионы. Рассмотрим, например, кольцо, на внутренней и внешней границах которого могут появляться и исчезать анионы. Тогда мы можем характеризовать унитарный оператор $\mathbf{T}_1$ как описывающий процесс, в котором анион обходит кольцо против часовой стрелки, а унитарный оператор $\mathbf{T}_2$ — как описывающий процесс, в котором анион появляется на внешней границе, распространяется к внутренней границе и там исчезает. Эти операторы $\mathbf{T}_1$ и $\mathbf{T}_2$ имеют такой же коммутатор, что и соответствующие операторы, определенные на торе. Таким образом, мы снова приходим к выводу, что при $\theta = \pi p/q$ основное состояние на кольце q-кратно вырожденно. Для диска с h дырками существует аналогичный $\mathbf{T}_1$ оператор, который оборачивает анион против часовой стрелки вокруг каждой дырки, и аналогичный $\mathbf{T}_2$ оператор, который переносит анион от внешней границы диска к границе дырки; таким образом, кратность вырождения равна q^h.

То, что мы здесь описали, представляет собой жесткую *топологическую квантовую память*. Фаза $e^{i2\theta} = e^{i2\pi p/q} \equiv \omega$, приобретаемая при обращении против часовой стрелки одного аниона вокруг другого, является первообразным q-м корнем из единицы, и в случае плоской системы с дырками оператор $\mathbf{T}_1$ может рассматриваться как закодированный оператор Паули $\bar{\mathbf{Z}}$, действующий на ассоциированную с данной дыркой q-мерную систему. С физической точки зрения, собственное значение ω^s оператора $\bar{\mathbf{Z}}$ лишь подсчитывает количество s анионов, «упакованных» внутри дырки.

Оператор $\mathbf{T}_2$ может рассматриваться как дополнительный оператор Паули $\bar{\mathbf{X}}$, увеличивающий значение s, переводя один анион от границы системы и помещая его в дырке. Поскольку квантовая информация закодирована в нелокальном свойстве системы, она хорошо защищена от декогерентизирующего действия окружения. Подобное помещение квантового состояния в память и его считывание может быть многообещающим для такой системы, так как, по крайней мере в принципе, $\bar{\mathbf{Z}}$ можно было бы измерить, скажем, выполняя интерференционный эксперимент, в котором налетающий анион рассеивается на дырке. Позднее мы увидим, что, используя неабелевы анионы, можно упростить считывание; кроме того, с неабелевыми анионами топологическими свойствами можно пользоваться как для *обработки* квантовой информации, так и для ее хранения.

Насколько устойчивой является эта квантовая память? Нас должны беспокоить ошибки, возникающие вследствие тепловых и квантовых флуктуаций. Тепловые флуктуации способны вызывать рождение анионов, которые могут диффундировать в окрестности одной из дырок в образце или от одной границы до другой, являясь причиной ошибок кодирования. Тепловые ошибки сильно подавляются больцмановским множителем $e^{-\Delta/T}$, если температура T достаточно мала по сравнению с энергетической щелью Δ (минимальной энергией, необходимой для рождения одного аниона на границе образца или анион-антианионной пары в объеме). Вредными квантовыми флуктуациями являются процессы туннелирования, в которых возникает виртуальная анион-антианионная пара и, прежде чем аннигилировать, анион распространяется вокруг дырки или виртуальный анион появляется на границе дырки и, прежде чем исчезнуть, распространяется до другой границы. Эти обусловленные квантовым туннелированием ошибки сильно подавляются, если дырки достаточно велики и достаточно далеко отделены друг от друга и от внешних границ.[1]

Заметим, что наш вывод о конечности топологического вырождения зависит от предположения о том, что угол θ является рациональным кратным π. Можно сказать, что теория анионов *рациональна*, если топологическое вырождение конечно для любой поверхности конечного рода (а для неабелевых анионов — если топологическое векторное пространство V_n конечномерно для любого конечного числа анионов n). Можно ожидать, что анионы, возникающие в любой физически разумной системе, будут в этом

[1]Если вы знакомы с методами евклидового интегрирования по траекториям, вы сможете просто проверить, что в главном полуклассическом приближении амплитуда A такого процесса туннелирования, в котором анион распространяется на расстояние L, имеет вид $A = Ce^{-L/L_0}$, где C — константа, а $L_0 = \hbar(2m^*\Delta)^{-1/2}$; здесь $\hbar$ — постоянная Планка, а m^* — эффективная масса аниона, определяемая таким образом, что кинетическая энергия аниона, движущегося со скоростью v, равна $m^*v^2/2$.

смысле рациональными, и, таким образом, следует ожидать, что они будут иметь обменные фазы, являющиеся корнями из единицы.

8.7. Еще раз о торических кодах

Если эти замечания о топологическом вырождении кажутся до боли знакомыми, то, возможно, потому, что аналогичные аргументы мы использовали при обсуждении торических кодов.

Торический код можно рассматривать как (вырожденное) основное состояние системы кубитов, заполняющих ребра квадратной решетки на торе, с гамильтонианом

$$\mathbf{H} = -\frac{1}{4}\Delta\left(\sum_P \mathbf{Z}_P + \sum_S \mathbf{X}_S\right), \tag{8.20}$$

здесь плакетный[1] оператор $\mathbf{Z}_P = \otimes_{\ell\in P}\mathbf{Z}_\ell$ представляет собой тензорное произведение операторов $\mathbf{Z}$, действующих на четыре кубита, расположенных на ребрах плакета P, а узельный оператор $\mathbf{X}_S = \otimes_{\ell\ni S}\mathbf{X}_\ell$ — тензорное произведение операторов $\mathbf{X}$, действующих на четыре кубита, расположенных на ребрах, сходящихся в узле S. Плакетные и узельные операторы являются в точности (коммутирующими) генераторами стабилизатора торического кода. Основное состояние одновременно является собственным состоянием всех генераторов стабилизатора с собственным значением $+1$.

В этой модели существует два типа возбуждений локализованных квазичастиц: плакетные возбуждения с $\mathbf{Z}_P = -1$,[2] которые можно представлять как магнитные флаксоны (кванты магнитного потока), и узельные возбуждения с $\mathbf{X}_S = -1$, которые можно толковать как электрические заряды. Действующая на ребре Z-ошибка порождает пару зарядов на двух связанных этим ребром узлах, тогда как действующая на ребре X-ошибка порождает пару флаксонов на двух разделенных этим ребром плакетах. Энергетическая щель Δ представляет собой плату за рождение пар любого из двух типов.

Заряды являются бозонами по отношению друг к другу (они имеют тривиальную обменную фазу $e^{i\theta} = 1$), флаксоны также являются бозонами относительно друг друга. Поскольку флаксоны отличимы от зарядов,

[1] Plaquette (франц.) — «плакет»; буквально: небольшая металлическая пластинка. В данном случае — элементарный квадрат решетки. Термин позаимствован из квантовой теории калибровочных полей на решетках. — *Прим. ред.*

[2] Это и следующее за ним аналогичное выражение представляют собой символическую запись того, что данное элементарное возбуждение является собственным состоянием оператора $\mathbf{Z}_P$ или $\mathbf{X}_S$ с собственным значением -1. — *Прим. ред.*

не имеет смысла рассматривать перестановки между ними. Но фаза (-1), приобретаемая при обходе заряда вокруг потока, делает эту модель анионной. Кратность вырождения основного состояния (размерность кодового пространства) можно интерпретировать как следствие этого свойства частиц.

Поскольку в этой модели на торе существует два типа частиц, существует и два типа операторов $\mathbf{T}_1$: оператор $\mathbf{T}_{1,S}$, который переносит заряд (узельный дефект) по 1-циклу C_1, и оператор $\mathbf{T}_{1,P}$, который переносит флаксон (плакетный дефект) по C_1. Аналогично, существует два типа операторов $\mathbf{T}_2$: операторы $\mathbf{T}_{2,S}$ и $\mathbf{T}_{2,P}$. Оба нетривиальных коммутатора

$$\mathbf{T}_{2,P}^{-1}\mathbf{T}_{1,S}^{-1}\mathbf{T}_{2,P}\mathbf{T}_{1,S} = -1 = \mathbf{T}_{2,S}^{-1}\mathbf{T}_{1,P}^{-1}\mathbf{T}_{2,S}\mathbf{T}_{1,P} \tag{8.21}$$

возникают в процессах, в которых мировые линии зарядов и флаксонов зацепляются только один раз. Таким образом, $\mathbf{T}_{1,S}$ и $\mathbf{T}_{2,S}$ могут быть одновременно диагонализованы и могут рассматриваться как закодированные операторы Паули $\bar{\mathbf{Z}}_1$ и $\bar{\mathbf{Z}}_2$, действующие на два защищаемых кубита. Оператор $\mathbf{T}_{2,P}$, коммутирующий с $\bar{\mathbf{Z}}_1$ и антикоммутирующий с $\bar{\mathbf{Z}}_2$, может рассматриваться как закодированный оператор $\bar{\mathbf{X}}_1$, и аналогично $\mathbf{T}_{1,P}$ представляет собой закодированный оператор $\bar{\mathbf{X}}_2$.

Для сконструированного нами на торе идеального гамильтониана вырождение четырех основных состояний является точным (частицы имеют бесконечные эффективные массы). Слабые локальные возмущения будут снимать вырождение, но только на величину, экспоненциально убывающую с ростом линейного размера тора L. Чтобы быть конкретнее, предположим, что возмущение представляет собой направленное вдоль оси $\hat{z}$ однородное «магнитное поле», взаимодействующее с магнитными моментами кубитов,

$$\mathbf{H}' = -h\sum_{\ell}\mathbf{Z}_{\ell}. \tag{8.22}$$

Поскольку энергетическая щель отлична от нуля, для вычисления главного вклада в расщепление вырождения по теории возмущений достаточно рассматривать действие возмущения в четырехмерном подпространстве, натянутом на основные состояния невозмущенной системы. В торическом коде операторами с нетривиальными матричными элементами в этом подпространстве являются те, чьи сомножители $\mathbf{Z}_{\ell}$ действуют на ребрах, образующих замкнутые петли, которые обходят вокруг тора (или чьи сомножители $\mathbf{X}_{\ell}$ действуют на ребрах, дуальных ребрам, которые образуют замкнутые петли, обходящие вокруг тора). Для решетки размера $L \times L$ на торе минимальная длина такой замкнутой петли равна L; следовательно, отличные от

нуля матричные элементы возникают, начиная лишь с L-го порядка теории возмущений, и подавляются множителем h^L. Таким образом, при малых h и больших L ошибки памяти, вызванные квантовыми флуктуациями, возникают только с экспоненциально малой амплитудой вероятности.

8.8. Неабелев эффект Аронова – Бома

Существует красивая абстрактная теория неабелевых анионов, и в свое время мы немного в ней покопаемся. Но я предпочел бы приступить к изучению этого предмета с описания более конкретной модели.

Имея в виду эту цель, вспомним некоторые особенности *хромодинамики*, теории кварков и глюонов, содержащихся внутри атомных ядер и других сильно взаимодействующих частиц. В реальном мире кварки постоянно связаны друг с другом и никогда не встречаются в свободном виде. Однако для нашего обсуждения представим фантастический мир, в котором силы между кварками настолько слабы, что характерный масштаб расстояний конфайнмента кварков очень велик.

Кварки имеют степень свободы, которую на образном языке называют *цветом*. Существует три типа кварков, которые, пользуясь этой метафорой, мы называем красным («red», R), желтым («yellow», Y) и синим («blue», B). Кварки всех трех цветов физически идентичны, и лишь когда мы сводим их вместе, можно говорить о том, являются ли их цвета одинаковыми (взаимодействие между одинаковыми цветами имеет характер отталкивания) или различными (разные цвета притягиваются друг к другу). Нет ничего, что помешало бы мне создать в моей лаборатории *кварковое бюро стандартов*, где раскрашенные кварки сортируются по трем корзинам; все кварки, лежащие в одной корзине, имеют одинаковые цвета, а кварки из разных корзин — разные. Мы можем (произвольным образом) снабдить три корзины метками — R, Y, B.

Если, прогуливаясь за пределами лаборатории, я обнаружу ранее не замеченный кварк, то на первых порах я буду неуверен относительно его цвета. Но это можно выяснить. Я захватываю кварк с собой и осторожно, чтобы по дороге не разрушить его цвет (в хромодинамике существует понятие *параллельного переноса* цвета), возвращаюсь в свою лабораторию. Вернувшись в кварковое бюро стандартов, я могу сравнить этот новый кварк с ранее калиброванными кварками в корзинах и определить, как должен быть помечен новый кварк: R, Y или B.

Это звучит просто, но есть одна загвоздка: в хромодинамике параллельный перенос цвета *зависит от пути*, благодаря влияющему на цвет

эффекту Ааронова–Бома. Допустим, что в кварковом бюро стандартов приготовлен кварк, цвет которого описывается квантовым состоянием

$$|\psi_q\rangle = q_R|R\rangle + q_Y|Y\rangle + q_B|B\rangle; \tag{8.23}$$

это когерентная суперпозиция с амплитудами q_R, q_Y и q_B для красного, желтого и синего состояний. Кварк движется вдоль пути, который обходит *трубку цветового магнитного потока* и возвращается в кварковое бюро стандартов, где его цвет может быть снова калиброван. По возвращении его цветовое состояние оказывается повернутым

$$\begin{pmatrix} q'_R \\ q'_Y \\ q'_B \end{pmatrix} = \mathbf{U} \begin{pmatrix} q_R \\ q_Y \\ q_B \end{pmatrix}, \tag{8.24}$$

где $\mathbf{U}$ — (специальная) унитарная 3×3-матрица. Подобно этому, если вновь обнаруженный кварк будет перенесен в бюро стандартов, результат измерения его цвета будет зависеть от того, обошел ли он во время своего вояжа трубку потока слева или справа.

Эта зависимость от пути параллельного переноса цвета очень похожа на зависимость от пути параллельного переноса касательного вектора на искривленном римановом многообразии. В хромодинамике магнитное поле представляет собой *кривизну*, величина которой определяет степень зависимости от пути.

В общем случае SU(3)-матрица $\mathbf{U}$, описывающая результат параллельного переноса цвета вдоль замкнутого пути, зависит от *стартовой точки* x_0, в которой этот путь начинается и заканчивается, а также и от замкнутой петли C, которую он описывает. В тех случаях, когда важно указывать петлю и базисную точку, мы будем использовать обозначение $\mathbf{U}(C, x_0)$. Собственные числа матрицы $\mathbf{U}$ имеют инвариантный «геометрический» смысл, характеризующий параллельный перенос, но сама $\mathbf{U}$ зависит от соглашений, принятых нами в стартовой точке. Возможно, вы предпочтете выбрать отличный от принятого мной ортонормированный базис для пространства цветов в стартовой точке x_0, так что ваши стандартные цвета R, Y и B будут отличаться от моих действием SU(3)-матрицы $\mathbf{V}(x_0)$. Тогда, если я описываю влияние параллельного переноса вокруг петли C матрицей $\mathbf{U}$, вы характеризуете его другой матрицей

$$\mathbf{V}(x_0)\mathbf{U}(C, x_0)\mathbf{V}^{-1}(x_0), \tag{8.25}$$

отличающейся от моей сопряжением матрицей $\mathbf{V}(x_0)$. Физики иногда говорят о свободе переопределения соглашений как о выборе *калибровки* и го-

ворят, что сама $\mathbf{U}$ *зависит от калибровки*, тогда как ее собственные числа — *калибровочно инвариантны*.

Хромодинамика на рассматриваемых здесь масштабах (гораздо меньших характерной длины конфайнмента кварков) представляет собой теорию типа электродинамики с дальнодействующими кулоновскими взаимодействиями между кварками, переносимыми посредством «глюонных» полей. Мы предпочтем рассматривать теорию, которая сохраняет некоторые черты хромодинамики (в частности, зависимость от пути переноса цвета), но без легко возбуждаемых легких глюонов. В случае электродинамики мы исключали легкие фотоны, рассматривая «сверхпроводник», в котором заряженные частицы образуют конденсат, магнитные поля вытеснены, а магнитный поток изолированного объекта квантуется. Обратимся к этой идее и здесь. Рассмотрим *неабелев сверхпроводник* в двух пространственных измерениях. Этот мир содержит частицы, несущие «кванты магнитного потока» (подобные квантам цветового магнитного потока в хромодинамике), и частицы, несущие заряд (подобные цветным кваркам хромодинамики). Поток принимает значения в *конечной неабелевой группе* G, а зарядам соответствуют *унитарные неприводимые представления* группы G. В такой постановке можно сформулировать некоторые интересные модели неабелевых анионов.

Пусть R обозначает конкретное неприводимое представление группы G, размерность которого обозначается как $|R|$. Мы можем открыть «зарядовое бюро стандартов» и определить произвольно выбранный ортонормированный базис в $|R|$-мерном векторном пространстве, в котором действует R:

$$|R,i\rangle, \quad i=1,2,\ldots|R|. \tag{8.26}$$

Если заряд R обходит замкнутый путь, окружающий поток $a \in G$, то возникает нетривиальный эффект Ааронова – Бома — базис для R поворачивается унитарной матрицей $\mathbf{D}^R(a)$, представляющей a:

$$|R,j\rangle \longmapsto \sum_{i=1}^{|R|} |R,i\rangle D^R_{ij}(a). \tag{8.27}$$

В принципе, матричные элементы $D^R_{ij}(a)$ измеримы, например, в интерференционных экспериментах, в которых пучок калиброванных зарядов может пройти по любую сторону от кванта потока. (Фаза комплексного числа $D^R_{ij}(a)$ определяет величину *сдвига* интерференционных полос, а модуль $D^R_{ij}(a)$ — их *контрастность*.) Таким образом, как только для зарядов выбран стандартный базис, их можно использовать для того, чтобы присвоить метки (элементы G) всем квантам потока. Это соответствие однозначно,

коль скоро представление R является точным и исключены любые групповые автоморфизмы (порождающие неоднозначности, которые мы свободны разрешать по своему усмотрению).

Однако групповые элементы, которые мы сопоставляем квантам потока, зависят от наших соглашений. Допустим, что мне предоставили k флаксонов (частиц, несущих кванты потока), и я использую мои стандартные заряды для измерения кванта потока каждой частицы. Я ставлю в соответствие этим k флаксонам элементы группы $a_1, a_2, \ldots, a_k \in G$. Затем прошу вас измерить квант потока, чтобы проверить мое распределение. Но ваши стандартные заряды отличаются от моих, вследствие того, что их незаметно перенесли вокруг другого потока (который я бы обозначил как $g \in G$). Следовательно, этим же k флаксонам вы сопоставите элементы группы $ga_1g^{-1}, ga_2g^{-1}, \ldots, ga_kg^{-1}$; то есть наши сопоставления отличаются сопряжением по элементу g.

Урок, извлекаемый из этого примера, состоит в том, что сопоставление флаксонам групповых элементов по сути неоднозначно и не имеет инвариантного смысла. Но поскольку правильные соответствия между элементами группы и флаксонами отличаются лишь сопряжением по некоторому элементу $g \in G$, действительно инвариантный смысл, с которым будут согласны все наблюдатели, имеют *классы сопряженных элементов*, соответствующих квантам потока в G. Действительно, даже если мы зафиксируем наши соглашения в зарядовом бюро стандартов, сопоставляемый конкретному флаксону групповой элемент может измениться, если этот флаксон примет участие в физическом процессе, в котором он сплетется с другими флаксонами. По этой причине флаксоны, принадлежащие одному классу сопряженных элементов, должны рассматриваться как неразличимые частицы, даже если они предстают во множестве разных лиц (по одному для каждого представителя класса), которые можно различить, выполняя измерения в нужное время, в нужном месте: флаксоны представляют собой *неабелевы анионы*.

8.9. Сплетение неабелевых флаксонов

На примере неабелева сверхпроводника с подходящими свойствами мы увидим, что отказоустойчивый универсальный квантовый компьютер может функционировать, оперируя флаксонами. Самое главное — понять, что происходит при перестановке двух флаксонов.

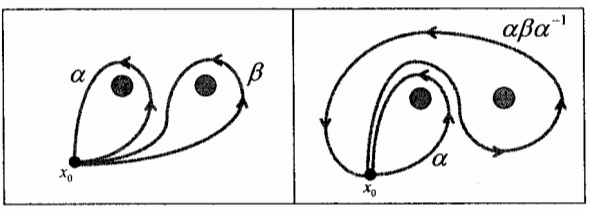
Представим с этой целью, что мы тщательно калибруем два флаксона и сопоставляем им элементы группы G. Это соответствие определяется вы-

бором стандартного базиса для заряженных частиц в начальной точке x_0. Затем выбирается типичный путь, обозначаемый как α, который начинается в точке x_0, обходит против часовой стрелки находящийся *слева* флаксон и возвращается в x_0. Итак, переместив заряженные частицы вдоль замкнутого пути α, мы обнаруживаем, что при этом параллельном переносе их состояния преобразуются матрицей $\mathbf{D}(a)$, где D — представление группы G, по которому преобразуются состояния заряженных частиц, а $a \in G$ — конкретный элемент группы, сопоставляемый флаксону. Аналогично, выбирается другой типичный путь, обозначаемый как β, который начинается в точке x_0, обходит против часовой стрелки флаксон, находящийся *справа*, и возвращается в x_0. Результатом параллельного переноса вдоль β является преобразование состояния заряженной частицы матрицей $\mathbf{D}(b)$, и, таким образом, находящемуся справа флаксону сопоставляется элемент $b \in G$.

Теперь представим, что выполняется перестановка двух флаксонов против часовой стрелки, после чего процедура калибровки повторяется. Какие групповые элементы будут сопоставлены флаксонам теперь?

Чтобы найти ответ, рассмотрим путь $\alpha\beta\alpha^{-1}$; здесь мы используем α^{-1} для обозначения пути α, пройденного в противоположном направлении, и принимаем соглашение, что $\alpha\beta\alpha^{-1}$ обозначает путь, в котором сначала проходится α^{-1}, затем — β и, наконец, α. Теперь видно, что если при перестановке двух флаксонов против часовой стрелки так деформировать пути, чтобы они нигде не пересекались флаксонами, то путь $\alpha\beta\alpha^{-1}$ преобразуется в α, тогда как путь α — в β:

$$\alpha\beta\alpha^{-1} \longmapsto \alpha, \quad \alpha \longmapsto \beta. \tag{8.28}$$



Отсюда следует, что после перестановки результат обхода заряда вдоль пути α эквивалентен результату его перемещения вдоль пути $\alpha\beta\alpha^{-1}$ до перестановки; аналогично, результат обхода вдоль пути β после перестановки эквивалентен результату перемещения вдоль α до перестановки. Мы приходим к выводу, что представляющий перестановку против часовой стрелки оператор сплетения $\mathbf{R}$ действует на флаксоны по правилу

$$\mathbf{R} : |a, b\rangle \longmapsto |aba^{-1}, a\rangle. \tag{8.29}$$

Конечно, если сопоставляемые этим флаксонам a и b являются коммутирующими элементами группы G, то единственным результатом сплетения будет взаимный обмен позициями этих частиц. Но если a и b не коммутируют, результат перестановки более тонкий и интересный. Ассиметрия действия оператора $\mathbf{R}$ является следствием наших соглашений, а также направления перестановки (против часовой стрелки); обратный оператор $\mathbf{R}^{-1}$, представляющий перестановку по часовой стрелке, действует как

$$\mathbf{R}^{-1} : |a, b\rangle \longmapsto |b, b^{-1}ab\rangle. \tag{8.30}$$

Отметим, что суммарный поток пары флаксонов можно детектировать с помощью заряженной частицы, которая обходит путь $\alpha\beta$, окружающий оба элемента данной пары. Поскольку заряд, детектирующий этот суммарный поток, в принципе, может находиться очень и очень далеко, перестановка не должна изменять общий поток; действительно, мы обнаруживаем, что результирующий поток ab сохраняется как оператором $\mathbf{R}$, так и оператором $\mathbf{R}^{-1}$.

Результатом двух последовательных перестановок против часовой стрелки является оператор «монодромии» $\mathbf{R}^2$, представляющий оборот одного флаксона вокруг другого против часовой стрелки и действующий следующим образом:

$$\mathbf{R}^2 : |a, b\rangle \longmapsto |(ab)a(ab)^{-1}, (ab)b(ab)^{-1}\rangle; \tag{8.31}$$

оба потока сопрягаются полным потоком ab, то есть оборот a вокруг b против часовой стрелки ведет к сопряжению b потоком a (и аналогично, оборот b вокруг a по часовой стрелке ведет к сопряжению a потоком b^{-1}). Нетривиальная монодромия означает, что если на плоскости распределено множество флаксонов, один из которых отправляется в мою лабораторию для анализа, то сопосталяемый этому флаксону групповой элемент, вообще говоря, зависит от траектории его перемещения в лабораторию. Если при одном выборе пути квант потока маркируется элементом $a \in G$, то при другом ему, в принципе, может быть сопоставлен любой элемент вида bab^{-1}. Таким образом, инвариантом является представляющий флаксон класс сопряженных элементов в группе G, тогда как конкретный представитель этого класса определяется неоднозначно.

Предположим, например, что G представляет собой S_3 — группу перестановок трех объектов. Один из ее классов сопряженных элементов $\{(12), (23), (31)\}$ имеет порядок три, то есть содержит три элемента — все циклы длины два (перестановки или транспозиции двух объектов —

два-циклы). При объединении (слиянии) пары сопоставляемых этим циклам флаксонов (или, для краткости, два-флаксонов), для суммарного потока существует три возможности — тривиальный квант потока e, или один из квантов потока, сопоставляемых циклам длины три (три-циклам): (123)[1] или (132). Если суммарный поток тривиален, сплетение двух квантов потока также тривиально (a и $b = a^{-1}$ коммутируют). Но если суммарный поток нетривиален, то оператор сплетения $\mathbf{R}$ имеет орбиту длины три:

$$\begin{aligned} &\mathbf{R} : |(12),(23)\rangle \longmapsto |(31),(12)\rangle \longmapsto |(23),(31)\rangle \longmapsto |(12),(23)\rangle, \\ &\mathbf{R} : |(23),(12)\rangle \longmapsto |(31),(23)\rangle \longmapsto |(12),(31)\rangle \longmapsto |(23),(12)\rangle. \end{aligned} \tag{8.32}$$

Таким образом, если два флаксона перестанавливаются трижды, они обмениваются положениями (число перестановок нечетное), но обозначение состояния остается неизменным. Это наблюдение говорит о возможности существования квантовой интерференции между «прямым» и «обменным» рассеянием двух флаксонов, которым сопоставляются разные элементы одного и того же класса сопряженных элементов; это укрепляет представление о том, что флаксоны, переносящие сопряженные «метки», должны рассматриваться как неразличимые частицы.

Поскольку действующий на пары два-флаксонов оператор сплетения удовлетворяет равенству $\mathbf{R}^3 = \mathbf{1}$, его собственные значения равны кубическому корню из единицы. Например, взяв линейные комбинации трех состояний с суммарным потоком (123), мы получим собственные состояния оператора $\mathbf{R}$

$$\begin{aligned} \mathbf{R} = 1: &\quad |(12),(23)\rangle + |(31),(12)\rangle + |(23),(31)\rangle, \\ \mathbf{R} = \omega: &\quad |(12),(23)\rangle + \bar{\omega}|(31),(12)\rangle + \omega|(23),(31)\rangle, \\ \mathbf{R} = \bar{\omega}: &\quad |(12),(23)\rangle + \omega|(31),(12)\rangle + \bar{\omega}|(23),(31)\rangle, \end{aligned} \tag{8.33}$$

где $\omega = e^{2\pi i/3}$, а $\bar{\omega} = \omega^2 = e^{-2\pi i/3}$.

Несмотря на то, что пара флаксонов $|a, a^{-1}\rangle$ с тривиальным суммарным потоком имеет тривиальные свойства сплетения, она интересна по другой причине — она несет *заряд*. Для определения заряда объекта нужно обнести вокруг него квант потока b (против часовой стрелки); это изменяет объект действием матрицы $\mathbf{D}^R(b)$ для некоторого представления R группы G. Если заряд равен нулю, то это представление тривиально: $\mathbf{D}(b) = \mathbf{1}$ для всех значений $b \in G$. Но если мы переносим против часовой стрелки

[1] Напомним, что символ $(n_1 n_2 n_3 \ldots n_k)$ обозначает циклическую перестановку k элементов (k-цикл) $n_1 \to n_2 \to n_3 \to \ldots \to n_k \to n_1$. — *Прим. ред.*

квант потока b вокруг состояния $|a, a^{-1}\rangle$, то это состояние трансформируется в

$$|a, a^{-1}\rangle \longmapsto |bab^{-1}, ba^{-1}b^{-1}\rangle, \tag{8.34}$$

что представляет собой нетривиальное действие (по крайней мере, для некоторых значений b), если a принадлежит классу, содержащему более одного элемента. Действительно, для каждого класса сопряженных элементов α существует единственное состояние $|0;\alpha\rangle$ с нулевым зарядом, однородная суперпозиция представителей класса:

$$|0;\alpha\rangle = \frac{1}{\sqrt{|\alpha|}} \sum_{a\in\alpha} |a, a^{-1}\rangle, \tag{8.35}$$

где $|\alpha|$ означает порядок α. Пара флаксонов класса α, которую можно создать в локальном процессе, не должна нести каких-либо сохраняющихся зарядов и, следовательно, должна находиться в состоянии $|0;\alpha\rangle$. Другие линейные комбинации, ортогональные состоянию $|0,\alpha\rangle$, несут ненулевой заряд. Этот переносимый парой флаксонов заряд можно детектировать с помощью других флаксонов, но, как это ни странно, его нельзя локализовать ни на одной из частиц пары. Скорее, это коллективное свойство пары. При столкновении двух флаксонов с отличным от нуля полным зарядом аннигиляция пары запрещена законом сохранения заряда, даже если их полный поток равен нулю.

Например, для пары флаксонов, соответствующих классу два-циклов группы $G = S_3$, существует двумерное подпространство с тривиальным суммарным потоком и нетривиальным зарядом, в котором можно выбрать базис

$$\begin{aligned} |0\rangle &= |(12),(12)\rangle + \bar{\omega}|(23),(23)\rangle + \omega|(31),(31)\rangle, \\ |1\rangle &= |(12),(12)\rangle + \omega|(23),(23)\rangle + \bar{\omega}|(31),(31)\rangle. \end{aligned} \tag{8.36}$$

При обходе квантом потока b вокруг данной пары, оба ее потока сопрягаются элементом b; следовательно, действие (посредством сопряжения) группы S_3 на эти состояния определяется матрицами

$$\begin{gathered} \mathbf{D}(12) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{D}(23) = \begin{pmatrix} 0 & \bar{\omega} \\ \omega & 0 \end{pmatrix}, \quad \mathbf{D}(31) = \begin{pmatrix} 0 & \omega \\ \bar{\omega} & 0 \end{pmatrix}, \\ \mathbf{D}(123) = \begin{pmatrix} \omega & 0 \\ 0 & \bar{\omega} \end{pmatrix}, \quad \mathbf{D}(132) = \begin{pmatrix} \bar{\omega} & 0 \\ 0 & \omega \end{pmatrix}, \end{gathered} \tag{8.37}$$

образующими ее двумерное неприводимое представление $R = [2]$. Таким образом, мы приходим к выводу, что заряд этой пары флаксонов равен $[2]$.

Более того, при сплетении этот переносимый парой флаксонов заряд может перейти на другие частицы. Рассмотрим, например, пару частиц, каждая из которых несет заряд, но не имеет кванта потока (я буду называть такие частицы *чарджионами*), так что полный заряд пары тривиален. Если состояние одного из чарджионов преобразуется по унитарному неприводимому представлению R группы G, то существует единственное сопряженное представление $\bar{R}$, которое можно скомбинировать с R, чтобы получить тривиальное представление; если $\{|R,i\rangle\}$ является базисом для R, то для $\bar{R}$ можно выбрать такой базис $\{|\bar{R},i\rangle\}$, в котором состояние пары чарджионов с тривиальным зарядом представляется в виде

$$|0;R\rangle = \frac{1}{\sqrt{|R|}} \sum_i |R,i\rangle \otimes |\bar{R},i\rangle. \tag{8.38}$$

Предположим, что мы создали пару флаксонов в состоянии $|0;\alpha\rangle$ и пару чарджионов в состоянии $|0;R\rangle$. Затем обводим против часовой стрелки чарджион с зарядом R вокруг флаксона с квантом потока класса α и снова сталкиваем два чарджиона, чтобы посмотреть, не аннигилируют ли они. Что же происходит?

При фиксированном значении кванта потока $a \in \alpha$ влияние обхода на состояние пары чарджионов согласно (8.27) состоит в следующем:

$$|0;R\rangle \longmapsto \frac{1}{\sqrt{|R|}} \sum_{i,j} |R,j\rangle \otimes |\bar{R},i\rangle D^R_{ji}(a); \tag{8.39}$$

если сейчас измерить заряд пары, то вероятность получения нулевого полного заряда равна квадрату модуля перекрытия этого состояния с $|0;R\rangle$, то есть

$$\mathrm{Prob}(0) = \left| \frac{\chi^R(a)}{|R|} \right|^2, \tag{8.40}$$

где

$$\chi^R(a) = \sum_i D^R_{ii}(a) = \mathrm{tr}\mathbf{D}^R(a), \tag{8.41}$$

— *характер* представления R, вычисленный для данного элемента a. Действительно, характер (след) инвариантен относительно операции сопряжения — он принимает одно и то же значение для всех $a \in \alpha$. Следовательно, (8.40) также представляет собой вероятность того, что полный заряд пары чарджионов останется равным нулю, если один из чарджионов (элемент соответствующей пары в исходном состоянии $|0;R\rangle$) обойдет вокруг

одного из флаксонов (элемента соответствующей пары в состоянии $|0;\alpha\rangle$). Конечно, поскольку полный заряд всех четырех частиц равен нулю, а заряд сохраняется, то после обхода эти две пары приобретают противоположные заряды: если пара чарджионов имеет заряд R', то пара флаксонов должна приобрести заряд $\bar{R}'$, который, комбинируясь с R', дает в результате тривиальный полный заряд. Пара частиц с нулевым полным зарядом и квантом потока может аннигилировать, не оставляя после себя ни одной стабильной частицы, тогда как пара с ненулевым зарядом полностью аннигилировать не может. Таким образом, если мировые линии пар флаксонов и чарджионов сцепляются один раз, вероятность того, что обе пары будут способны аннигилировать, задается уравнением (8.40). Эта вероятность меньше единицы, при условии, что представление R не одномерно, а класс α представлен нетривиально. Таким образом, сцепление мировых линий индуцирует обмен зарядами между двумя парами.

Например, в случае, когда α является классом два-циклов группы $G = S_3$, а $R = [2]$ (двумерное неприводимое представление группы S_3), из уравнений (8.37) следует, что $\chi^{[2]}(\alpha) = 0$. Таким образом, в этом случае вероятность обмена зарядом равна единице; после оборота обе пары флаксонов и чарджионов преобразуются по $R' = [2]$.

8.10. Суперотборные секторы неабелева сверхпроводника

В предыдущем обсуждении неабелева сверхпроводника рассматривались два типа частиц: *флаксоны*, несущие квант потока, но не имеющие заряда, и *чарджионы*, несущие заряд, но не поток. Это не самые общие возможные частицы. Поучительно рассмотреть, что получится, если создать сложную частицу, составленную из флаксона и чарджиона. В частности, чему будет равен заряд такого композита? Это на редкость тонкий вопрос; чтобы дать обоснованный ответ, нужно как следует подумать о том, как можно измерить заряд.

В принципе, заряд можно измерить с помощью интерференционного эксперимента Аронова–Бома. Объект, заряд которого мы хотим определить, можно поместить непосредственно за экраном между двумя (прорезанными в нем параллельными) щелями, направить на этот экран пучок тщательно откалиброванных флаксонов и детектировать их по ту сторону экрана. Выявленные таким образом сдвиг и контрастность интерференционной картины позволяют определить $\mathbf{D}^R(b)$ для каждого $b \in G$ и таким образом определить R.

Однако все не так просто, если вместе с зарядом анализируемый объект несет нетривиальный поток $a \in G$. Поскольку перенос потока b во-

круг a меняет a на bab^{-1}, то в случае некоммутирующих a и b два возможных пути для потока b не интерферируют. Действительно, после детектирования потока b можно проверить, не изменился ли поток a, и, следовательно, определить, прошел поток b сквозь щель слева или справа от a. Поскольку поток (a или bab^{-1}) коррелирует с информацией «какой путь» (левая или правая щель), интерференция разрушается.

Следовательно, этот эксперимент раскрывает информацию о заряде, только при коммутирующих a и b. Поэтому прикрепленный к потоку a заряд не описывается неприводимым представлением группы G; вместо этого он описывается как неприводимое представление подгруппы группы G, нормализатора $N(a)$ потока a в группе G, который определяется как

$$N(a) = \{b \in G | ab = ba\}. \tag{8.42}$$

Нормализаторы $N(a)$ и $N(bab^{-1})$ изоморфны, поэтому нормализатор естественнее ассоциировать с классом α группы G, а не с конкретным ее элементом, и обозначать как $N(\alpha)$. Следовательно, каждый тип частиц, возникающих в нашем неабелевом сверхпроводнике, в действительности имеет две метки: класс сопряженных элементов α, характеризующий поток, и описывающее заряд неприводимое представление $R^{(\alpha)}$ нормализатора $N(\alpha)$. Принято говорить, что α и $R^{(\alpha)}$ характеризуют *суперотборные секторы* теории, поскольку они представляют свойства локализованного объекта, которые обязаны сохраняться в любых локальных физических процессах. Для частиц, несущих метки $(\alpha, R^{(\alpha)})$, можно учредить «бюро стандартов», обращаясь к которому в конкретное время и в конкретном месте, можно распознать всего $|\alpha| \cdot |R^{(\alpha)}| \equiv d_{(\alpha, R(\alpha))}$ различных видов частиц — это число называется *размерностью* сектора. Если эти частицы сплетаются с другими, их виды могут меняться, тогда как метки $(\alpha, R^{(\alpha)})$ остаются неизменными.

В любой теории анионов каждому типу частиц можно приписать размерность, хотя, как мы увидим, в общем случае она не обязана быть целой и толковаться как количество различных видов частиц одного типа (одного сектора). *Полную размерность* $\mathcal{D}$ можно определить, просуммировав по всем типам; в случае неабелева сверхпроводника мы имеем

$$\mathcal{D}^2 = \sum_{\alpha} \sum_{R^{(\alpha)}} d^2_{(\alpha, R^{(\alpha)})} = \sum_{\alpha} |\alpha|^2 \sum_{R^{(\alpha)}} |R^{(\alpha)}|^2. \tag{8.43}$$

Поскольку сумма квадратов размерностей всех неприводимых представлений конечной группы равна ее порядку, а порядок нормализатора $N(\alpha)$ ра-

вен $|G|/|\alpha|$, мы получаем

$$\mathcal{D}^2 = \sum_{\alpha} |\alpha| \cdot |G| = |G|^2, \tag{8.44}$$

а полная размерность $\mathcal{D} = |G|$.

В случае $G = S_3$ существует восемь типов частиц, которые перечислены ниже:

Тип	Поток	Заряд	Размерность
A	e	$[+]$	1
B	e	$[-]$	1
C	e	[2]	2
D	(12)	$[+]$	3
E	(12)	$[-]$	3
F	(123)	[1]	2
G	(123)	$[\omega]$	2
H	(123)	$[\bar{\omega}]$	2

Если поток тривиален (e), тогда заряд может быть одним из трех неприводимых представлений группы S_3: тривиальным одномерным представлением $[+]$, нетривиальным одномерным представлением $[-]$ или двумерным представлением [2]. Если поток представляет собой два-цикл, то нормализатором является Z_2, а зарядом может быть либо тривиальное представление $[+]$, либо нетривиальное $[-]$. Если же потоком является 3-цикл, тогда нормализатором будет Z_3, а заряд может быть либо тривиальным представлением [1], либо нетривиальным представлением $[\omega]$, либо сопряженным ему представлением $[\bar{\omega}]$. Вы можете проверить, что полная размерность, как и ожидалось, равна $\mathcal{D} = |S_3| = 6$.

Отметим, что, поскольку a по определению коммутирует со всеми элементами нормализатора $N(a)$, матрица $\mathbf{D}^{R^{(a)}}(a)$, представляющая a в неприводимом представлении $R^{(a)}$, коммутирует со всеми матрицами в этом представлении; следовательно, согласно лемме Шура она кратна единице:

$$\mathbf{D}^{R^{(a)}}(a) = \exp(i\theta_{R^{(a)}})\mathbf{1}. \tag{8.45}$$

Чтобы понять смысл фазы $\exp(i\theta_{R^{(a)}})$, рассмотрим композит поток–заряд, в котором чарджион в представлении $R^{(a)}$ привязан к потоку a, и представим поворот против часовой стрелки этого композитного объекта на 2π.

Этот поворот переносит заряд вокруг потока, генерируя фазу

$$e^{-2\pi i\mathbf{J}} = e^{i\theta_{R^{(a)}}}; \tag{8.46}$$

следовательно, каждому суперотборному сектору соответствует определенное значение *топологического спина*, определяемое как $\theta_{R^{(a)}}$.

При слиянии двух разных типов частиц могут получиться композитные объекты различных типов, а какие типы возможны, устанавливают *правила слияния* данной теории. Квант потока композита может принадлежать любому из классов сопряженных элементов, которые можно получить перемножением представителей классов, соответствующих двум составляющим рассматриваемого объекта. Определение заряда композита особенно сложно, поскольку для этого необходимо разложить тензорное произведение представлений двух разных нормализаторов на сумму представлений нормализатора результирующего потока. Например, в случае $G = S_3$ правило, управляющее слиянием двух частиц типа D, выглядит следующим образом:

$$\mathrm{D} \times \mathrm{D} = \mathrm{A} + \mathrm{C} + \mathrm{F} + \mathrm{G} + \mathrm{H}. \tag{8.47}$$

Мы уже отмечали, что слияние пары потоков, сопоставляемых два-циклам, может дать либо тривиальный суммарный поток, либо поток соответствующий три-циклу. Заряд композита с тривиальным суммарным потоком может быть либо [+], либо [2]. Если же суммарный поток, соответствует три-циклу, то собственные зарядовые состояния в точности совпадают с собственными состояниями оператора сплетения, построенными в уравнении (8.33).

Почему собственные состояния полного заряда системы двух анионов должны также являться и собственными состояниями оператора сплетения? Мы можем понять эту связь в более общем смысле, если подумаем об угловом моменте двух-анионного составного объекта. Оператор монодромии $\mathbf{R}^2$ заключает в себе результат обхода против часовой стрелки одной частицы вокруг другой. Этот обход — практически то же самое, что и поворот против часовой стрелки композитной системы на угол 2π, за исключением того, что при повороте композитной системы вращаются и обе ее составляющие. Вращение этих составляющих будет компенсироваться, если поворот композита против часовой стрелки будет сопровождаться поворотом его составляющих по часовой стрелке. Следовательно, оператор монодромии можно представить в виде

$$(\mathbf{R}^c_{ab})^2 = e^{-2\pi i\mathbf{J}_c} e^{2\pi i\mathbf{J}_a} e^{2\pi i\mathbf{J}_b} = e^{i(\theta_c - \theta_a - \theta_b)}, \tag{8.48}$$

где $\mathbf{R}^c_{ab}$ обозначает оператор сплетения для перестановки против часовой стрелки частиц типа a и b, соединеных в композит типа c. Здесь мы поль-

зуемся более коротким, чем раньше, обозначением, в котором a, b, c представляют набор меток для суперотборных секторов (определяющих в модели неабелева сверхпроводника и поток, и заряд). Поскольку каждый суперотборный сектор имеет определенный топологический спин, а оператор монодромии диагонален в базисе топологического спина, мы видим, что собственные состояния заряда и оператора сплетения совпадают. Отметим, что уравнение (8.48) обобщает наши предшествующие наблюдения относительно абелевых анионов, а именно: композит из двух идентичных анионов имеет топологический спин $e^{i4\theta}$, а обменная фаза анион-антианионнной пары (с тривиальным полным спином) равна $e^{-i\theta}$.

8.11. Квантовые вычисления с неабелевыми флаксонами

Модель анионов характеризуется ответами на два основных вопроса: (1) Что происходит при комбинировании двух анионов (в чем состоят *правила слияния*)? (2) Что происходит при перестановке двух анионов (в чем состоят *правила сплетения*)? Мы обсудили ответы на эти вопросы в частном случае модели неабелева сверхпроводника, ассоциируемой с конечной неабелевой группой G, а сейчас мы хотели бы понять, как эти правила слияния и сплетения можно применить для моделирования квантовых схем.

Описывая это моделирование, мы будем предполагать следующие физические возможности:

Рождение и идентификация пары. Мы можем создавать пары частиц и определять тип частиц каждой пары [класс сопряженных элементов α потока каждой частицы в паре, а также заряд частицы — неприводимое представление $R^{(\alpha)}$ нормализатора группы потока $N(\alpha)$]. Это предположение естественно, поскольку не существует симметрии, связывающей частицы разных типов; они имеют различимые физические свойства — например, различные энергетические щели и эффективные массы. В самом деле, единственные типы частиц, которые нам потребуются, — это не имеющие заряда флаксоны и не несущие потока чарджионы.

Аннигиляция пары. Мы можем сталкивать две частицы друг с другом и наблюдать, аннигилирует ли эта пара полностью. Таким образом, мы получаем ответ на вопрос: имеет эта пара частиц тривиальный поток и заряд или нет? Это предположение естественно, поскольку если пара имеет нетривиальное значение какой-либо сохраняющейся величины, то после ее слияния должно остаться локализованное возбуждение, а эта остаточная частица, в принципе, обнаружима.

Сплетение. Мы можем вести частицы вдоль установленных траекторий и, таким образом, выполнять их перестановки. Квантовые вентили будут моделироваться выбором мировых линий частиц, реализующих конкретные косы.

Эти элементарные операции позволяют реализовать некоторые выводимые ниже возможности, которыми мы будем регулярно пользоваться. Во-первых, мы можем использовать чарджионы для калибровки флаксонов и учредить бюро стандартов для потоков. Предположим, нам вручили две пары флаксонов в состояниях $|a, a^{-1}\rangle$ и $|b, b^{-1}\rangle$. Мы хотим сравнить потоки a и b. Для этого создадим пару чарджион-античарджион, где заряд чарджиона представляет собой неприводимое представление R группы G. Затем пронесем чарджион вдоль замкнутого пути, окружающего первый элемент первой пары флаксонов и второй элемент второй пары флаксонов. И наконец, вновь соединим чарджион с античарджионом и посмотрим, аннигилируют они или нет. Поскольку ограниченный траекторией чарджиона суммарный поток равен ab^{-1}, пара чарджион-античарджион аннигилирует с вероятностью

$$\text{Prob}(0) = \left| \frac{\chi^R(ab^{-1})}{|R|} \right|^2, \tag{8.49}$$

что меньше единицы, если поток ab^{-1} не равен единице (при условии, что представление R не одномерное и представляет ab^{-1} нетривиально). Таким образом, если аннигиляция не происходит, мы знаем наверняка, что потоки a и b не совпадают; если же чарджионные пары всякий раз аннигилируют, то равенство a и b становится все более вероятным. Повторив эту процедуру умеренное количество раз, можно с высокой статистической достоверностью сделать вывод о равенстве a и b.

Эта процедура позволяет рассортировать пары флаксонов по корзинам так, чтобы все пары, лежащие в одной корзине, имели одинаковый квант потока. Если корзина содержит n пар, то в общем случае ее состояние представляет собой смесь состояний вида

$$\sum_{a \in G} \psi_a |a, a^{-1}\rangle^{\otimes n}. \tag{8.50}$$

После удаления из корзины лишь одной пары каждое такое состояние становится смесью

$$\sum_{a \in G} \rho_a (|a, a^{-1}\rangle\langle a, a^{-1}|)^{\otimes (n-1)}; \tag{8.51}$$

каждую корзину можно рассматривать как содержащую $(n-1)$ пару с одним и тем же определенным, но пока неизвестным потоком.

«Кто есть кто» в этих корзинах? Мы хотим маркировать корзины элементами группы G. Чтобы прийти к согласованной маркировке, извлечем по паре флаксонов из трех разных корзин. Предположим, что эти три пары имеют вид $|a, a^{-1}\rangle$, $|b, b^{-1}\rangle$ и $|c, c^{-1}\rangle$, и попытаемся проверить, справедливо ли равенство $c = ab$. Для этого создадим пару чарджион-античарджион и пронесем чарджион вдоль замкнутого пути, окружающего первый элемент первой пары флаксонов, первый элемент второй пары флаксонов и второй элемент третьей пары флаксонов, после чего посмотрим, аннигилирует вновь объединенная чарджионная пара или нет. Поскольку суммарный поток, окруженный траекторией чарджиона равен abc^{-1}, повторяя эту процедуру, можно с высокой статистической уверенностью определить, равны ли ab и c. Эти наблюдения позволяют маркировать корзины в соответствии с групповым правилом композиции. Такая маркировка является однозначной, за исключением групповых автоморфизмов (а неоднозначности, возникающие от любых автоморфизмов, можно разрешить произвольным образом).

Как только бюро стандартов для потоков создано, его можно использовать для измерения неизвестного потока немаркированной пары. Если состояние измеряемой пары — $|d, d^{-1}\rangle$, мы можем извлечь из корзины маркированную пару $|a, a^{-1}\rangle$ и использовать чарджионные пары для измерения потока ad^{-1}. Повторяя эту процедуру с другими маркированными потоками и, в конечном счете, осуществляя проецирующее измерение потока, мы можем определить значение d.

Для моделирования квантовых схем с помощью флаксонов нам потребуется выполнять логические вентили, действующие на значение потока. Основной вентиль, который мы будем использовать, реализуется путем оборота против часовой стрелки пары флаксонов в состоянии $|a, a^{-1}\rangle$ вокруг первого элемента другой пары флаксонов в состоянии $|b, b^{-1}\rangle$. Поскольку пара $|a, a^{-1}\rangle$ имеет тривиальный полный поток, данная процедура никак не влияет на пару $|b, b^{-1}\rangle$. Но поскольку в результате поток b обходит против часовой стрелки вокруг обоих элементов пары, начальным состоянием которой было $|a, a^{-1}\rangle$, то эта пара преобразуется как

$$|a, a^{-1}\rangle \longmapsto |bab^{-1}, ba^{-1}b^{-1}\rangle. \tag{8.52}$$

Будем называть это преобразование *операцией сопряжения*, действующей на пару флаксонов.

Подведем итог вышесказанному. Наши исходные и выведенные возможности позволяют: (1) выполнять проецирующее измерение потока,

(2) выполнять разрушающее измерение, определяющее, являются ли поток и заряд пары тривиальными и (3) осуществлять сопрягающий элемент. Сейчас мы должны обсудить, как моделировать квантовые схемы, используя эти возможности.

Нужно решить, как кодировать кубиты с помощью флаксонов. Подходящее кодирование можно выбрать разными способами; мы остановимся на одном конкретном выборе, иллюстрирующем ключевые идеи, а именно: будем кодировать кубит, используя пару флаксонов, полный поток которой тривиален. Выберем два некоммутирующих элемента $a, b \in G$, где $b^2 = e$, и вычислительный базис для кубита

$$|\bar{0}\rangle = |a, a^{-1}\rangle, \quad |\bar{1}\rangle = |bab^{-1}, ba^{-1}b^{-1}\rangle. \tag{8.53}$$

Важный нюанс: отдельный изолированный флаксон с квантом потока a выглядит идентичным флаксону с сопряженным квантом потока bab^{-1}. Следовательно, если два флаксона из одной пары удерживать далеко друг от друга, локальные взаимодействия с окружающей средой не вызовут декогерентизации суперпозиции состояний $|\bar{0}\rangle$ и $|\bar{1}\rangle$. Квантовая информация защищена от повреждения, поскольку она хранится нелокальным образом, благодаря топологическому вырождению состояний, в которых флаксон и антифлаксон закреплены в фиксированных пространственно разделенных позициях.

Но в отличие от топологического вырождения в системах с абелевыми анионами, этот защищенный кубит можно сравнительно просто измерить, не прибегая к тонким интерферометрическим процедурам, выделяющим фазы Ааронова–Бома. Мы уже описывали, как измерить поток, используя предварительно откалиброванные флаксоны; следовательно, мы можем выполнить проецирующее измерение закодированного оператора Паули $\bar{\mathbf{Z}}$ (проекцию на базис $\{|\bar{0}\rangle, |\bar{1}\rangle\}$). Мы также можем измерить дополнительный оператор Паули $\bar{\mathbf{X}}$, пусть даже разрушающим методом и неидеально. Собственными состояниями оператора $\bar{\mathbf{X}}$ являются

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|\bar{0}\rangle \pm |\bar{1}\rangle) \equiv \frac{1}{\sqrt{2}}(|a, a^{-1}\rangle \pm |bab^{-1}, ba^{-1}b^{-1}\rangle); \tag{8.54}$$

следовательно, состояние $|-\rangle$ *ортогонально* состоянию нулевого заряда

$$|0; \alpha\rangle = \frac{1}{\sqrt{|\alpha|}} \left(\sum_{c \in \alpha} |c, c^{-1}\rangle \right), \tag{8.55}$$

где α — класс сопряженных элементов, содержащий a. С другой стороны, состояние $|+\rangle$ имеет ненулевое перекрытие с $|0;\alpha\rangle$

$$\langle +|0;\alpha\rangle = \sqrt{2/|\alpha|}. \tag{8.56}$$

Следовательно, при столкновении частиц, образующих флаксонную пару, полная аннигиляция невозможна, если пара находится в состоянии $|-\rangle$, если же состоянием пары является $|+\rangle$, то аннигиляция происходит с вероятностью $\mathrm{Prob}(0) = 2/|\alpha|$.

Отметим, что можно также приготовить флаксонную пару в состоянии $|+\rangle$. Один из возможных способов состоит в следующем: сначала создается пара в состоянии $|0;\alpha\rangle$. Если α содержит только два элемента a и bab^{-1}, то задача решена. В противном случае вновь созданная пара сравнивается с калиброванными парами в каждом из состояний $|c, c^{-1}\rangle$, где $c \in \alpha$ и отличается как от a, так и от bab^{-1}. Если она не согласуется ни с одной из этих $|c, c^{-1}\rangle$-пар, то ее состоянием должно быть $|+\rangle$.

Чтобы продвинуться дальше, необходимо определить вычислительные возможности операции сопряжения. Будем использовать более компактное обозначение, в котором состояние $|x, x^{-1}\rangle$ флаксонной пары обозначается просто как $|x\rangle$, и рассмотрим преобразования состояния $|x, y, z\rangle$, которые можно построить с помощью операции сопряжения. Пронося (по замкнутому пути) третью пару сквозь первую против или по часовой стрелке, можно реализовать следующие вентили:

$$|x, y, z\rangle \longmapsto |x, y, xzx^{-1}\rangle, \quad |x, y, z\rangle \longmapsto |x, y, x^{-1}zx\rangle. \tag{8.57}$$

Точно так же, пронося третью пару сквозь вторую против или по часовой стрелки, можно реализовать вентили

$$|x, y, z\rangle \longmapsto |x, y, yzy^{-1}\rangle, \quad |x, y, z\rangle \longmapsto |x, y, y^{-1}zy\rangle. \tag{8.58}$$

Более того, позаимствовав в бюро стандартов пару с квантом потока $|c\rangle$, можно осуществить

$$|x, y, z\rangle \longmapsto |x, y, czc^{-1}\rangle \tag{8.59}$$

для любой постоянной величины $c \in G$. Комбинация этих элементарных операций позволяет реализовать любой вентиль вида

$$|x, y, z\rangle \longmapsto |x, y, fzf^{-1}\rangle, \tag{8.60}$$

где функцию $f(x, y)$ можно представить в *мультипликативной форме*, то есть в виде конечного произведения, сомножителями которого могут быть

входящие данные x и y, обратные им значения x^{-1} и y^{-1} или постоянные элементы группы G, причем каждый из них может встречаться в произведении любое количество раз.

Что представляют собой функции $f(x,y)$, которые можно представить в таком виде? Ответ зависит от структуры группы G, но для наших целей достаточно следующей характеристики. Вспомним, что подгруппа H конечной группы G является *нормальной*, если для любой величины $h \in H$ и $g \in G$, $ghg^{-1} \in H$;[1] вспомним также, что конечная группа G является *простой*, если G не имеет нормальных подгрупп за исключением самой группы G и тривиальной подгруппы $\{e\}$. Оказывается, если G является простой неабелевой конечной группой, то *любую* функцию $f(x,y)$ можно представить в мультипликативной форме. В литературе по теории вычислительных систем родственный этому утверждению результат часто называют *теоремой Баррингтона*.

В частности, если группа G представляет собой неабелеву простую группу, то существует представимая в мультипликативной форме функция f, такая, что

$$f(a,a) = f(a,bab^{-1}) = f(bab^{-1},a) = e, \quad f(bab^{-1},bab^{-1}) = b. \qquad (8.61)$$

Таким образом, для $x,y,z \in \{a,bab^{-1}\}$ действие уравнения (8.60) «переворачивает» поток третьей пары, если и только если $x = y = bab^{-1}$; из наших элементарных операций мы построили вентиль Тоффоли в вычислительном базисе. Следовательно, для реализации универсальных обратимых *классических* вычислений достаточно операций сопряжения, действующих на стандартные базисные состояния.

A_5 — неабелева простая группа минимального порядка $|A_5| = 60$, группа четных перестановок пяти объектов. Следовательно, одна конкретная реализация универсальных классических вычислений, использующих операции сопряжения, получается при выборе в качестве a три-цикла элемента $a = (345) \in A_5$, а в качестве b — произведения пары два-циклов $b = (12)(34) \in A_5$, так что $bab^{-1} = (435)$.

При таком разумном выборе группы G мы добиваемся топологической реализации универсальных классических вычислений, но как продвинуться еще дальше и осуществить универсальные квантовые вычисления? Мы имеем возможность готовить состояния вычислительного базиса, измерять в вычислительном базисе и выполнять вентили Тоффоли, но все это чисто классические средства. Единственными неклассическими приемами, которыми мы располагаем, являются способность готовить собственные состо-

[1] Такие подгруппы называют также *инвариантными*, или *нормальными, делителями*.

яния $\bar{\mathbf{X}} = 1$ и умение выполнять неидеальное разрушающее измерение $\bar{\mathbf{X}}$. К счастью, этих дополнительных возможностей оказывается достаточно.

В предыдущем обсуждении квантовой отказоустойчивости мы отмечали, что если можно выполнять классические вентили Тоффоли и CNOT, то для универсальных квантовых вычислений достаточно уметь применять каждый из операторов Паули $\mathbf{X}$, $\mathbf{Y}$ и $\mathbf{Z}$, а также выполнять проецирующие измерения любого из этих операторов. Мы уже знаем, как применить классический вентиль $\mathbf{X}$ и как измерить $\mathbf{Z}$ (то есть спроецировать на вычислительный базис). В нашем арсенале до сих пор отсутствуют проецирующие измерения $\mathbf{X}$ и $\mathbf{Y}$ и выполнение $\mathbf{Z}$. (Конечно, если мы можем применять операторы $\mathbf{X}$ и $\mathbf{Z}$, то в состоянии применять и их произведение $\mathbf{ZX} = i\mathbf{Y}$.)

Посмотрим, как превратить неидеальное разрушающее измерение $\mathbf{X}$ в надежное проецирующее измерение. Вспомним действие сопряжения CNOT на операторы Паули:

$$\text{CNOT} : \mathbf{X1} \longmapsto \mathbf{XX}, \tag{8.62}$$

где первый кубит является управляющим, а второй — целью CNOT. Следовательно, для осуществления проецирующего измерения $\mathbf{X}$ достаточно вентилей CNOT, а также способности готовить собственные состояния $\mathbf{X} = 1$ и выполнять разрушающие измерения $\mathbf{X}$. Мы можем приготовить служебный кубит в собственном состоянии $\mathbf{X} = 1$, выполнить вентиль CNOT со служебным кубитом в качестве управляющего и измеряемыми данными в качестве цели, а затем провести разрушающее измерение этого служебного кубита. Это измерение готовит данные в собственном состоянии оператора $\mathbf{X}$, собственное значение которого совпадает с результатом измерения служебного кубита. В нашем случае разрушающее измерение не вполне надежно, но его можно многократно повторить. Каждый раз мы готовим и измеряем новый служебный бит и после нескольких повторений получаем результат измерения с приемлемой статистической достоверностью.

Теперь, когда мы можем выполнить проецирующее измерение $\mathbf{X}$, мы способны приготавливать не только собственные состояния $\mathbf{X} = 1$, но и собственные состояния $\mathbf{X} = -1$ (например, повторяя следующие друг за другом измерения операторов $\mathbf{Z}$ и $\mathbf{X}$ до тех пор, пока в конце концов не будет получен результат $\mathbf{X} = -1$). Затем, выполняя вентиль CNOT, целью которого является собственное состояние $\mathbf{X} = -1$, мы можем выполнить оператор Паули $\mathbf{Z}$, действующий на управляющий кубит. Остается лишь продемонстрировать, что измерение $\mathbf{Y}$ тоже может быть реализовано.

На первый взгляд, измерение оператора $\mathbf{Y}$ выглядит проблематичным, поскольку наши физические возможности не позволяют различать собственные состояния $\mathbf{Y} = 1$ и $\mathbf{Y} = -1$ (то есть отличать состояние ψ от

его комплексно-сопряженного ψ^*). Но эта неопределенность не представляет серьезной трудности, поскольку на самом деле не важно, как она будет разрешена. Если в процедуре моделирования унитарного преобразования $\mathbf{U}$ заменить измерение оператора $\mathbf{Y}$ на измерение оператора $-\mathbf{Y}$, то результатом такой замены будет моделирование $\mathbf{U}^*$ вместо $\mathbf{U}$; эта замена не меняет распределение вероятностей результатов измерений в стандартном вычислительном базисе.

Чтобы быть точнее, можно сформулировать протокол измерения оператора $\mathbf{Y}$, в первую очередь отметив, что применение вентиля Тоффоли, целевым кубитом которого является собственное состояние $\mathbf{X} = -1$, реализует вентиль контролируемой фазы $\Lambda(\mathbf{Z})$, действующий на два управляющих кубита. Объединив этот вентиль с CNOT вентилем $\Lambda(\mathbf{X})$, мы получаем вентиль $\Lambda(i\mathbf{Y})$, действующий как

$$\Lambda(i\mathbf{Y}): \quad \begin{cases} |X=+1\rangle \otimes |Y=+1\rangle \longmapsto |Y=+1\rangle \otimes |Y=+1\rangle, \\ |X=+1\rangle \otimes |Y=-1\rangle \longmapsto |Y=-1\rangle \otimes |Y=-1\rangle, \\ |X=-1\rangle \otimes |Y=+1\rangle \longmapsto |Y=-1\rangle \otimes |Y=+1\rangle, \\ |X=-1\rangle \otimes |Y=-1\rangle \longmapsto |Y=+1\rangle \otimes |Y=-1\rangle, \end{cases} \tag{8.63}$$

где первый кубит является управляющим, а второй — целью. Теперь предположим, что мой надежный друг дает мне только один кубит, который, как он меня уверяет, был приготовлен в состоянии $|Y=1\rangle$. Я знаю, как самостоятельно приготовить состояния $|X=1\rangle$, и могу выполнять вентили $\Lambda(i\mathbf{Y})$; следовательно, поскольку вентиль $\Lambda(i\mathbf{Y})$ с состоянием $|Y=1\rangle$ в качестве его цели преобразует $|X=1\rangle$ в $|Y=1\rangle$, я могу сделать множество копий полученного от моего друга состояния $|Y=1\rangle$. Желая измерить $\mathbf{Y}$, я применяю обратный по отношению к $\Lambda(i\mathbf{Y})$ вентиль, целью которого является измеряемый кубит, а управляющим кубитом — одно из моих состояний $\mathbf{Y}=1$; затем я выполняю $\mathbf{X}$-измерение служебного кубита, считывающие результат $\mathbf{Y}$-измерения другого кубита.

А что, если мой друг солгал и вместо этого дал мне копию состояния $|Y=-1\rangle$? Тогда я сделаю множество копий состояния $|Y=-1\rangle$ и буду измерять $-\mathbf{Y}$, считая, что измеряю $\mathbf{Y}$. Мое моделирование будет работать точно так же, как и раньше; фактически я буду моделировать комплексно-сопряженное значение идеальной схемы, но это не изменит конечного результата квантового вычисления. Если мой друг подбросит в воздух монетку, чтобы решить, дать ли мне состояние $|Y=1\rangle$ или $|Y=-1\rangle$, это также не окажет никакого влияния на точность моего моделирования. То есть получается, что помощь друга мне вообще не нужна — вместо использования состояния $|Y=1\rangle$, которое я мог бы получить от него, я могу

использовать случайное состояние $\rho = 1/2$ (равновзвешенную смесь состояний $|Y = 1\rangle$ и $|Y = -1\rangle$), способ приготовления которого мне известен.

Этим завершается доказательство того, что по крайней мере в случае, когда G является простой неабелевой конечной группой,[1] мы можем эффективно моделировать отказоустойчивые квантовые схемы, используя флаксоны и чарджионы неабелева сверхпроводника. Рассматриваемое в целом, включая все приготовления состояний и калибровку потоков, это моделирование можно описать следующим образом: приготавливается множество пар анионов (флаксонов и чарджионов), мировые линии анионов следуют определенной косе, анионные пары сталкиваются с целью увидеть, аннигилируют они или нет. Это моделирование недетерминированно в том смысле, что реализованная анионами фактическая коса зависит от результатов измерений, выполненных (благодаря столкновениям) в процессе моделирования. Оно устойчиво к малым возмущениям, если температура низка по сравнению с энергетической щелью, а частицы удерживаются достаточно далеко друг от друга (за исключением момента рождения и столкновения пар), чтобы подавить обмен виртуальными анионами. Пока сплетение частиц принадлежит правильному топологическому классу, малые деформации их мировых линий не влияют на результат вычислений.

8.12. Обобщенные анионные модели

Наше обсуждение модели неабелева сверхпроводника обеспечивает доказательство существования использующих анионы отказоустойчивых квантовых вычислений. Но эта модель, естественно, имеет недостатки. Описанной нами схеме недостает красоты, элегантности, простоты.

Я рассмотрел эту модель так подробно, поскольку она достаточно конкретна и, следовательно, позволяет нам приобрести интуицию в понимании свойств неабелевых анионов. Но сейчас, когда мы уже лучше понимаем ключевые идеи сплетения и слияния в анионных моделях, мы готовы подойти к этим вопросам с более общих и абстрактных позиций. Это приведет нас к новым моделям, в том числе и к существенно более простым по сравнению рассмотренными ранее. Мы сможем выбросить большую часть лишнего багажа, обременяющего модель неабелева сверхпроводника (например, различие между флаксонами и чарджионами, калибровку флаксонов, а также измерения, требуемые для моделирования неклассических

[1] Мошон показал, что универсальные квантовые вычисления возможны для более широкого класса групп. См.: C. Mochon, *Anyons from Non-Solvable Finite Groups are Sufficient for Universal Quantum Computation*, Phys. Rev. **A67**, 022315 (2003).

вентилей). Простейшие модели, с которыми мы сейчас познакомимся, органичнее вписываются в схемы отказоустойчивых вычислений и выглядят более правдоподобными с точки зрения их возможной реализации в системах с разумными физическими свойствами.

Анионная модель — это теория частиц на двумерной поверхности (которую мы будем считать плоской), переносящих локально сохраняющиеся *заряды*. Также предполагается, что эта теория имеет массовую щель, то есть дальнодействующие взаимодействия, переносимые безмассовыми частицами, отсутствуют. Эта модель характеризуется тремя определяющими свойствами:

1. Перечень *типов* частиц. Типами являются метки, устанавливающие возможные значения сохраняющегося заряда, который может нести частица.

2. Правила *композиции* и *расщепления*, определяющие возможные значения заряда, которые могут быть получены при объединении двух частиц с известными зарядами, а также возможные способы расщепления на две части заряда, переносимого одной частицей.

3. Правила *сплетения*, устанавливающие, что происходит при перестановке двух частиц (или при повороте одной частицы на 2π).

Теперь обсудим каждое из этих свойств более подробно.

8.12.1. Метки

Для меток, определяющих различные типы частиц, я буду использовать латинские буквы $\{a, b, c, \ldots\}$. (В случае неабелева сверхпроводника метка $(\alpha, R^{(\alpha)})$ обозначала класс сопряженных элементов и неприводимое представление нормализатора класса, но теперь наше обозначение будет более компактным.) Предположим, что совокупность возможных меток конечна. Символ a представляет величину сохраняющегося заряда, переносимого частицей. Иногда мы говорим, что эта метка определяет *суперотборный сектор* теории. Этот термин означает лишь то, что метка a является свойством локализованного объекта, которое нельзя изменить локальным физическим процессом, то есть если одна частица всегда хорошо изолирована от других, ее метка никогда не изменится. В частности, никакие локальные взаимодействия частицы с ее окружением не могут изменить ее метку. Это локальное сохранение заряда является основной причиной того, почему анионы подходят для отказоустойчивой обработки квантовой информации.

Существует одна частная, единичная метка 1. Наличие частицы с меткой 1 фактически эквивалентно отсутствию частицы вообще. Более того, для каждой метки частицы a существует сопряженная метка $\bar{a}$, а также существует операция зарядового сопряжения C (где $C^2 = I$), действие которой отображает метки на сопряженные им:

$$C : a \longmapsto \bar{a} \longmapsto a. \tag{8.64}$$

Метка может быть самосопряженной, так что $\bar{a} = a$. Например, $\bar{1} = 1$.

Нам потребуется рассматривать определенным образом упорядоченные состояния n-частиц. Их удобно представлять расположенными друг за другом вдоль некоторой линии (например, вдоль вещественной оси) слева направо; n частиц помечены как $(a_1, a_2, a_3, \ldots, a_n)$, где a_1 присвоена крайней слева частице, а a_n — крайней справа.

8.12.2. Пространства композитных состояний

Композитный объект, образующийся при объединении двух частиц, также имеет заряд. Возможные значения полного заряда c при значениях зарядов составляющих a и b определяют *правила композиции* модели. Их можно записать как

$$a \times b = \sum_c N_{ab}^c \, c, \tag{8.65}$$

где каждое N_{ab}^c — неотрицательное целое число, а суммирование ведется по полному набору меток c. Отметим, что a, b и c — это метки, а не векторные пространства; произведение слева не является тензорным произведением, а сумма справа не является прямой суммой. Скорее правила композиции можно рассматривать как абстрактное соотношение для набора меток, отображающее упорядоченную тройку $(a, b; c)$ на N_{ab}^c. Это соотношение симметрично по a и b ($a \times b = b \times a$) — возможные заряды композита не зависят от положения a (справа или слева). Прочитанные в обратном направлении, правила композиции определяют возможные способы расщепления заряда c на две части с зарядами a и b.

Если $N_{ab}^c = 0$, то при комбинировании a и b заряд c не может быть получен. Если $N_{ab}^c = 1$, то c *можно* получить единственным способом. Если $N_{ab}^c > 1$, то c можно получить N_{ab}^c различимыми способами. Представление о том, что объединение двух зарядов может дать третий заряд несколькими способами, должно быть знакомо из теории представлений групп. Например, правило, управляющее композицией двух октетных представлений группы SU(3), выглядит следующим образом:

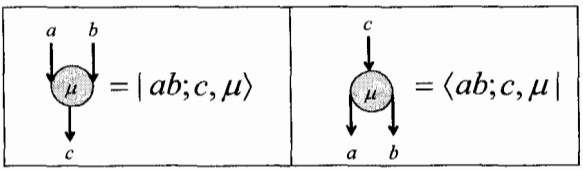
$$8 \times 8 = 1 + 8 + 8 + 10 + \overline{10} + 27, \tag{8.66}$$

так что $N_{88}^{8} = 2$. Но подчеркнем еще раз: в то время как правила композиции представлений групп можно интерпретировать как разложение тензорного произведения векторных пространств на прямую сумму векторных пространств, в общем случае правила композиции в анионной модели такой интерпретации не имеют.

N_{ab}^{c} различимых способов, которыми при объединении a и b может возникнуть c, можно рассматривать как ортонормированные базисные состояния гильбертова пространства V_{ab}^{c}. Мы называем V_{ab}^{c} *пространством композитных состояний*, а состояния, которые оно включает, — *композитными состояниями*. Базисные векторы V_{ab}^{c} можно обозначить как

$$\{|ab; c, \mu\rangle, \quad \mu = 1, 2, \ldots, N_{ab}^{c}\}. \tag{8.67}$$

Для состояний композитного базиса удобно ввести графическое обозначение:



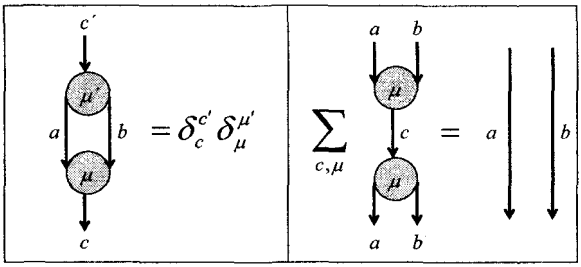
Состояние $|ab; c, \mu\rangle$ представлено в виде окружности, содержащей символ μ; к окружности присоединены обозначенные метками a и b ориентированные входящие линии, представляющие объединяемые заряды, и маркированная меткой c выходящая ориентированная линия, представляющая результат объединения. Существует дуальное векторное пространство V_{c}^{ab}, описывающее состояния, возникающие при расщеплении заряда c на заряды a и b, и дуальный базис с инвертированными линиями (входящей c и выходящими a и b). Пространства V_{ab}^{c} с различными значениями c взаимно ортогональны, так что элементы базиса композитных состояний удовлетворяют уравнению

$$\langle ab; c', \mu' | ab; c, \mu\rangle = \delta_{c}^{c'} \delta_{\mu}^{\mu'}, \tag{8.68}$$

а полноту базиса композитных состояний можно выразить как

$$\sum_{c,\mu} |ab; c, \mu\rangle\langle ab; c, \mu| = I_{ab}, \tag{8.69}$$

где I_{ab} обозначает проектор на пространство $\bigoplus_{c} V_{ab}^{c}$, полное гильбертово пространство для анионной пары ab.



Среди пространств композитных состояний существует несколько естественных изоморфизмов. Во-первых, $V_{ab}^c \cong V_{ba}^c$; эти векторные пространства ассоциируются с различными маркировками двух частиц (при $a \neq b$) и, следовательно, должны рассматриваться как различимые, но они являются изоморфными пространствами, поскольку объединение симметрично. Мы можем также «поднимать и опускать индексы» пространства композитных состояний, заменяя метку ее сопряженной величиной, например,

$$V_{ab}^c \cong V_{a\bar{c}}^{\bar{b}} \cong V_{ab\bar{c}}^1 \cong V_a^{\bar{b}c} \cong V_{\bar{c}}^{\bar{a}\bar{b}} \cong \cdots ; \tag{8.70}$$

на графическом языке этому соответствует обращение соответствующих ориентированных линий с одновременным сопряжением их меток. Пространство $V_{ab\bar{c}}^1$, изображаемое диаграммой с тремя входящими линиями, представляет собой пространство, натянутое на базис различимых способов получения тривиального полного заряда 1 при композиции трех частиц с метками a, b и $\bar{c}$.

Заряд 1 достоин своего имени, поскольку с другими частицами он комбинируется тривиальным образом:

$$a \times 1 = a. \tag{8.71}$$

Вследствие изоморфизма $V_{a1}^a \cong V_{a\bar{a}}^1$, мы делаем вывод, что $\bar{a}$ — единственная метка, которая дает 1 при композиции с a, и что эта композиция может возникнуть лишь единственном способом. Аналогично, $V_{a1}^a \cong V_1^{a\bar{a}}$ означает, что рожденные из вакуума пары частиц имеют сопряженные заряды.

Анионная модель является *неабелевой*, если

$$\dim\left(\bigoplus_c V_{ab}^c\right) = \sum_c N_{ab}^c \geqslant 2 \tag{8.72}$$

по крайней мере для некоторых пар с метками ab; в противном случае модель является *абелевой*. В абелевой модели любые две частицы объеди-

няются единственным способом, а в неабелевой модели существуют некоторые пары частиц, способные объединяться более чем одним способом, и существует гильбертово пространство размерностью два и более, натянутое на эти различимые состояния. Мы будем говорить об этом пространстве как о «топологическом гильбертовом пространстве» анионной пары, подчеркивая тем самым, что данная квантовая информация закодирована нелокально — это коллективное свойство пары, не локализованное ни на какой из образующих ее частиц. Действительно, когда две частицы с метками a и b находятся далеко друг от друга, разные состояния в топологическом гильбертовом пространстве локально выглядят идентичными. Следовательно, эта квантовая информация хорошо спрятана и неуязвима для декогерентизации, обусловленной локальными взаимодействиями с окружающей средой.

Именно по этой причине мы предлагаем использовать неабелевы анионы в работе квантового компьютера. Конечно, нелокально закодированная информация спрятана не только от окружения; она недоступна и для нас. Однако с помощью неабелевых анионов можно убить двух зайцев сразу! По завершении квантового вычисления, когда мы готовы считать результат, мы можем объединить анионы в пары и пронаблюдать результат такой композиции. В сущности, будет достаточно отличать случай, когда заряд композита $c = 1$, от случая $c \neq 1$, то есть отличать остаточную частицу (неспособную распадаться вследствие ее нетривиального сохраняющегося заряда) от отсутствия частицы вообще.

Отметим, что для каждой анионной пары это топологическое гильбертово пространство конечномерно. Обладающая этим свойством анионная модель *рациональна*. Как и в обсуждавшемся случае топологически вырожденного основного состояния абелевой модели, анионы рациональных неабелевых моделей всегда имеют топологические спины, представляющие собой корни из единицы.

8.12.3. Сплетение: R-матрица

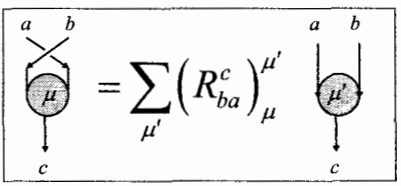
При перестановке против часовой стрелки двух частиц с метками a и b их полный заряд c не изменяется. Следовательно, поскольку две частицы меняются местами на линии, обмен индексами является естественным изоморфизмом, отображающим гильбертово пространство V_{ba}^c на V_{ab}^c; это отображение представляет собой оператор сплетения

$$\mathbf{R} : V_{ba}^c \to V_{ab}^c. \tag{8.73}$$

Если в этих двух пространствах выбрать канонические базисы $\{|ba;c,\mu\rangle\}$ и $\{|ab;c,\mu'\rangle\}$, то операцию (8.73) можно представить как унитарное преобразование

$$\mathbf{R}:|ba;c,\mu\rangle \longmapsto \sum_{\mu'}|ab;c,\mu'\rangle(R^c_{ab})^{\mu'}_{\mu}; \tag{8.74}$$

отметим, что $\mathbf{R}$ может оказывать нетривиальное действие на композитные состояния. Представляя действие $\mathbf{R}$ в графической форме, удобно зафиксировать положения меток a и b на входящих линиях и повернуть против часовой стрелки линии, входящие в вершину композита (μ). График с перекрещенными линиями представляет состояние в пространстве V^c_{ab}, полученное в результате применения матрицы $\mathbf{R}$ к $|ba;c,\mu\rangle$, которое может быть разложено в каноническом базисе пространства V^c_{ab}:



Оператор *монодромии*

$$\mathbf{R}^2 : V^c_{ab} \to V^c_{ab} \tag{8.75}$$

является изоморфизмом V^c_{ab} на самого себя, представляя результат полного оборота a вокруг b против часовой стрелки. Как уже отмечалось при обсуждении неабелева сверхпроводника, оператор монодромии эквивалентен повороту c на 2π с одновременным поворотом a и b на -2π; следовательно, собственные значения оператора монодромии определяются *топологическими спинами* частиц:

$$(R^c_{ab})^2 = e^{-2\pi i J_c} e^{2\pi i J_a} e^{2\pi i J_b} \equiv e^{i(\theta_c - \theta_a - \theta_b)}. \tag{8.76}$$

Более того, как и в случае абелевых анионов, топологический спин определяется оператором сплетения, действующим на пару частица-античастица с тривиальным полным зарядом:

$$e^{-i\theta_a} = R^1_{a\bar{a}} \tag{8.77}$$

(поскольку рождение пары, перестановка и аннигиляция пары эквивалентны повороту частицы на -2π).

8.12.4. Ассоциативность композитных состояний: F-матрица

Композитные состояния ассоциативны:

$$(a \times b) \times c = a \times (b \times c). \tag{8.78}$$

С математической точки зрения, это аксиома, которой удовлетворяют правила композиции анионной модели. С физической точки зрения, это необходимое условие, поскольку полный заряд системы из трех частиц — это ее внутреннее свойство, которое не должно зависеть от того, объединяются ли сначала a и b, а затем эта пара — с c, или же сначала комбинируются b и c, а затем результат композиции — с a.

Следовательно, топологическое гильбертово пространство композитных состояний с полным зарядом d, образующихся при объединении трех частиц с зарядами a, b и c, допускает два естественных способа разложения в прямую сумму тензорных произведений пространств композитных состояний пар частиц:

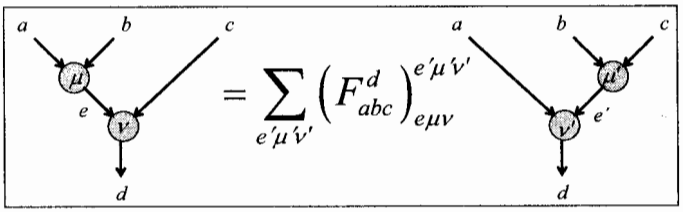
$$V_{abc}^{d} \cong \bigoplus_{e} V_{ab}^{e} \otimes V_{ec}^{d} \cong \bigoplus_{e'} V_{ae'}^{d} \otimes V_{bc}^{e'}. \tag{8.79}$$

Соответственно, в пространстве V_{abc}^{d} существует два естественных ортонормированных базиса

$$\begin{aligned} |(ab)c \to d; e\mu\nu\rangle &\equiv |ab; e, \mu\rangle \otimes |ec; d, \nu\rangle, \\ |a(bc) \to d; e'\mu'\nu'\rangle &\equiv |ae'; d, \nu'\rangle \otimes |bc; e', \mu'\rangle, \end{aligned} \tag{8.80}$$

связанных между собой унитарным преобразованием $\mathbf{F}$:

$$|(ab)c \to d; e\mu\nu\rangle = \sum_{e'\mu'\nu'} |a(bc) \to d; e'\mu'\nu'\rangle \left(F_{abc}^{d}\right)_{e\mu\nu}^{e'\mu'\nu'}. \tag{8.81}$$



Унитарные матрицы F_{abc}^{d} иногда называют *матрицами композиции*; однако чтобы не рисковать вызвать путаницу между ними и правилами композиции N_{ab}^{c}, я лучше буду называть их *F-матрицами*.

8.12.5. Множество анионов: стандартный базис

В анионном квантовом компьютере мы оперируем с n-анионным топологическим квантовым состоянием путем сплетения анионов. Чтобы описывать эти вычисления, удобно выбрать стандартный базис в таком гильбертовом пространстве.

Пусть n анионов с полным зарядом c, расположенные последовательно вдоль линии, имеют метки $a_1, a_2, a_3, \ldots, a_n$. Рассмотрим процесс,в котором сначала происходит объединение анионов 1 и 2, затем образовавшаяся пара объединяется с анионом 3, тройка — с анионом 4, и так далее. Совершаемому в этом порядке объединению соответствует разложение n-анионного топологического гильбертова пространства:

$$V^c_{a_1a_2a_3\cdots a_n} \cong \bigoplus_{b_1,b_2,\ldots,b_{n-2}} V^{b_1}_{a_1a_2} \otimes V^{b_2}_{b_1a_3} \otimes V^{b_3}_{b_2a_4} \otimes \cdots \otimes V^c_{b_{n-2}a_n}. \quad (8.82)$$

Отметим, что это пространство *не имеет* естественного разложения на тензорное произведение подсистем, соответствующих локализованным частицам; скорее, мы представили его как прямую сумму множества тензорных произведений. Для неабелевых анионов его размерность экспоненциальна по n:

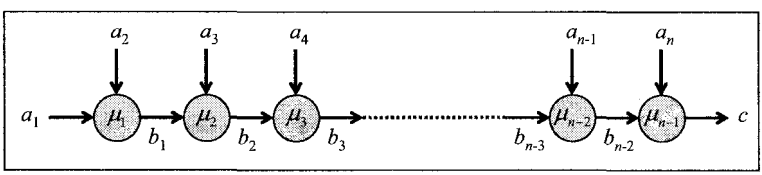
$$\dim(V^c_{a_1a_2a_3\cdots a_n}) \equiv N^c_{a_1a_2a_3\cdots a_n} = \\ = \sum_{b_1,b_2,b_3,\ldots b_{n-2}} N^{b_1}_{a_1a_2} N^{b_2}_{b_1a_3} N^{b_3}_{b_2a_4} \cdots N^c_{b_{n-2}a_n}. \quad (8.83)$$

Таким образом, топологическое гильбертово пространство является подходящей ареной для успешной обработки квантовой информации.

С этим разложением ассоциируется стандартный базис, элементы которого нумеруются промежуточными зарядами $b_1, b_2, \ldots, b_{n-2}$ и представляют собой произведения базисных векторов $\{|\mu_j\rangle\}$ пространств композитных состояний $V^{b_j}_{b_{j-1},a_{j+1}}$:

$$\{|a_1a_2; b_1, \mu_1\rangle|b_1a_3; b_2, \mu_2\rangle \cdots |b_{n-3}a_{n-1}; b_{n-2}, \mu_{n-2}\rangle|b_{n-2}a_n; c, \mu_{n-1}\rangle\}, \quad (8.84)$$

или на графическом языке:



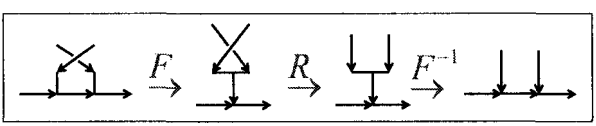
Конечно, выбор этого базиса неоднозначен. При желании можно представить процесс объединения частиц, протекающий в другом порядке, и получить другой стандартный базис, связанный с предыдущим унитарной F-матрицей.

8.12.6. Сплетение в стандартном базисе: B-матрица

Рассмотрим, что произойдет с состояниями топологического векторного пространства n анионов $V^c_{a_1a_2a_3\cdots a_n}$ при перестановке частиц друг с другом. Поскольку перестановки могут менять позиции частиц с различными метками, они могут отображать одно топологическое векторное пространство на другое путем перестановки меток. Тем не менее, можно рассмотреть прямые суммы векторных пространств, ассоциируемых со всеми возможными перестановками меток, представляющими группу кос B_n.

Попробуем описать, как это представление действует на стандартные базисы в этих пространствах. Достаточно сказать, как представляются перестановки соседних частиц, то есть определить действие генераторов группы кос. Однако до настоящего времени мы обсуждали действие группы кос лишь на пару частиц с определенным полным зарядом (R-матрица), что по сути своей не дает информации о ее действии на стандартные базисы.

Выход из этого затруднительного положения дает следующее наблюдение: применяя F-матрицу, можно перейти от стандартного базиса к базису, в котором R-матрица блочно-диагональна, применить $\mathbf{R}$, а затем, применяя $\mathbf{F}^{-1}$, вернуться в стандартный базис:



Композиция этих трех операций, выражающая результат сплетения в стандартном базисе, обозначается символом $\mathbf{B}$ и иногда называется «матрицей сплетения»; но чтобы избежать путаницы между $\mathbf{B}$ и $\mathbf{R}$, я буду называть ее просто *B-матрицей*.

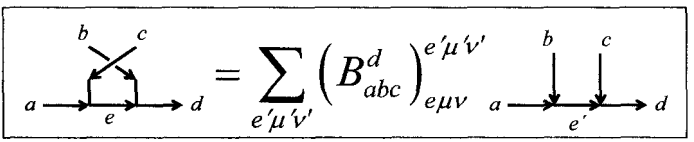
Рассмотрим перестановку анионов в позициях j и $j+1$ вдоль линии. В нашем разложении пространства $V^c_{a_1a_2a_3\cdots a_n}$ эта перестановка действует на пространство

$$V^{b_j}_{b_{j-2},a_j,a_{j+1}} = \bigoplus_{b_{j-1}} V^{b_{j-1}}_{b_{j-2},a_j} \otimes V^{b_j}_{b_{j-1},a_{j+1}}. \tag{8.85}$$

Сокращая число подстрочных индексов, будем называть это пространство V^d_{acb}, которое преобразуется перестановкой как

$$\mathbf{B} : V^d_{acb} \to V^d_{abc}. \tag{8.86}$$

Выразим действие $\mathbf{B}$ в стандартных базисах двух пространств V^d_{acb} и V^d_{abc}:



Чтобы избежать нагромождения в уравнениях, я опускаю метки для базисных элементов пространства композитных состояний (расположение этих меток очевидно). Следовательно, мы пишем

$$\begin{aligned}
\mathbf{B}|(ac)b \to d; e\rangle &= \sum_f \mathbf{B}|a(cb) \to d; f\rangle \left(F^d_{acb}\right)^f_e = \\
&= \sum_f |a(bc) \to d; f\rangle R^f_{bc} \left(F^d_{acb}\right)^f_e = \\
&= \sum_{f,g} |(ab)c \to d; g\rangle \left[(F^{-1})^d_{abc}\right]^g_f R^f_{bc} \left(F^d_{acb}\right)^f_e ,
\end{aligned} \tag{8.87}$$

или

$$\mathbf{B} : |(ac)b \to d; e\rangle \longmapsto \sum_g |(ab)c \to d; g\rangle \left(B^d_{abc}\right)^g_e , \tag{8.88}$$

где

$$\left(B^d_{abc}\right)^g_e = \sum_f \left[(F^{-1})^d_{abc}\right]^g_f R^f_{bc} \left(F^d_{acb}\right)^f_e . \tag{8.89}$$

Мы, как и хотели, выразили действие B-матрицы в стандартном базисе через F- и R-матрицы.

Таким образом, реализованное n анионами представление группы кос полностью характеризуется F-матрицей и R-матрицей. Более того, мы увидели, что R-матрица также определяет топологические спины анионов, так что мы фактически построили представление более широкой группы, генераторы которой включают как перестановки соседних частиц, *так и* повороты частиц на 2π. Подходящим названием для этой группы было бы «группа лент», поскольку ее элементы скорее находятся во взаимно однозначном

соответствии с топологическими классами сплетенных лент (которые могут скручиваться), нежели сплетенных струн; однако математики уже назвали это «группой классов отображений для сферы с n проколами».

На данном этапе мы заканчиваем наше общее описание анионной модели. Модель определяется следующими параметрами: (1) набор меток, (2) правила композиции, (3) R-матрица и (4) F-матрица.

Построенный нами математический объект называется *унитарным топологическим модулярным функтором*. Он тесно связан с двумя другими объектами, которые уже были хорошо изучены: *теорией топологического квантового поля* в $2+1$-мерном пространстве-времени, а также *теорией конформного поля* в $1+1$-мерном пространстве-времени. Однако мы будем называть его просто *анионной моделью*.

8.13. Моделирование анионов квантовой схемой

Топологические квантовые вычисления выполняются в три этапа:

1. *Инициализация*: Создаются пары частица-античастица $c_1\bar{c}_1$, $c_2\bar{c}_2$, $c_3\bar{c}_3$, $\ldots, c_m\bar{c}_m$. Каждая пара определенного типа и с тривиальным полным зарядом.

2. *Обработка*: Количество $n = 2m$ частиц проводятся вдоль траекторий, их мировые линии следуют определенной косе.

3. *Считывание*: Пары оказавшихся рядом частиц объединяются и ведется запись, аннигилирует каждая такая пара полностью или нет. Эта запись является результатом вычислений.

(В случае вычислительной модели неабелева сверхпроводника предполагалось, что сплетение определяется результатом объединения, выполняемого на стадии обработки. Но здесь рассматривается модель, в которой все измерения откладываются вплоть до финального считывания).

Какой мощностью обладает данная модель вычислений? Я утверждаю, что этот топологический квантовый компьютер можно эффективно смоделировать с помощью квантовой схемы, поскольку топологическое гильбертово пространство n-анионов не имеет простого и естественного разложения на тензорное произведение малых подсистем. Возможно, это утверждение непосредственно не очевидно. Для доказательства, необходимо объяснить следующее:

1. как с помощью обычных кубитов закодировать топологическое гильбертово пространство,

2. как, используя квантовые вентили, эффективно представить сплетение,

3. как смоделировать объединение анионной пары.

Кодирование. Поскольку каждая образованная в процессе инициализации пара имеет тривиальный полный заряд, исходное n-анионное состояние также имеет тривиальный полный заряд. Следовательно, для определенного набора меток $a_1, a_2, a_3, \ldots, a_n$ топологическое гильбертово пространство имеет вид

$$V^1_{a_1 a_2 a_3 \cdots a_n} \cong \bigoplus_{b_1, b_2, \ldots, b_{n-3}} V^{b_1}_{a_1 a_2} \otimes V^{b_2}_{b_1 a_3} \otimes \cdots \otimes V^{\bar{a}_n}_{b_{n-3} a_{n-1}}; \tag{8.90}$$

имеется $n-3$ промежуточных зарядов и $n-2$ пространств композитных состояний, возникающих в каждом слагаемом. Перестановки частиц меняют местами метки, но после каждой перестановки векторное пространство попрежнему имеет вид (8.90), отличаясь от него лишь порядком следования меток a_j.

Хотя каждое n-анионное топологическое гильбертово пространство само по себе не является тензорным произведением подсистем, все эти пространства содержатся в

$$(\mathcal{H}_d)^{\otimes(n-2)}, \tag{8.91}$$

где

$$\mathcal{H}_d = \bigoplus_{a,b,c} V^1_{abc}. \tag{8.92}$$

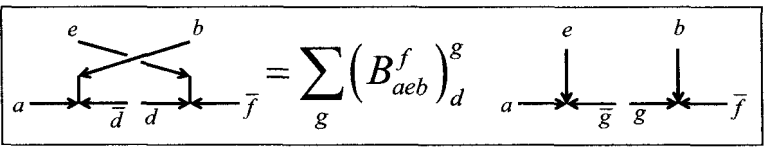
Здесь a, b, c суммируются по полному набору меток модели (которая по предположению является конечной), так что пространство $\mathcal{H}_d$ содержит все возможные композитные состояния трех частиц, а его размерность d равна

$$d = \sum_{a,b,c} N^1_{abc}. \tag{8.93}$$

Таким образом, n-анионное состояние можно закодировать в гильбертовом пространстве $(n-2)$ *кудитов* некоторой постоянной размерности d (которая зависит от анионной модели, но не зависит от n). Базисные состояния этого кудита можно задать в виде $\{|a, b, c; \mu\rangle\}$, где μ маркирует элемент базиса пространства композитных состояний V^1_{abc}.

Сплетение. В топологическом квантовом компьютере коса сплетается путем выполнения ряда перестановок, каждая из которых действует на пару соседних частиц. Результат каждой перестановки в стандартном базисе

ывается B-матрицей. Как она действует в закодированном (с помощью тов) топологическом векторном пространстве? Опуская для краткости и композитных состояний, наш базис для двухкудитовых состояний но обозначить как $|a,b,c\rangle|d,e,\bar{f}\rangle$. Но в топологическом квантовом компьютере метки d и $\bar{c}$ всегда совпадают, и, следовательно, для выполнения елирования сплетения нужно рассмотреть лишь такие двухкудитовые ояния, метки которых совпадают в следующем смысле:

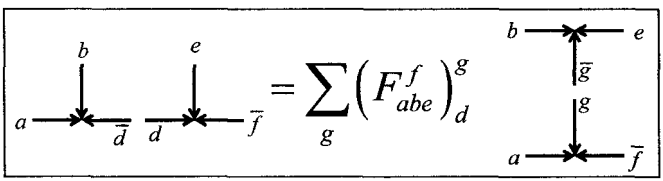


а действие B-матрицы на эти базисные состояния имеет вид

$$\mathbf{B} : |a,b,\bar{d}\rangle|d,e,\bar{f}\rangle \longmapsto \sum_g |a,e,\bar{g}\rangle|g,b,\bar{f}\rangle \left(B^f_{aeb}\right)^g_d . \qquad (8.94)$$

и требовалось, мы представили $\mathbf{B}$ как $d^2 \times d^2$-матрицу, действующую ару соседних кудитов.

Объединение. Объединение анионной пары можно смоделировать кудитовым измерением, которое с помощью F-матрицы можно свести нокудитовому измерению:



мотрим базисное состояние $|a,b,\bar{d}\rangle|d,e,\bar{f}\rangle$ пары соседних кудитов; ка- амплитуда вероятности того, что анионная пара (be) имеет тривиный полный заряд? Используя F-преобразование, ее состояние можно ставить в виде разложения

$$\begin{aligned} \mathbf{F} : |a,b,\bar{d}\rangle|d,e,\bar{f}\rangle &\longmapsto \sum_g |a,g,\bar{f}\rangle|b,\bar{g},e\rangle \left(F^f_{abe}\right)^g_d = \\ &= |a,1,\bar{f}\rangle|b,1,e\rangle \left(F^f_{abe}\right)^1_d + \sum_{g\neq 1} |a,g,\bar{f}\rangle|b,\bar{g},e\rangle \left(F^f_{abe}\right)^g_d ; \end{aligned} \qquad (8.95)$$

мы явно выделили из суммы по g слагаемое, отвечающее композитному состоянию с тривиальным зарядом 1. После F-преобразования (которое является лишь специфическим двухкудитовым унитарным вентилем) можно определить вероятность того, что (be) комбинируется к 1, выполняя проекционное измерение второго кудита в базисе $\{|b, \bar{g}, e\rangle\}$, и записывая, действительно ли $g = 1$.

Этим завершается доказательство того, что квантовая схема может эффективно моделировать топологический квантовый компьютер.

8.14. Анионы Фибоначчи

Мы установили, что топологические квантовые вычисления не мощнее модели квантовой схемы — любая задача, которую можно эффективно решить путем сплетения неабелевых анионов, может быть также эффективно решена с помощью квантовой схемы. Но насколько мощными являются топологические вычисления? Можно ли с помощью сплетения анионов смоделировать универсальный квантовый компьютер? Ответ зависит от специфических свойств анионов: одни модели неабелевых анионов универсальны, другие — нет. Чтобы найти ответ для конкретной анионной модели, необходимо понять свойства представлений группы кос, которые определяются F- и R-матрицами.

Вместо того чтобы заниматься общим обсуждением, изучим одну особенно простую модель неабелевых анионов и продемонстрируем ее вычислительную универсальность. Это простейшая из неабелевых моделей — специалисты в области конформной теории поля называют ее «моделью Янга–Ли», но я буду называть ее «моделью Фибоначчи» по причинам, которые вскоре станут ясны.

В модели Фибоначчи имеется лишь две метки: тривиальная метка, которую я сейчас обозначу как 0, и единственная нетривиальная метка, которую я обозначу как 1, где $\bar{1} = 1$. И существует только одно нетривиальное правило композиции:

$$1 \times 1 = 0 + 1; \tag{8.96}$$

при столкновении два аниона либо аннигилируют, либо превращаются в изолированный анион. Модель неабелева, поскольку два аниона могут объединяться двумя различимыми способами.

Рассмотрим стандартный базис для n-анионного гильбертова пространства V_1^{bn}, каждый базисный элемент которого описывает различимый способ, которым могут комбинироваться n анионов, выдавая в итоге полный заряд $b \in \{0, 1\}$. Пусть сначала объединяются два крайних слева ани-

она; зарядом возникающего в результате состояния может быть 0 или 1. Затем этот заряд объединяется с третьим слева анионом, давая полный заряд 0 или 1, и так далее. Наконец, последний анион объединяется с полным зарядом первых $n-1$ анионов и выдает полный заряд b. В этом описании процесса появляется всего $n-2$ промежуточных зарядов $b_1, b_2, b_3, \ldots, b_{n-2}$; таким образом, соответствующие элементы базиса можно обозначить двоичной строкой длины $n-2$. Если полный заряд равен 0, результат композиции первых $n-1$ анионов *должен* быть 1, так что базисные состояния маркируются строками длиной $n-3$.

Однако *не все* бинарные строки допустимы — за 0 всегда должна следовать 1. Не существует строк, содержащих два нуля подряд, поскольку единственно возможным результатом объединения зарядов 0 и 1 является полный заряд 1. Во всем остальном, в последовательности не имеется ограничений. Следовательно, базисные состояния находятся во взаимно однозначном соответствии с двоичными строками, не содержащими ни одной пары следующих друг за другом нулей.

Таким образом, размерность $N_n^0 \equiv N_{1^n}^0$ топологического гильбертова пространства $V_{1^n}^0$ подчиняется простому рекуррентному соотношению. Если объединение первых двух частиц дает тривиальный полный заряд, то оставшиеся $n-2$ частиц могут объединяться N_{n-2}^0 различимыми способами. Если же объединение первых двух частиц дает анион с нетривиальным зарядом, то этот анион может объединиться с другими $n-2$ анионами N_{n-1}^0 способами; следовательно,

$$N_n^0 = N_{n-1}^0 + N_{n-2}^0. \tag{8.97}$$

Поскольку $N_1^0 = 0$ и $N_2^0 = 1$, решением данного рекуррентного соотношения является

$$\begin{array}{rcccccccccc} n & = & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\ldots \\ N_n^0 & = & 0 & 1 & 1 & 2 & 3 & 5 & 8 & 13 & 21\ldots \end{array} \tag{8.98}$$

— последовательность чисел Фибоначчи (вот почему я называю эту модель «моделью Фибоначчи»).

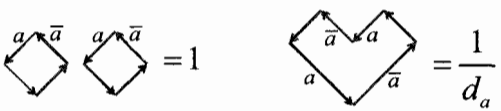
Числа Фибоначчи растут вместе с n со скоростью $N_n^0 \approx C\phi^n$, где ϕ — золотое сечение $\phi = (1+\sqrt{5})/2 \approx 1.618$. Поскольку ϕ управляет скоростью, с которой расширяется гильбертово пространство при добавлении анионов, мы говорим, что $d = \phi$ — *квантовая размерность* аниона Фибоначчи. Иррациональность этой «размерности» служит яркой иллюстрацией того, что топологическое гильбертово пространство не имеет естественного разложения на тензорное произведение подсистем — напротив, топологически

закодированная квантовая информация является коллективным свойством n анионов.

8.15. Квантовая размерность

Мы еще вернемся к свойствам модели Фибоначчи, но сначала более глубоко изучим концепцию квантовой размерности. Как определить размерность d_a метки a для общей анионной модели? С этой целью удобно представить физический процесс, в котором рождаются две пары $a\bar{a}$ (каждая с тривиальным полным зарядом); затем частица a из пары справа объединяется с античастицей $\bar{a}$ из пары слева. Аннигилируют ли эти частицы?

При соответствующем соглашении относительно фаз, *амплитуда* вероятности аннигиляции является действительным числом из единичного интервала [0,1]. Определим его как $1/d_a$, где d_a — квантовая размерность a (а $1/d_a^2$ — *вероятность* аннигиляции). Отметим, что из этого определения становится очевидным равенство $d_a = d_{\bar{a}}$. В том случае, когда метка a является меткой неприводимого представления R_a группы G, ее размерность в точности равна $d_a = |R_a|$, то есть размерности представления. Графически это понять проще:



Если рождаются две пары, а затем каждая пара немедленно аннигилирует, их мировые линии образуют две замкнутые петли, а $|R|$ подсчитывает количество распространяющихся вдоль каждой петли различимых «цветов». Но если частица из каждой пары аннигилирует с античастицей другой пары, то существует только одна замкнутая петля и, следовательно, одно суммирование по цветам; если мы нормируем процесс слева на единицу, амплитуда для процесса справа подавляется коэффициентом $1/|R|$. Выражая эту же идею на языке уравнений, запишем нормированное состояние $R\bar{R}$-пары в виде

$$|R\bar{R}\rangle = \frac{1}{\sqrt{|R|}} \sum_i |i\rangle|\bar{i}\rangle, \tag{8.99}$$

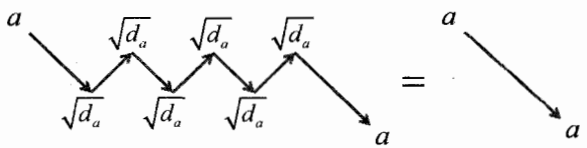
где $\{|i\rangle\}$ обозначает ортонормированный базис для R, а $\{|\bar{i}\rangle\}$ — базис для $\bar{R}$. Предположим, что рождаются пары $|R\bar{R}\rangle$ и $|R'\bar{R}'\rangle$; если после обмена эле-

ментами пары объединяются, то амплитуда аннигиляции равна

$$\langle R\bar{R}, R'\bar{R}'|R\bar{R}', R'\bar{R}\rangle = \frac{1}{|R|^2}\sum_{i,i',j,j'}\langle j\bar{j}, j'\bar{j}'|i\bar{i}', i'\bar{i}\rangle = \\ = \frac{1}{|R|^2}\sum_{i,i',j,j'}\delta_{ji}\delta_{ji'}\delta_{j'i'}\delta_{j'i} = \frac{1}{|R|^2}\sum_{i}\delta_{ii} = \frac{1}{|R|}. \tag{8.100}$$

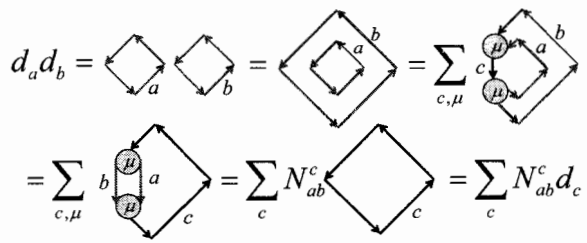
Однако в общем случае квантовая размерность не имеет прямой интерпретации как «счетчика цветов» и нет причин, по которым она должна быть целым числом.

Как такие квантовые размерности связаны с размерностями топологических гильбертовых пространств? Чтобы увидеть эту связь, полезно изменить наши соглашения о нормировке. Обратите внимание на то, что, создавая множество пар $a\bar{a}$ и аннигилируя частицу каждой пары с античастицей следующей, в мировую линию аниона a можно ввести множество «зигзагов». Но каждый зигзаг понижает амплитуду на очередной множитель $1/d_a$. Его можно скомпенсировать, приписывая каждому рождению и аннигиляции пары весовой весовой множитель $\sqrt{d_a}$. С учетом этого соглашения мы можем, не опасаясь неприятных последствий, изгибать мировую линию частицы вперед и назад во времени:



Теперь приписанный мировой линии вес является топологическим инвариантом (при деформации линии он остается неизменным), а мировая линия частицы типа a, образующая замкнутую петлю, приобретает весовой множитель d_a.

С учетом этих новых соглашений можно объяснить следующую последовательность действий:



Каждая диаграмма представляет скалярное произведение двух (нормированных нестандартно) состояний. Мы вставили сумму по всем меткам (c) и соответствующим композитным состояниям (μ), которые могут возникнуть при объединении a и b. Затем, используя топологическую инвариантность диаграммы, мы вывернули ее наизнанку и сократили композитные состояния (приобретая множитель N_{ab}^{c}, подсчитывающий возможные значения μ).

Выведенное нами уравнение

$$d_a d_b = \sum_c N_{ab}^{c} d_c \equiv \sum_c (N_a)_b^c d_c \tag{8.101}$$

означает, что вектор $\vec{d}$, компонентами которого являются квантовые размерности, является собственным вектором с собственным значением d_a матрицы N_a, описывающей правила композиции метки a с другими метками:

$$N_a \vec{d} = d_a \vec{d}. \tag{8.102}$$

Более того, поскольку матрица N_a имеет неотрицательные элементы, а все компоненты вектора $\vec{d}$ положительны, размерность d_a равна наибольшему собственному значению матрицы N_a и является невырожденной. (Это простое замечание иногда называют *теоремой Перрона–Фробениуса*.) Для n анионов с одинаковыми метками a топологическое гильбертово пространство $V_{aaa\cdots a}^{b}$ сектора с полным зарядом b имеет размерность

$$N_{aaa\cdots a}^{b} = \sum_{\{b_i\}} N_{aa}^{b_1} N_{ab_1}^{b_2} N_{ab_2}^{b_3} \cdots N_{ab_{n-2}}^{b} = \langle b|(N_a)^{n-1}|a\rangle. \tag{8.103}$$

Матрица N_a может быть диагонализована и представлена в виде

$$N_a = |v\rangle d_a \langle v| + \cdots, \tag{8.104}$$

где

$$|v\rangle = \frac{\vec{d}}{\mathcal{D}}, \quad \mathcal{D} = \sqrt{\sum_c {d_c}^2}, \tag{8.105}$$

и опущены собственные значения более высоких порядков малости; следовательно:

$$N_{aaa\cdots a}^{b} = d_a^n d_b / \mathcal{D}^2 + \cdots, \tag{8.106}$$

где многоточие обозначает слагаемые, экспоненциально малые при больших n. Таким образом, квантовая размерность d_a определяет скорость роста n-частичного гильбертова пространства для анионов типа a.

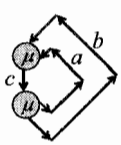
Поскольку метка 0 с тривиальным зарядом комбинируется тривиально, мы имеем $d_0 = 1$. В случае модели Фибоначчи из правила композиции $1 \times 1 = 0 + 1$ следует, что $d_1^2 = 1 + d_1$, решением которого является $d_1 = \phi$, как мы обнаружили ранее; следовательно, $\mathcal{D}^2 = d_0^2 + d_1^2 = 1 + \phi^2 = 2 + \phi$. Наша формула принимает вид

$$N_{111\cdots1}^{0} = \left(\frac{1}{2+\phi}\right)\phi^n, \tag{8.107}$$

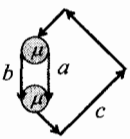
что является прекрасным приближением к числам Фибоначчи даже при умеренных значениях n.

Предположим, что одновременно рождаются пары $a\bar{a}$ и $b\bar{b}$. Какова вероятность $p(ab \to c)$ того, что при объединении частиц a и b возникнет состояние с полным зарядом c? На этот вопрос можно ответить, используя те же графические манипуляции:

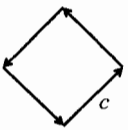
$$d_a d_b p(ab \to c) = \sum_{\mu}$$



$$= \sum_{\mu}$$



$$= N_{ab}^{c}$$



$$= N_{ab}^{c} d_c$$

Деля на $d_a d_b$ с целью восстановить подходящую нормировку скалярного произведения, мы приходим к выводу, что

$$p(ab \to c) = \frac{N_{ab}^{c} d_c}{d_a d_b}, \tag{8.108}$$

что обобщает формулу $p(a\bar{a} \to 1) = 1/d_a^2$, которую мы использовали для определения квантовой размерности, и удовлетворяет условию нормировки

$$\sum_{c} p(ab \to c) = 1. \tag{8.109}$$

Чтобы прийти к другой интерпретации квантовой размерности, представим рождение плотного анионного газа, который затем некоторое время

отжигается — анионы сталкиваются и объединяются, постепенно понижая количество частиц. В конечном счете (но задолго до достижения теплового равновесия) частота столкновений становится настолько низкой, что процессы объединения практически останавливаются. К этому моменту, каким бы ни было начальное распределение типов частиц, достигается устойчивое распределение состояний, которое сохраняется столкновениями. Если в устойчивом состоянии частицы типа a появляются с вероятностью p_a, то

$$\sum_{ab} p_a p_b p(ab \to c) = p_c. \tag{8.110}$$

Используя уравнение

$$\sum_a N_{ab}^c d_a = \sum_a N_{b\bar{c}}^a \, d_{\bar{a}} = d_b d_{\bar{c}} = d_b d_c, \tag{8.111}$$

нетрудно убедиться в том, что это условие удовлетворяется уравнением

$$p_a = \frac{d_a^2}{\mathcal{D}^2}. \tag{8.112}$$

Таким образом, в случайном процессе рождения анионов бóльшую вероятность генерации имеют анионы, несущие метки бóльшей квантовой размерности, в соответствии с тем свойством, что анионы с бóльшей размерностью имеют больше квантовых состояний.

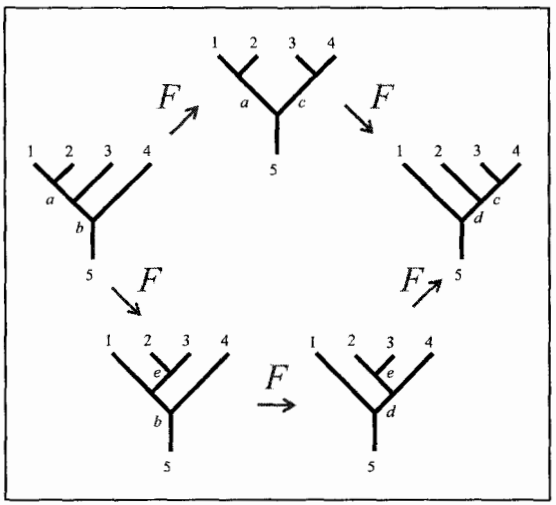
8.16. Уравнения пяти- и шестиугольника

Чтобы оценить вычислительную мощность анионной модели (такой, как модель Фибоначчи), необходимо знать свойства сплетений анионов, которые определяются R- и F-матрицами. Мы увидим, что правила сплетения существенно ограничиваются алгебраическими условиями совместимости. В случае модели Фибоначчи этих условий согласования достаточно для определения единственного правила сплетения, совместимого с правилами композиции.

Условия совместимости возникают потому, что изоморфизм, связывающий два топологических пространства, можно получить, выполняя последовательность «F-преобразований» и «R-преобразований». Изоморфизм можно рассматривать как унитарную матрицу, которая связывает канонические ортонормированные базисы двух разных пространств; это унитар-

ное преобразование не зависит от конкретной последовательности преобразований, из которых конструируется изоморфизм, а лишь от начального и конечного базисов.

Например, существует пять связанных F-преобразованиями разных способов объединения четырех частиц (без их перестановок):



Изображенный крайним слева на этой пятиугольной схеме базис — «левый стандартный базис» $\{|\text{left}; a, b\rangle\}$, в котором сначала объединяются частицы 1 и 2, результирующий заряд a объединяется с частицей 3 и в результате выдает заряд b, а затем, наконец, b объединяется с частицей 4 и дает полный заряд 5. Изображенный крайним справа базис — «правый стандартный базис» $\{|\text{right}; c, d\rangle\}$, в котором частицы объединяются не слева направо, а наоборот — справа налево. Через вершину пятиугольника эти два базиса связаны двумя F-преобразованиями, то есть

$$|\text{left}; a, b\rangle = \sum_{c,d} |\text{right}; c, d\rangle (F^5_{12c})^d_a (F^5_{a34})^c_b. \tag{8.113}$$

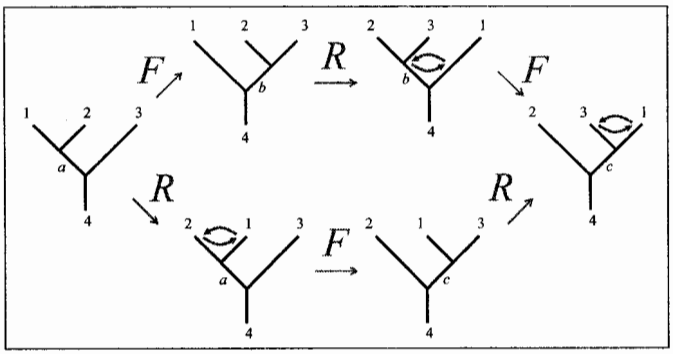
Через нижнюю часть пятиугольника они связаны тремя F-преобразованиями, другими словами,

$$|\text{left}; a, b\rangle = \sum_{c,d,e} |\text{right}; c, d\rangle (F^d_{234})^c_e (F^5_{1e4})^d_b (F^b_{123})^e_a. \tag{8.114}$$

Сравнение этих двух выражений для $|\text{left}; a, b\rangle$ дает *уравнение пятиугольника*:

$$(F^5_{12c})^d_a (F^5_{a34})^c_b = \sum_e (F^d_{234})^c_e (F^5_{1e4})^d_b (F^b_{123})^e_a. \tag{8.115}$$

Другое нетривиальное условие совместимости можно получить, если рассмотреть различные способы объединения трех частиц:



Крайний слева на этой шестиугольной схеме базис $\{|\text{left}; a\rangle\}$ получается, если частицы располагаются в порядке 123 и сначала объединяются частицы 1 и 2. Крайний справа базис $\{|\text{right}; c\rangle\}$ получается при расстановке частиц в порядке 231, и если сначала комбинируются частицы 1 и 3. Через верхнюю часть шестиугольника два этих базиса связаны последовательностью преобразований FRF:

$$|\text{left}, a\rangle = \sum_{b,c} |\text{right}; c\rangle (F^4_{231})^c_b R^4_{1b} (F^4_{123})^b_a. \tag{8.116}$$

Через нижнюю часть шестиугольника они связаны последовательностью преобразований RFR, то есть

$$|\text{left}, a\rangle = \sum_c |\text{right}; c\rangle R^c_{13} (F^4_{213})^c_a R^a_{12}. \tag{8.117}$$

Сравнение этих двух выражений для $|\text{left}, a\rangle$ дает *уравнение шестиугольника*:

$$R^c_{13} (F^4_{213})^c_a R^a_{12} = \sum_b (F^4_{231})^c_b R^4_{1b} (F^4_{123})^b_a. \tag{8.118}$$

Красивая теорема, которую здесь я не стану доказывать, утверждает, что не существует иных условий, обеспечивающих совместимость сплетения и объединения. То есть при любом выборе начального и конечного базисов для n анионов *все* последовательности R- и F-преобразований, переводящих исходный базис в конечный, дают *один и тот же* изоморфизм при условии справедливости уравнений пяти- и шестиугольника. Данная теорема является частным случаем *теоремы когерентности Маклейна*, фундаментального результата теории категорий. Вместе уравнения пяти- и шестиугольника называются *полиномиальными уравнениями Мура–Сейберга* — их важность для физики была впервые осознана в 80-е годы при изучении (1+1)-мерной конформной теории поля.

Решение полиномиальных уравнений определяет жизнеспособные анионные модели. Следовательно, существует систематическая процедура их конструирования:

1. выбрать набор меток и принять правило композиции,
2. решить полиномиальные уравнения для $\mathbf{R}$ и $\mathbf{F}$.

Если решений не существует, то гипотетическое правило композиции несовместимо с принципами локальной квантовой физики и должно быть забраковано. Если существует более одного решения (причем решения не связаны друг с другом перестановкой меток, переопределением базисов и т. п.), то каждое решение определяет отдельную модель с принятым правилом композиции.

Чтобы проиллюстрировать эту процедуру, рассмотрим полиномиальные уравнения для правила композиции Фибоначчи. Здесь фигурируют только две F-матрицы, которые мы будем обозначать как

$$\mathbf{F}_{0111} \equiv \mathbf{F}_0, \quad \mathbf{F}_{1111} \equiv \mathbf{F}_1. \tag{8.119}$$

$\mathbf{F}_0$ в действительности является матрицей размерности 1×1

$$(F_0)_a^b = \delta_a^1 \delta_1^b, \tag{8.120}$$

тогда как $\mathbf{F}_1$ — матрица размерности 2×2. Уравнение пятиугольника приобретает вид

$$(F_c)_a^d (F_a)_b^c = \sum_e (F_d)_e^c (F_e)_b^d (F_b)_a^c. \tag{8.121}$$

Общее решение для $\mathbf{F} \equiv \mathbf{F}_1$:

$$\mathbf{F} = \begin{pmatrix} \tau & e^{i\phi}\sqrt{\tau} \\ e^{-i\phi}\sqrt{\tau} & -\tau \end{pmatrix}, \tag{8.122}$$

где $e^{i\phi}$ — произвольный фазовый множитель (который при подходящем соглашении о фазах можно положить равным единице), а $\tau = (\sqrt{5}-1)/2 \approx \approx 0.618$, что удовлетворяет уравнению

$$\tau^2 + \tau = 1. \tag{8.123}$$

2×2-матрица $\mathbf{R}$, описывающая перестановку против часовой стрелки двух анионов Фибоначчи, имеет два собственных значения: R^0 для случая, когда полный заряд анионной пары тривиален, и R^1 в случае нетривиального заряда. Уравнение шестиугольника принимает вид

$$R^c(F)^c_a R^a = (F)^c_0 R^0 (F)^0_a + (F)^c_1 R^1 (F)^1_a. \tag{8.124}$$

Используя выражение для $\mathbf{F}$, полученное при решении уравнения пятиугольника, можно решить уравнение шестиугольника для $\mathbf{R}$ и найти

$$\mathbf{R} = \begin{pmatrix} e^{4\pi i/5} & 0 \\ 0 & -e^{2\pi i/5} \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \tau & \sqrt{\tau} \\ \sqrt{\tau} & -\tau \end{pmatrix}. \tag{8.125}$$

Единственным альтернативным решением является комплексно-сопряженное полученному; это второе решение на самом деле описывает ту же модель с точностью до выбора направления обхода в процессе сплетения: по или против часовой стрелки. Следовательно, анионная модель с правилом композиции Фибоначчи *действительно* существует и по сути она единственна.

8.17. Моделирование квантовой схемы с анионами Фибоначчи

Теперь мы знаем достаточно, чтобы задуматься о том, возможно ли моделирование универсального квантового компьютера с помощью анионов Фибоначчи. Необходимо объяснить, как с помощью анионов можно закодировать кубиты и как можно реализовать универсальный набор квантовых вентилей.

Сначала отметим, что гильбертово пространство $V_4^0 \equiv V_{1111}^0$ имеет размерность $N_4^0 = 2$; следовательно, кубит можно закодировать четырьмя анионами с тривиальным полным зарядом. Анионы выстроены в ряд в порядке 1234, пронумерованные слева направо; в стандартном базисном состоянии $|0\rangle$ анионы 1 и 2 объединяются и дают в результате полный заряд 0, тогда как в стандартном базисе $|1\rangle$ анионы 1 и 2 при объединении дают полный заряд 1. Действующий в этом стандартном базисе генератор

группы кос σ_1 (перестановка частиц 1 и 2 против часовой стрелки) представляется матрицей

$$\sigma_1 \longmapsto \mathbf{R} = \begin{pmatrix} e^{4\pi i/5} & 0 \\ 0 & -e^{2\pi i/5} \end{pmatrix}, \tag{8.126}$$

тогда как генератор σ_2 представляется матрицей

$$\sigma_2 \longmapsto \mathbf{B} = \mathbf{F}^{-1}\mathbf{R}\mathbf{F}, \quad \mathbf{F} = \begin{pmatrix} \tau & \sqrt{\tau} \\ \sqrt{\tau} & -\tau \end{pmatrix}. \tag{8.127}$$

Эти матрицы генерируют представление группы кос B_3 на трех нитях, имеющее плотный в SU(2) образ. Действительно, матрицы $\mathbf{R}$ и $\mathbf{B}$ генерируют Z_{10} — подгруппы группы U(2) вокруг двух различных осей, и не существует конечной подгруппы группы U(2), содержащей обе эти подгруппы; следовательно, представление является плотным на группе, содержащей все элементы группы U(2) с детерминантом, равным корню десятой степени из единицы. Аналогично, для n анионов с тривиальным полным зарядом образ представлений группы кос является плотным в $SU(N_n^0)$.

Чтобы смоделировать квантовую схему, действующую на n кубитов, всего используется $4n$ анионов. Мы уже видели, что путем сплетения внутри каждого кластера из четырех анионов можно реализовать произвольные однокубитовые вентили. Чтобы укомплектовать универсальный набор, нам еще потребуются двухкубитовые вентили. Но два соседних кубита кодируются восемью анионами, а перестановки этих анионов генерируют представление группы B_8, образ которого является плотным в $\mathrm{SU}(N_8^0) = \mathrm{SU}(13)$, что, конечно, включает группу SU(4), действующую на два закодированных кубита. Следовательно, каждый вентиль из универсального набора может быть сколь угодно точно смоделирован некоторой конечной косой.

Поскольку мы можем плести косы как по часовой стрелке, так и против, мы также имеем в своем распоряжении обратное значение каждого вентиля перестановок. Следовательно, можно применить теорему Соловея–Китаева и сделать вывод о том, что при длине кос $\mathrm{poly}(\log(1/\varepsilon))$ универсальные вентили модели квантовых схем можно моделировать с точностью ε. Отсюда следует, что идеальная квантовая схема с L вентилями, действующими на все n кубитов сразу, может быть смоделирована с установленной точностью с помощью $4n$ анионов и косы длины $O(L \cdot \mathrm{poly}(\log(L)))$. Как и требовалось, мы показали, что универсальный квантовый компьютер можно эффективно моделировать с помощью анионов Фибоначчи. Отметим, что, в отличие от моделирования с использованием модели неабелева сверхпроводника, для реализации универсальных вентилей не нужны никакие промежуточные измерения.

В вышеприведенном анализе мы предположили полное отсутствие ошибок моделирования, за исключением тех, что ограничивают точность приближения Соловея–Китаева к идеальным вентилям. Следовательно, подразумевается, что температура достаточно низкая по сравнению с энергетической щелью модели, что термически возбужденные анионы слишком редки, чтобы вызвать возмущение, что анионы удерживаются на достаточном расстоянии друг от друга и можно пренебречь неконтролируемой перестановкой зарядов и, наконец, что в топологических квантовых вычислениях ошибки вообще несущественны. Если частота появления ошибок достаточно мала, но не настолько, чтобы ею можно было пренебречь, то для поддержания необходимой точности моделирования можно обратиться к стандартным методам теории квантовой отказоустойчивости, требующим дополнительных полилогарифмических по L накладных расходов. Отказоустойчивая процедура должна включать в себя метод контролирования «утечки» закодированных кубитов, то есть предотвращать уход четырехкубитовых кластеров из двумерного вычислительного пространства V_4^0 в его трехмерное ортогональное дополнительное пространство V_4^1.

8.18. Заключение

Здесь я вкратце коснусь нескольких других тем, которые я мог бы осветить, если бы не кончилось время.

8.18.1. Теория Черна–Саймонса

Мы уже обсуждали, как можно конструировать анионные модели путем прямолинейного решения полиномиальных уравнений. Этот метод прост, но на практике модели часто конструируются с использованием других, более эффективных методов. Действительно, большинство известных анионных моделей были открыты как частные случаи *теории Черна–Саймонса*.

Правила композиции теории Черна–Саймонса представляют собой усеченную версию правил композиции для неприводимых представлений группы Ли. Например, существует ассоциированное с группой SU(2) множество теорий Черна–Саймонса, нумеруемых положительными целыми числами k. Для SU(2) неприводимые представления имеют метки $j = 0, 1/2, 1, 3/2, 2, 5/2, \ldots$, а правила композиции имеют вид

$$j_1 \times j_2 = \sum_{j=|j_2-j_1|}^{j_1+j_2} j. \qquad (8.128)$$

В теории Черна–Саймонса, обозначенной $\mathrm{SU}(2)_k$, полуцелые метки ограничены величиной $j \leqslant k/2$, а метка j содержится в $j_1 \times j_2$, если только $j_1 + j_2 + j \leqslant k$.

Например, модель группы $\mathrm{SU}(2)_1$ — абелева, а нетривиальные правила композиции модели группы $\mathrm{SU}(2)_2$ следующие:

$$\begin{aligned} \tfrac{1}{2} \times \tfrac{1}{2} &= 0 + 1, \\ \tfrac{1}{2} \times 1 &= \tfrac{1}{2}, \\ 1 \times 1 &= 0. \end{aligned} \tag{8.129}$$

Следовательно, метка $1/2$ имеет квантовую размерность $d_{1/2} = \sqrt{2}$, а топологическое гильбертово пространство $2m$ таких анионов с полным зарядом 0 имеет размерность

$$N^0_{(\frac{1}{2})^{2m}} = 2^{m-1}. \tag{8.130}$$

Полиномиальные уравнения для этих правил композиции имеют множество решений (из которых лишь одно описывает свойства сплетения модели группы $\mathrm{SU}(2)_2$), но ни одна из результирующих моделей не имеет правил сплетения, универсальных с точки зрения реализации квантовых вычислений. Пространство $V^0_{\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}}$ двумерно, а 2×2-матрицы $\mathbf{F} \equiv \mathbf{F}_{\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}}$ и $\mathbf{R} \equiv \mathbf{R}_{\frac{1}{2}\frac{1}{2}}$, с точностью до общих фаз и комплексного сопряжения, имеют вид

$$\mathbf{F} = \mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{R} = \mathbf{P} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{8.131}$$

Это квантовые вентили группы Клиффорда, не отвечающие требованиям универсальности.

Однако при $k \geqslant 3$ $\mathrm{SU}(2)_k$-модели *вычислительно универсальны*. Нетривиальные правила композиции группы $\mathrm{SU}(2)_3$ имеют вид

$$\begin{aligned} \tfrac{1}{2} \times \tfrac{1}{2} &= 0 + 1, \\ \tfrac{1}{2} \times 1 &= \tfrac{1}{2} + \tfrac{3}{2}, \\ \tfrac{1}{2} \times \tfrac{3}{2} &= 1, \end{aligned}$$

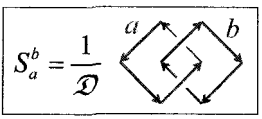
$$
\begin{aligned}
1 \times 1 &= 0 + 1, \\
1 \times \frac{3}{2} &= \frac{1}{2}, \\
\frac{3}{2} \times \frac{3}{2} &= 0.
\end{aligned}
\tag{8.132}
$$

Рассмотренная нами модель Фибоначчи (Янга–Ли) получена путем усечения группы $\mathrm{SU}(2)_3$, состоящем в исключении нецелых меток $1/2$ и $3/2$ (то есть, это теория Черна–Саймонса $\mathrm{SO}(3)_3$); тогда единственным остающимся нетривиальным правилом композиции является $1 \times 1 = 0 + 1$.

Уонг (неопубликовано) недавно сконструировал все анионные модели с количеством меток не более четырех и обнаружил, что все они тесно связаны с открытыми теорией Черна–Саймонса моделями.

8.18.2. S-матрица

Модулярную S-матрицу анионной модели можно определить на языке мировых линий двух анионов, образующих *зацепление Хопфа*:



Здесь $\mathcal{D}$ — полная квантовая размерность модели, и мы воспользовались нормировкой, в которой расцепленные петли имели бы значение $d_a d_b$; тогда матрица S_a^b симметрична и унитарна. В абелевых анионных моделях понятие зацепления Хопфа возникло при обсуждении топологического вырождения, которым определялось воздействие на вакуумное состояние анионной модели, возникающее при обходе аниона вокруг одного из циклов тора. В неабелевом случае S-матрица имеет аналогичную интерпретацию. С помощью элементарных рассуждений S может быть связана с правилами композиции:

$$
(N_a)_b^c = \sum_d S_b^d \left(\frac{S_a^d}{S_1^d} \right) (S^{-1})_d^c; \tag{8.133}
$$

то есть S-матрица одновременно диагонализует все матрицы $\{N_a\}$ (соотношение Верлинде). Отметим, что из определения следует, что $S_1^a = d_a/\mathcal{D}$.

8.18.3. Краевые возбуждения

В нашей формулировке анионных моделей мы обсудили объединение и сплетение частиц в двумерном *объеме*. Но существует другой аспект фи-

зики двумерных сред, который до сих пор не обсуждался, — свойства одномерной *границы* образца. Обычно, если двумерная система содержит анионы в объеме, то имеются и *киральные безмассовые возбуждения*, распространяющиеся вдоль одномерной границы. При ненулевой температуре T вдоль границы существует поток энергии, определяемый выражением

$$J = \frac{\pi}{12} c_- T^2; \tag{8.134}$$

здесь постоянная c_-, называемая *киральным центральным зарядом* границы, является универсальным свойством, неуязвимым для небольших изменений в гамильтониане системы.

В то время как этот киральный центральный заряд является внутренним свойством двумерной среды, свойства анионов в объеме не определяют его полностью; скорее, мы имеем

$$\frac{1}{\mathcal{D}} \sum_a d_a^2 e^{2\pi i J_a} = e^{(2\pi i/8)c_-}, \tag{8.135}$$

где суммирование ведется по полному набору меток анионной модели, а $e^{2\pi i J_a} = R^1_{a\bar{a}}$ — топологический спин метки a. Это выражение связывает величину c_-, характеристику краевой теории, с квантовыми размерностями и топологическими спинами объемной теории, но определяет c_- только по модулю восемь. Следовательно, по крайней мере, в принципе, возможно существование множества краевых теорий, соответствующих единственной теории анионов в объеме.

8.19. Библиографические замечания

Некоторые из пионерских работ по теории анионов напечатаны в [1].

Описанная мной модель «неабелева сверхпроводника» в литературе часто упоминается как «квантовый дубль», она изучается с использованием теории представлений алгебры Хопфа (см., например, [2]).

В [3] было впервые предложено использование неабелевых анионов для отказоустойчивых квантовых вычислений. В этой работе также обсуждается торический код и связанные с ним решеточные модели, имеющие неабелевы фазы. Конкретная реализация универсальных квантовых вычислений в неабелевом сверхпроводнике была рассмотрена в [4, 5]. Мое обсуждение набора универсальных вентилей опирается на [6], где также рассматриваются более общие модели. Другие схемы, которые обеспечивают более экстенсивное использование электрических зарядов и являются универсальными для меньших групп (таких как S_3), описаны в [7].

Графические методы, подобные тем, что я использовал в обсуждении квантовой размерности, широко используются для вывода свойств анионов в [8]. Роль полиномиальных уравнений (уравнений пяти- и шестиугольника) в (1+1)-мерной конформной теории поля обсуждается в [9].

Моделирование анионов с использованием квантовой схемы описано в [10]. Моделирование универсального квантового компьютера с использованием анионов $SU(2)_{k=3}$-теории Черна–Саймонса обсуждается в [11]. В [12] было отмечено, что модель Янга–Ли также является универсальной.

В моих лекциях я не обсуждал физическую реализацию, но в любом случае здесь я перечислю несколько относящихся к данной проблеме работ. Идеи реализации абелевых и неабелевых анионов с помощью массивов сверхпроводящих джозефсоновских контактов описаны в [13]. Спиновая модель со взаимодействием между ближайшими соседями, которая имеет неабелевы анионы (хотя не только они являются вычислительно универсальными), предложена и решена в [14], а предложение реализовать эту модель с использованием холодных атомов, захваченных в оптической решетке, описано в [15]. Некоторые идеи реализации (вычислительно универсальной) модели $SU(2)_{k=3}$ в системе взаимодействующих электронов обсуждаются в [16].

Своим пониманием теории квантовых вычислений с помощью неабелевых анионов я обязан многочисленным полезным дискуссиям с Алексеем Китаевым.

Литература

[1] F. Wilczek, *Fractional statistics and anyon superconductivity* (World Scientific, Singapore, 1990).

[2] M. de Wild Propitius and F. A. Bais, Discrete gauge theories, arXiv: hep-th/9511201 (1995); CRM-CAP Summer School «Particles and Fields 94», Bannf, Alberta, Canada, August 16–24, (1994).

[3] A. Yu. Kitaev, Fault-tolerant quantum computation by anyons, Annals Phys. **303**, 2–30 (2003), arXiv: quant-ph/9707021.

[4] R. W. Ogburn and H. Preskill, Topological quantum computation, in *Lect. Notes in Comp. Sci.* C.P. Williams (ed.) **1509**, 341–356, Springer–Verlag (1999).

[5] J. Preskill, Fault-tolerant quantum computation, arXiv: quant-ph/9712048 (1997); В книге: *Introduction to Quantum Computation*, H.-K. Lo, S. Popesku, T.P. Spiller (Eds.), World Scientific, Singapore et al. (1998); пере-

вод: Дж. Прескилл, Отказоустойчивые квантовые вычисления (в этом издании).

[6] C. Mochon, Anyons from non-solvable finite groups are sufficient for universal quantum computation Phys. Rev. A **67**, 022315 (2003), quant-ph/0206128.

[7] C. Mochon, Anyon computers with smaller groups, Phys. Rev. A **69**, 032306 (2004), arXiv: quant-ph/0306063.

[8] J. Fröhlich and F. Gabbiani, Braid statistics in local quantum theory, Rev. Math. Phys. **2:3**, 251–353 (1990).

[9] G. Moore and N. Seiberg, Classical and quantum conformal field theory, Comm. Math. Phys. **123**, 171–254 (1989).

[10] M. H. Freedman, A. Kitaev, and Z. Wang, Simulation of topological field theories by quantum computers, Comm. Math. Phys. **227**, 587–603 (2002), arXiv: quant-ph/0001071.

[11] M. H. Freedman, M. Larsen, and Z. Wang, A modular functor which is universal for quantum computation, Comm. Math. Phys. **227**, 605–622 (2002), arXiv: quant-ph/0001108 (2000).

[12] G. Kuperberg, unpublished.

[13] B. Doucot, L. B. Iofie, and J. Vidal, Discrete non-Abelian gauge theories in two-dimensional lattices and their realizations in Josephson-junction arrays, Phys. Rev. B **69**, 214501 (2004), arXiv: cond-mat/0302104.

[14] A. Kitaev, Anyons in a spin model on the honeycomb lattice, unpublished. [Обсуждение точно решаемой спиновой модели Китаева на решетке «пчелиных сот» см. в: A. Kitaev, C. Laumann, Topological phases and quantum computation, cond-mat/0904.2771; A. Kitaev, Anyons in an exactly solved model and beyond, *Ann. Phys.*, **321**, 2–111 (2006); cond-mat/0506438]

[15] L.-M. Duan, E. Demler, and M. D. Lukin, Controlling spin exchange interactions of ultracold atoms in optical lattices, Phys. Rev. Lett. **91**, 090402 (2003), arXiv: cond-mat/0210564.

[16] M. Freedman, C. Nayak, K. Shtengel, K. Walker, and Zhenghan Wang, A class of P; T-invariant topological phases of interacting electrons, *Ann. Phys.*, NY, **310**, 428–492; cond-mat/0307511 (2003).

Приложение

Отказоустойчивые квантовые вычисления

Надежные квантовые компьютеры[1]

Джон Прескилл[2]

Новая область науки *коррекция квантовых ошибок* получила заметное развитие с момента своего возникновения менее двух лет назад. Закодированную квантовую информацию можно защитить от ошибок, возникающих вследствие неконтролируемых взаимодействий с окружающей средой. Исправление может эффективно выполняться даже при возникновении случайных ошибок во время самой процедуры восстановления. Более того, закодированная квантовая информация может *обрабатываться* без существенного размножения ошибок. Следовательно, квантовые вычисления произвольной продолжительности могут надежно выполняться при условии, что средняя вероятность одной ошибки на квантовый вентиль меньше некоторого критического значения, называемого *порогом безошибочности*. Квантовый компьютер, вмещающий порядка 10^6 кубитов информации, при вероятности ошибки на квантовый вентиль порядка 10^{-6}, мог бы стать внушительной вычислительной машиной. Даже меньший и менее точный квантовый компьютер смог бы выполнить множество полезных задач.

Данная работа основана на материалах доклада, представленного на Конференции по квантовой когерентности и декогерентизации, Институт теоретической физики, Санта Барбара, 15–18 декабря 1996 г.

1. Золотой век коррекции квантовых ошибок

Многие из нас рассчитывают на то, что в XXI веке квантовые компьютеры, наконец, станут практичными и полезными вычислительными

[1] J. Preskill, Reliable Quantum Computers, *Proc. Roy. Soc. London* A, **454**, pp. 385–410 (1998).

[2] Калифорнийский технологический институт, Пасадена, CA 91125, США.

устройствами. Однако справедливости ради следует отметить, что сейчас никто из нас не может точно предсказать, что будет представлять собой «железо» этих машин будущего; возможно, оно будет значительно отличаться от аппаратных средств, исследуемых в наши дни физиками-экспериментаторами. Но в одном можно быть вполне уверенным — работа реального квантового компьютера будет включать в себя определенный тип коррекции ошибок. Квантовые компьютеры гораздо более восприимчивы к возникновению ошибок по сравнению с традиционными цифровыми компьютерами, поэтому для предотвращения сбоев в их работе потребуются определенные методы контроля и исправления этих ошибок.

Еще в середине 90-х годов мы не имели четкого представления о том, как будет работать коррекция квантовых ошибок и будет ли она работать вообще. Действительно, был ряд причин для пессимизма относительно реальной возможности исправления квантовых ошибок (Unruh 1995; Landauer 1995, 1996, 1997). Во-первых, несмотря на то, что были разработаны достаточно изощренные методы исправления ошибок в классической информации (MacWilliams & Sloane 1977), было далеко не очевидно, как их адаптировать для исправления атакующих квантовые системы *фазовых* ошибок. Во-вторых, в квантовом компьютере, как и в классическом аналоговом компьютере, с течением времени малые ошибки могут накапливаться и, в конечном счете, превращаться в большие, а найти методы, способные предотвращать или исправлять подобные малые ошибки, достаточно сложно. В-третьих, чтобы исправить ошибку, необходимо сначала получить определенную информацию о ее природе, осуществив измерение, которое влечет за собой опасность повреждения закодированной в устройстве чувствительной квантовой информации. Наконец, чтобы защититься от ошибок, необходимо кодировать информацию в избыточном виде. Однако известная теорема (Wootters & Zurek 1982; Dieks 1982) говорит о том, что квантовую информацию нельзя клонировать и, следовательно, не ясно, как хранить (квантовую) информацию с требуемой избыточностью.

Но к настоящему моменту все эти трудности преодолены — теперь мы знаем, что коррекция квантовых ошибок действительно возможна. Ключевая идея, осознанная нами, состоит в том, что с *запутыванием можно бороться с помощью запутывания*. Запутанность квантовых состояний может быть нашим врагом, поскольку запутывание состояния нашего устройства с окружающей средой может скрывать от нас квантовую информацию и, таким образом, служить причиной ошибок. Но запутанность может быть и нашим союзником — информацию, которую мы хотим защитить, можно закодировать в запутанном состоянии, то есть в корреляциях, охватывающих большое количество кубитов. Тогда эту информацию невозможно из-

влечь, измеряя лишь несколько кубитов. Впрочем, по той же причине она не может быть и повреждена, если окружающая среда взаимодействует лишь с несколькими кубитами. Более того, мы узнали, что хотя квантовый компьютер является до известной степени аналоговым устройством, совершаемые им ошибки можно *оцифровать*. Мы боремся с малыми ошибками, выполняя подходящие измерения, которые проецируют состояние квантового компьютера или на исходное неповрежденное состояние, или на состояние с большой ошибкой, которую можно исправить известными методами. Мы осознали, что возможно измерение ошибок без измерения самой информации — можно получить информацию об истинной природе ошибки, ничего при этом не узнав о закодированной в нашем устройстве квантовой информации (что могло бы повлечь за собой декогерентизацию и сбой в нашем вычислении). Основная идея коррекции квантовых ошибок состоит в следующем: маленькое подпространство гильбертова пространства нашего устройства определяется как *кодовое подпространство*. Оно выбирается таким образом, чтобы любая ошибка, которую мы хотим исправить, перемещала его в одно из взаимно ортогональных *подпространств ошибок*. После того как наша система провзаимодействовала с окружением, выполняется измерение, результат которого сообщает о том, в каком из этих взаимно ортогональных подпространств находится система, и, следовательно, точно определяет тип возникшей ошибки. После этого ошибку можно исправить, применяя подходящее унитарное преобразование.

Если мы и были скептически настроены в 1995 году, то к концу 1996 года появилась веская причина для оптимизма. Этот год стал поворотным пунктом в истории квантовой информации; это момент, когда мы узнали, как противостоять и обращать эффекты декогерентизации. Это открытие имеет важное значение для квантовых вычислений, но на самом деле его последствия гораздо шире. Вот несколько основных этапов, пройденных в этом году:

- Осенью 1995 года Питер Шор (Shor 1995) и Эндрю Стин (Steane 1996a) впервые указали на то, что коды, корректирующие квантовые ошибки, существуют.

- К началу 1996 года Стин (Steane 1996b), а также Колдербэнк и Шор (Calderbank & Shor 1996) показали существование *хороших* кодов, то есть кодов, способных корректировать большое количество ошибок.

- Из работы по случайным кодам Ллойда (Lloyd 1997) и Беннетта, Дивинченцо, Смолина и Вутерса (Bennett, *et al.* 1996) мы узнали, что

квантовая информация может храниться в течение достаточно длительного времени. То есть существует код, который позволяет с высокой точностью вопроизведения восстанавливать хранящуюся информацию при условии, что вероятность ошибки на кубит ориентировочно менее 19%.

Но эта оценка допустимой частоты ошибок (19%) довольно обманчива, поскольку она применима только при весьма нереалистичном допущении, согласно которому кодирование информации и ее восстановление выполняются безупречно, то есть без каких-либо ошибок. В действительности же кодирование и восстановление сами по себе являются сложными квантовыми вычислениями, и в процессе их выполнения неизбежно появление ошибок. Таким образом, нам необходимо найти достаточно устойчивые методы избавления от ошибок, чтобы можно было по-прежнему с высокой точностью восстанавливать квантовую информацию, даже если ошибки совершаются в самом процессе восстановления. Это задача *отказоустойчивого восстановления*, и Питер Шор (Shor 1996) показал в своей пионерской работе, написанной в мае 1996 года, что оно возможно, если частота возникновения ошибок не слишком велика.[1] Конечно, нам необходимо больше, чем просто хранение информации; мы хотим иметь возможность обрабатывать данную информацию и выполнять интересные квантовые вычисления. Поэтому мы должны показать возможность разработки квантовых вентилей, эффективно работающих с квантовой информацией, закодированной так, чтобы быть защищенной от ошибок. В той же работе Шор показал, что отказоустойчивые *вычисления* действительно возможны.

В августе того же года Мэнни Нилл и Раймонд Лафламм (Knill & Laflamme 1996) показали существование порога безошибочности хранения квантовой информации: если частота появления ошибок ниже некоторого критического значения, то возможно сколь угодно длительное хранение неизвестного квантового состояния с высокой точностью воспроизведения. В калтеховской группе (Gottesman *et al.* 1996) мы быстро поняли, что можно объединить идеи Шора и Нилла и Лафламма и показать, что порог безошибочности существует и для вычислений; если частота появления ошибок ниже некоторого критического значения, то возможно выполнение сколь угодно продолжительного квантового вычисления при ничтожной вероятности возникновения ошибок. Аналогичные выводы были сделаны Нилом, Лафламмом и Зуреком (Knill *et al.* 1996, 1997), Аароновым и Бен-Ором (Aharonov & Ben-Or 1996), а также Китаевым (Kitaev 1996b).

[1] Методы отказоустойчивого восстановления независимо развивались Алексеем Китаевым (Kitaev 1996a).

Итак, мы столкнулись с острой проблемой разработки и создания квантового компьютера. Мы можем лишь сказать, насколько хорошо должно работать аппаратное обеспечение этой машины, чтобы оно могло быть использовано для выполнения интересных квантовых вычислений, конкурентноспособных с теми, что могут выполнять лучшие современные цифровые компьютеры. Начнем с обзора основных принципов отказоустойчивых вычислений.

2. Законы отказоустойчивых вычислений

Чтобы выполнять отказоустойчивые вычисления, необходимо: (1) обеспечить *надежное* исправление ошибок, (2) уметь применять вентили, способные *обрабатывать* закодированную информацию, (3) контролировать распространение ошибок. Когда вентиль применяется, скажем, к паре кубитов, один из которых содержит ошибку, то эта ошибка будет стремиться перейти на другой кубит. Необходимо соблюдать осторожность, чтобы сдерживать заражение. Правила, которым мы должны следовать при выполнении отказоустойчивых вычислений, можно систематизировать, как я это буду называть, в виде «законов отказоустойчивых вычислений». Эту информацию можно получить из пионерской работы Шора (Shor 1996).

Первый закон гласит: **(1) Не используйте один и тот же кубит дважды.**[1] Плохая сеть XOR-вентилей, нарушающая эту заповедь, изображена на рис. 2 (используются условные обозначения рис. 1). Бит, который я назвал служебным, является целью нескольких следующих друг за другом вентилей. Если в этом служебном бите есть хотя бы одна ошибка, такая сеть может распространить ее на некоторые другие биты, являющиеся источниками XOR-вентилей; следовательно, инфекция распространяется лавинообразно. Странная квантово-механическая особенность состоит в следующем: тогда как даже классический XOR-вентиль распространяет ошибки инвертирования бита от источника к цели, в случае квантовых вентилей следует помнить о фазовых ошибках, а они распространяются в противоположном направлении от цели к источнику. Итак, при выполнении квантовых вычислений необходимо быть особенно внимательным к возможности распространения ошибок. Следуя этому закону, сеть можно перестроить, как это показано на том же рисунке 2, увеличив количество служебных кубитов до нескольких, так чтобы ни один из них не действовал более одного раза.

Второй закон касается того, как должно выполняться исправление ошибок. С этой целью некоторая информация из блока данных копиру-

[1]Менее строгая, но более благоразумная версия этого закона: избегайте использования одного и того же кубита слишком часто.

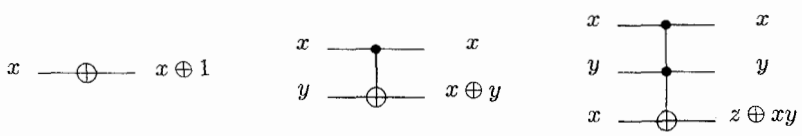


Рис. 1. Графическое изображение вентилей NOT, XOR (исключающее ИЛИ) и вентиля Тоффоли (дважды контролируемое НЕ)

Рис. 2. Первый закон. Плохая схема, которая несколько раз использует один и тот же служебный кубит, и хорошая схема, которая использует каждый вспомогательный бит только один раз

ется в блок служебных кубитов. Затем выполняется его измерение, чтобы найти *синдром ошибки*, который сообщает, какая операция восстановления необходима. Второй закон гласит: **(2) Копируйте ошибки, а не информацию**. Если какая-то часть закодированной информации копируется из блока данных в служебный, то результирующее запутывание между этими регистрами вызовет декогерентизацию и, следовательно, ошибки. Чтобы избежать этого, необходимо перед копированием какой-либо информации подготовить служебные кубиты в специальных состояниях, выбранных таким образом, чтобы результаты их измерения сообщали информацию лишь о возникших ошибках, а не о закодированных данных (см. рис. 3).

Третий и четвертый законы требуют: прежде чем что-либо делать, необходимо убедиться (если это возможно) в том, что это делается правильно. Часто требуется приготовить блок, кодирующий известное квантовое состояние, такое как закодированный нуль. Третий закон гласит: **(3) Кодируя *известное* квантовое состояние, *проверяйте* результат**. В процессе кодирования информация оказывается особенно подверженной действию ошибок: защитные механизмы кода еще не работают и одна-единственная ошибка может распространиться катастрофическим образом. Следователь-

Рис. 3. Второй закон. Служебный кубит должен быть подходящим образом приготовлен

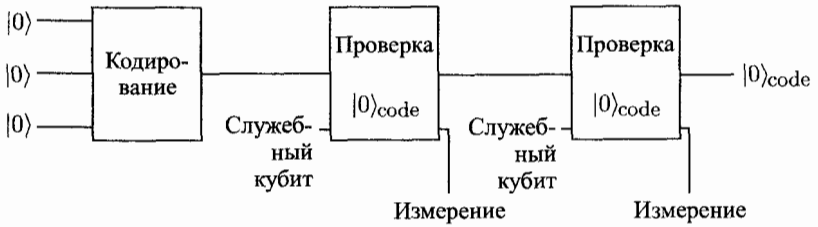


Рис. 4. Третий закон. Кодирующее устройство конструирует $|0\rangle_{\text{code}}$, а затем выполняется неразрушающее измерение (по крайней мере дважды), чтобы убедиться в успешности кодирования

но, необходимо выполнить измерение, проверяющее, правильность кодирования. Конечно, и сама проверка может быть ошибочной, поэтому ее следует повторить несколько раз, пока мы в достаточной мере не убедимся в правильности кодирования.

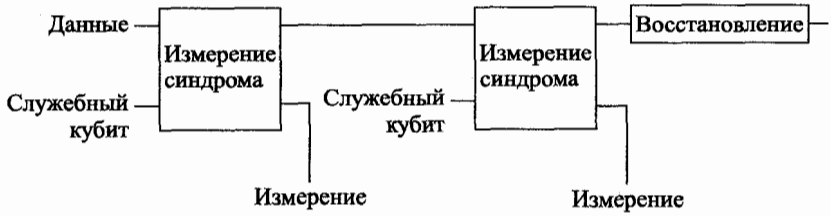


Рис. 5. Четвертый закон. Для обеспечения точности необходимо повторять операции, в частности, измерение синдрома

Необходимость повторения проверок фактически представляет собой частный случай четвертого закона, который гласит: **(4) *Повторяйте* опе-**

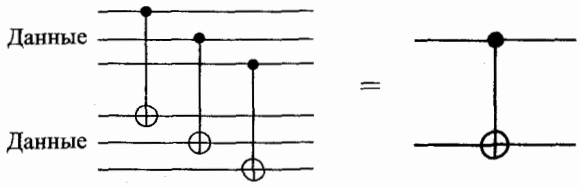


Рис. 6. Пятый закон. Трансверсальная (побитовая) реализация отказоустойчивого XOR-вентиля

рации. Этот закон важно применять к измерению синдрома, предшествующему процедуре исправления. Ошибка, появившаяся в процессе измерения синдрома, может *одновременно повредить данные и явиться причиной* ошибочного синдрома. Если мы по ошибке примем измеренный синдром и будем действовать в соответствии с ним, то вместо исправления это приведет к дальнейшему повреждению данных. Следовательно, перед выполнением восстановления необходимо быть в высшей степени уверенными в том, что измерение синдрома выполнено корректно. Для достижения достаточной уверенности необходимо, в соответствии с четвертым законом, повторить измерение синдрома несколько раз.

Пятый закон является, вероятно, наиболее важным и в то же время наиболее сложным для выполнения: **(5) Используйте *правильный* код.** Используемый для вычислений код должен обладать специальными качествами: допускать применение квантовых вентилей к закодированной информации, эффективно работать и отвечать первым четырем законам. Например, хороший для вычислений код может быть таким, чтобы действующий на закодированные кубиты XOR-вентиль применялся, как показано на рисунке 6: по одному XOR-вентилю, применяемому к каждому кубиту, как в блоке источников, так и в блоке целей. Тогда действующий на закодированные кубиты вентиль удовлетворяет первому закону и является отказоустойчивым.[1]

3. Пример: 7-кубитовый код Стина

Чтобы понять, как осуществляется коррекция квантовых ошибок, поучительно рассмотреть конкретный код. Простым и важным примером ко-

[1]Когда я читал лекцию в декабре, принцип применения отказоустойчивых вентилей для большинства квантовых кодов еще не был известен. Позднее Готтесман (Gottesman 1997a) показал, что отказоустойчивые вычисления возможны для всех «стабилизирующих кодов». Тем не менее, пятый закон — хороший закон; подойдет любой код, но для некоторых из них реализация отказоустойчивых вентилей проще.

да, корректирующего квантовые ошибки, является разработанный Эндрю Стином (Steane 1996ab) 7-кубитовый код. Он позволяет хранить один кубит квантовой информации (произвольное состояние в двумерном гильбертовом пространстве), в совокупности используя семь кубитов (погружая двумерное гильбертово пространство в 2^7-мерное пространство). Фактически код Стина тесно связан с известным нам классическим корректирующим ошибки [7,4,3]-кодом Хэмминга (MacWilliams & Sloane 1977). Чтобы понять, как работает код Стина, для начала полезно разобраться в устройстве кода Хэмминга.

Код Хэмминга использует блок из семи битов для кодирования четырех битов классической информации: имеется $16 = 2^4$ строк длины семь, которые представляют собой кодовые слова, характеризуемые с помощью матрицы контроля четности,

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{1}$$

Каждое кодовое слово является 7-битовой строкой v_{code}, удовлетворяющей уравнению

$$\sum_k H_{jk}(v_{\text{code}})_k = 0 \pmod 2; \tag{2}$$

то есть в арифметике по модулю 2 матрица H уничтожает каждое кодовое слово. Поскольку $Z_2 = \{0, 1\}$ представляет собой (конечное) поле, на нем применимы известные результаты линейной алгебры. Матрица H имеет три линейно независимых строки, а ее ядро[1] является линейной оболочкой четырех независимых векторов-столбцов. 16 кодовых слов представляют собой все возможные линейные комбинации этих четырех строк с коэффициентами, выбираемыми из $\{0, 1\}$.

Теперь предположим, что v_{code} — (неизвестное) кодовое слово, в котором возникла единственная (неизвестная) ошибка: один из семи битов инвертировался. Наша задача — определить поврежденный бит и исправить ошибку. Этот трюк можно выполнить с помощью матрицы контроля четности. Пусть e_i представляет собой строку с единицей в i-й позиции и нулями в остальных. Тогда при инвертировании i-го бита v_{code} становится равным $v_{\text{code}} + e_i$. Если мы подействуем на эту строку матрицей H, то получим

$$H(v_{\text{code}} + e_i) = He_i \tag{3}$$

[1] Ядро матрицы или ее *нуль-пространство* — пространство векторов, удовлетворяющих уравнению (2). — *Прим. ред.*

(поскольку H уничтожает v_{code}), что представляет собой именно i-й столбец матрицы H. А поскольку все столбцы матрицы H различны, то этим однозначно определяется значение i. Выяснив местоположение ошибки, ее можно исправить, инвертируя i-й бит в исходное состояние. Таким образом, если инвертируется только один бит, то можно однозначно восстановить закодированную информацию; но если происходит инвертирование двух или более различных битов, закодированная информация будет повреждена. Замечательно здесь то, что величина He_i выявляет позицию ошибки, ничего не сообщая о v_{code}, то есть не раскрывая закодированную информацию.

Код Стина представляет собой квантовое обобщение этого классического кода коррекции ошибок. Он использует 7-кубитовый «блок» для кодирования одного кубита квантовой информации, то есть произвольного состояния в двумерном гильбертовом пространстве, натянутом на два состояния: «логический нуль» $|0\rangle_{\text{code}}$ и «логическая единица» $|1\rangle_{\text{code}}$. Конструкция кода позволяет исправить любую ошибку, возникшую в любом из семи кубитов в блоке.

Что подразумевается под произвольной ошибкой? Вызывающий подозрение кубит мог подвергаться случайному *унитарному* преобразованию, или *потерять когерентность*, запутавшись с состояниями окружающей среды. Пусть неповрежденный кубит находится в состоянии $a|0\rangle + b|1\rangle$. (Конечно, этот отдельный кубит может быть запутан с другими, так что коэффициенты a и b не обязательно являются комплексными числами; они могут представлять собой векторы состояний, ортогональных как $|0\rangle$, так и $|1\rangle$, которые мы (пока) считаем неуязвимыми для ошибок.) Если теперь кубит поражен произвольной ошибкой, результирующее состояние можно разложить следующим образом:

$$\begin{aligned} a|0\rangle + b|1\rangle \;\rightarrow\; & (a|0\rangle + b|1\rangle) \otimes |A_{\text{no error}} \rangle_{\text{env}} + \\ & + (a|1\rangle + b|0\rangle) \otimes |A_{\text{bit-flip}} \rangle_{\text{env}} + \\ & + (a|0\rangle - b|1\rangle) \otimes |A_{\text{phase-flip}} \rangle_{\text{env}} + \\ & + (a|1\rangle - b|0\rangle) \otimes |A_{\text{both errors}} \rangle_{\text{env}}, \end{aligned} \tag{4}$$

где каждое $|A\rangle_{\text{env}}$ обозначает состояние окружающей среды. Мы не делаем никаких особенных предположений относительно ортогональности или нормировки состояний окружения $|A\rangle_{\text{env}}$,[1] поэтому уравнение (4) не приводит ни к какой потере общности. Таким образом, кубит эволюционирует

[1] Хотя, конечно, совместная эволюция кубита и окружающей среды должна быть унитарной.

к линейной суперпозиции четырех возможностей: (1) никакой ошибки не возникает, (2) возникает инвертирование бита $|0\rangle \leftrightarrow |1\rangle$, (3) обращается относительная фаза $|0\rangle$ и $|1\rangle$, (4) инвертирование бита и обращение фазы происходят одновременно.

Теперь понятно, как должен работать квантовый код коррекции ошибок (Steane 1996; Knill & Laflamme 1997). Совершая подходящее измерение, мы хотим диагностировать, какая из этих четырех возможностей случилась на самом деле. Конечно, в целом, состояние кубита будет представлять линейную комбинацию этих четырех состояний, но наше измерение должно спроецировать его на использованный в (4) базис. После этого мы можем приступить к исправлению ошибки с помощью одного из четырех унитарных преобразований:

$$(1)\quad I,\quad (2)\quad X \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\quad (3)\quad Z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\quad (4)\quad X \cdot Z, \quad (5)$$

(какое из них необходимо применить, укажет результат измерения). Применив это преобразование, мы возвращаем кубит в его исходное состояние и полностью распутываем квантовые состояния кубита и окружающей среды. Существенно, что при диагностике ошибки мы ничего не узнаем о закодированной информации, поскольку получение любой информации о коэффициентах a и b в уравнении (4) неизбежно разрушит когерентность кубита.

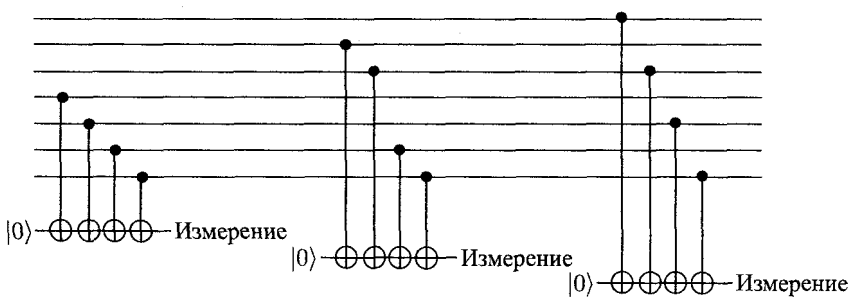


Рис. 7. Вычисление синдрома инвертирования бита для 7-кубитового кода Стина. Повторение вычисления в повернутом базисе диагностирует ошибку обращения фазы. Чтобы сделать процедуру отказоустойчивой, каждый служебный кубит должен быть заменен четырьмя кубитами в подходящих состояниях

Если мы используем код Стина, то удовлетворяющее этим критериям измерение возможно. Логический нуль является равновзвешенной суперпозицией всех кодовых слов Хэмминга (H) с четным весом (с четными количеством единиц),

$$\begin{aligned}|0\rangle_{\text{code}} &= \frac{1}{\sqrt{8}} \sum_{v_{\text{even}} \in \text{H}} |v_{\text{even}}\rangle = \\ &= \frac{1}{\sqrt{8}} \big(|0000000\rangle + |0001111\rangle + |0110011\rangle + |0111100\rangle + \\ &\quad + |1010101\rangle + |1011010\rangle + |1100110\rangle + |1101001\rangle\big),\end{aligned} \tag{6}$$

а логическая единица является равновзвешенной суперпозицией всех кодовых слов Хэмминга с нечетным весом (с нечетным количеством единиц),

$$\begin{aligned}|1\rangle_{\text{code}} &= \frac{1}{\sqrt{8}} \sum_{v_{\text{odd}} \in \text{H}} |v_{\text{odd}}\rangle = \\ &= \frac{1}{\sqrt{8}} \big(|1111111\rangle + |1110000\rangle + |1001100\rangle + |1000011\rangle + \\ &\quad + |0101010\rangle + |0100101\rangle + |0011001\rangle + |0010110\rangle\big).\end{aligned} \tag{7}$$

Поскольку все возникающие в уравнениях (6) и (7) состояния являются кодовыми словами Хэмминга, инвертирование единственного бита в блоке несложно детектировать, выполняя простое квантовое вычисление, как это показано на рисунке 7. Мы дополняем блок из семи кубитов тремя служебными кубитами[1] и выполняем унитарное преобразование:

$$|v\rangle \otimes |0\rangle_{\text{anc}} \rightarrow |v\rangle \otimes |Hv\rangle_{\text{anc}}, \tag{8}$$

где H — матрица контроля четности Хэмминга, а $|\cdot\rangle_{\text{anc}}$ обозначает состояние трех служебных кубитов. Если ошибка содержится лишь в одном из семи кубитов, то измерение служебного кубита проецирует его либо на инвертированное состояние, либо на исходное неповрежденное (но не на любую нетривиальную суперпозицию этих двух состояний). В первом случае результат измерения диагностирует поврежденный бит, ничего сообщая о закодированной в блоке квантовой информации.

Но чтобы выполнять коррекцию квантовых ошибок, нам понадобится диагностировать не только ошибки инвертирования битов, но и фазовые ошибки. Для достижения этой цели мы можем, следуя Стину

[1] Чтобы сделать процедуру отказоустойчивой, нам понадобится увеличить количество служебных кубитов, как это обсуждается в разделе 4.

(Steane 1996ab), изменить базис для каждого кубита, применив поворот Адамара

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{9}$$

Тогда ошибки, бывшие фазовыми в базисе $|0\rangle$, $|1\rangle$, в повернутом базисе

$$|\tilde{0}\rangle \equiv \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big), \quad |\tilde{1}\rangle \equiv \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big) \tag{10}$$

становятся ошибками инвертирования бита. Следовательно, будет достаточно, если наш код способен диагностировать ошибки инвертирования бита в этом повернутом базисе. Но если мы применяем поворот Адамара к каждому из семи кубитов, то логические нуль и единица кода Стина в новом базисе приобретают вид

$$\begin{aligned} |\tilde{0}\rangle &= \frac{1}{4}\sum_{v\in \mathrm{H}} |v\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big), \\ |\tilde{1}\rangle &= \frac{1}{4}\sum_{v\in \mathrm{H}} (-1)^{\mathrm{wt}(v)} |v\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big) \end{aligned} \tag{11}$$

(где $\mathrm{wt}(v)$ обозначает вес v). Важно то, что $|\tilde{0}\rangle$ и $|\tilde{1}\rangle$, как и $|0\rangle$ и $|1\rangle$, являются суперпозициями кодовых слов Хэмминга. Следовательно, в повернутом базисе, как и в исходном, мы можем выполнить контроль четности Хэмминга для диагностики инвертирования битов, которое в исходном базисе представляло собой обращение фазы. При условии, что поврежден только один кубит, выполнение контроля четности в обоих базисах полностью диагностирует ошибку и дает возможность ее исправить.

В описанной выше схеме исправления ошибок я предполагал, что ошибка воздействует только на один кубит в блоке. Очевидно, это предположение нереалистично; обычно все кубиты до некоторой степени запутываются с окружением. Но, как мы уже видели, процедура определения синдрома ошибки обычно проецирует каждый кубит на исходное неповрежденное состояние. Для каждого кубита существует ненулевая предполагаемая малой вероятность возникновения ошибки, которую мы обозначим как ϵ. Сейчас мы сделаем очень важное предположение: ошибки, действующие на разные кубиты в одном блоке, абсолютно не коррелируют между собой. При таком допущении вероятность возникновения двух ошибок имеет порядок ϵ^2 и, следовательно, гораздо меньше вероятности возникновения одной ошибки, если ϵ — достаточно малая величина. Поэтому, с точностью

порядка ϵ, мы можем уверенно сосредоточить наше внимание на случае, когда ошибку содержит максимум один кубит в блоке.

Но в случае (маловероятном), когда в одном и том же блоке кода возникает две ошибки, наша процедура восстановления, как правило, будет безуспешной. Если в одном блоке инвертируются два бита, то контроль четности Хэмминга неправильно диагностирует ошибку. Восстановление вернет квантовое состояние в кодовое подпространство, но в *закодированной* в блоке квантовой информации произойдет инвертирование бита:

$$|0\rangle_{\text{code}} \to |1\rangle_{\text{code}}, \quad |1\rangle_{\text{code}} \to |0\rangle_{\text{code}}. \tag{12}$$

Аналогично, если в одном блоке возникают две фазовые ошибки, то есть две ошибки инвертирования бита в повернутом базисе, то после восстановления в нем произойдет инвертирование бита в повернутом базисе или обращение фазы в исходном базисе:

$$|0\rangle_{\text{code}} \to |0\rangle_{\text{code}}, \quad |1\rangle_{\text{code}} \to -|1\rangle_{\text{code}}. \tag{13}$$

(Если один кубит в блоке содержит фазовую ошибку, а другой — ошибку инвертирования бита, то восстановление будет успешным.)

Итак, мы увидели, что код Стина способен повысить надежность хранения квантовой информации. Предположим, мы хотим сохранить один кубит в неизвестном чистом состоянии $|\psi\rangle$. Вследствие несовершенства запоминающего устройства состояние ρ_{out}, которое мы восстановим, будет иметь точность воспроизведения

$$F \equiv \langle\psi|\rho_{\text{out}}|\psi\rangle = 1 - \epsilon. \tag{14}$$

Но если мы храним кубит, используя 7-кубитовый блоковый код Стина, если каждый из семи кубитов хранится с точностью воспроизведения $F = 1-\epsilon$, если ошибки кубитов некоррелированы и, наконец, если мы в состоянии безукоризненно выполнять коррекцию ошибок, кодирование и декодирование (более подробно об этом ниже), то закодированная информация может храниться с улучшенной точностью воспроизведения $F = 1 - O(\epsilon^2)$.

Кубит в неизвестном состоянии можно закодировать, используя изображенную на рисунке 8 схему. Проще всего понять принцип работы кодирующего устройства, используя альтернативное выражение для матрицы контроля четности Хэмминга,

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{15}$$

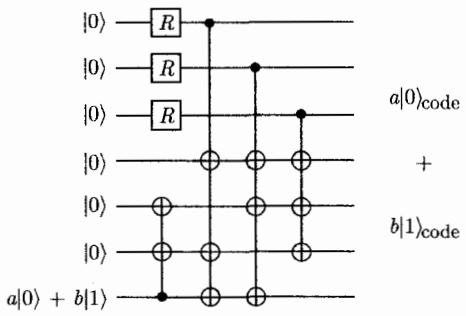


Рис. 8. Кодирующая схема для 7-кубитового кода Стина

(Эта форма матрицы H получена из (1) путем перестановки столбцов, то есть лишь изменением нумерации битов в блоке.) Четный субкод кода Хэмминга фактически является пространством, натянутым на строки матрицы H; итак, мы видим, что (в этом представлении матрицы H) первые три бита строки полностью характеризуют представленные в субкоде данные. Остальные четыре бита представляют собой биты четности, обеспечивающие необходимую для защиты от ошибок избыточность. При кодировании неизвестного состояния $a|0\rangle + b|1\rangle$ кодирующее устройство сначала использует два XOR-вентеля, чтобы приготовить состояние $a|0000000\rangle + b|0000111\rangle$, суперпозицию четного и нечетного слов Хэмминга. Далее по схеме к этому состоянию добавляется $|0\rangle_{\text{code}}$: повороты Адамара (R) готовят равновзвешенную суперпозицию всех восьми возможных значений первых трех битов в блоке, а остальные XOR-вентили включают предписываемые матрицей H биты контроля четности.

Мы также хотим иметь возможность измерять закодированный кубит, скажем, проецируя его на ортогональный базис $\{|0\rangle_{\text{code}}, |1\rangle_{\text{code}}\}$. Если нас не беспокоит разрушение закодированного блока в ходе измерения, то достаточно измерить каждый из семи кубитов в блоке, проецируя их на базис $\{|0\rangle, |1\rangle\}$, а затем, чтобы получить кодовое слово Хэмминга, выполнить классическую коррекцию ошибок в результатах измерения. Четность этого кодового слова является значением логического кубита. (Этап классической коррекции обеспечивает защиту от ошибок измерения. Например, если блок находится в состоянии $|0\rangle_{\text{code}}$, то при измерении элементарных кубитов для определения логического кубита могут возникнуть две независимые ошибки, в результате чего будет получено неверное значение $|1\rangle_{\text{code}}$.)

В применении к квантовым вычислениям, нам понадобится выполнять измерение, проецирующее на $\{|0\rangle_{\text{code}}, |1\rangle_{\text{code}}\}$, без разрушения блока. Это

выполняется путем копирования четности блока на служебный кубит с последующим его измерением. Схема, представляющая неразрушающее измерение кодового блока, показана на рисунке 9.

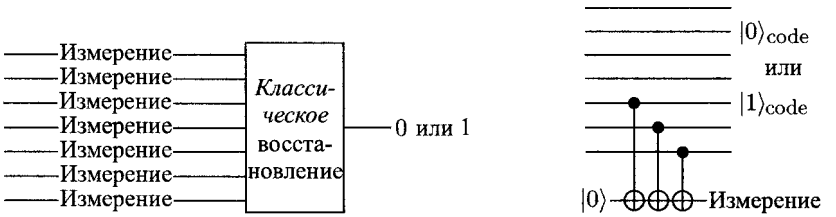


Рис. 9. Разрушающее и неразрушающее измерение логического кубита

7-кубитовый код Стина может может исправить только одну ошибку в кодовом блоке, но можно сконструировать более совершенные коды, способные защищать информацию от ошибок количеством до t внутри одного блока, так что закодированная информация может сохраняться с точностью воспроизведения $F = 1 - O(\epsilon^{t+1})$ (Steane 1996b; Calderbank & Shor 1996; Gottesman 1996; Calderbank *et al.* 1996, 1997). Обзор текущего состояния теории квантового кодирования сделан Шором в работе (Shor 1997).

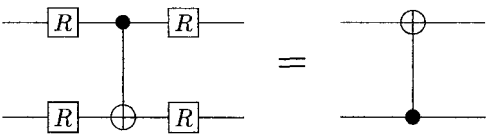


Рис. 10. Полезное тождество. Источник и цель XOR-вентиля меняются местами при изменении базиса с помощью поворотов Адамара

4. Отказоустойчивое восстановление

Конечно, исправление ошибок никогда не будет безупречным. Восстановление само по себе является предрасположенным к появлению ошибок квантовым вычислением. Необходимо позаботиться о создании процедуры восстановления для кода Стина, обеспечивающей вероятность возникновения ошибки порядка ϵ^2 (или порядка ϵ^{t+1} для кода, исправляющего t ошибок). Не менее серьезную проблему представляет контроль распространения ошибок в ходе процедуры восстановления.

Изображенная на рисунке 7 схема не является отказоустойчивой; она нарушает первый закон. XOR-вентили могут перенести единственную

ошибку, возникшую в одном из служебных кубитов, в два разных кубита в блоке данных, приводя в итоге к фазовой ошибке, появляющейся в закодированных данных с вероятностью порядка ϵ. Выполняя требования первого закона, следует увеличить количество служебных кубитов, так чтобы каждый из них являлся целью лишь одного XOR-вентиля. Но пока обратим внимание на второй закон: служебные кубиты должны запутываться с *ошибками* блока данных, а не с закодированной в нем квантовой информацией. Действительно, запутывание служебного кубита с закодированными данными разрушает их когерентность.

Выполняя это условие, мы, прежде чем приступать к процедуре вычисления синдрома, готовим *служебное состояние Шора* (Shor 1996). Это состояние четырех служебных кубитов, представляющее собой равновзвешенную суперпозицию всех строк с четным весом:

$$|\mathrm{Shor}\rangle_{\mathrm{anc}} = \frac{1}{\sqrt{8}} \sum_{v_{\mathrm{even}}} |v_{\mathrm{even}}\rangle_{\mathrm{anc}}. \qquad (16)$$

Чтобы вычислить каждый бит синдрома, мы готовим состояние Шора, выполняем четыре XOR-вентиля (с соответствующими кубитами в блоке данных в роли источников и с четырьмя служебными кубитами в состоянии Шора в роли целей), а затем измеряем служебное состояние. Бит синдрома извлекается из четности результата измерения состояния четырех служебных кубитов. Состояние Шора устроено таким образом, что из него можно извлечь *только* эту четность — никакой другой информации о блоке данных на нем не отпечатывается.

Всего имеется шесть битов синдрома (три для диагностики ошибок инвертирования бита и три для диагностики ошибок обращения фазы), так что измерение синдрома использует 24 служебных кубита, приготовленных в шести состояниях Шора, и 24 XOR-вентиля.

[Одним из способов получения синдрома обращения фазы может быть следующий: сначала применить семь параллельных R-вентилей к блоку данных, чтобы повернуть базис, затем, как показано на рисунке 7, применить XOR-вентили (но с приготовленными в состояниях Шора служебными кубитами) и, наконец, применить семь R-вентилей для обратного поворота данных. Но чтобы усовершенствовать эту процедуру, мы можем использовать представленное на рисунке 10 тождество. Обращая направление XOR-вентилей (то есть используя служебные кубиты в качестве источников, а информацию — в качестве цели), мы можем избежать применения R-вентилей к данным и, следовательно, понизить вероятность их повреждения неправильными вентилями (Zalka 1996; Steane 1997).]

Вследствие распространения ошибок, единственная возникающая в процессе приготовления состояния Шора ошибка может стать результатом возникновения двух фазовых ошибок в данном состоянии. Обе они могут распространиться на данные, если при измерении синдрома используется поврежденное служебное состояние. Следовательно, перед использованием состояния Шора оно должно быть протестировано на наличие нескольких фазовых ошибок (пример третьего закона), как это показано на рисунке 11. Если состояние не проходит тест, его следует уничтожить, и приготовить новое.

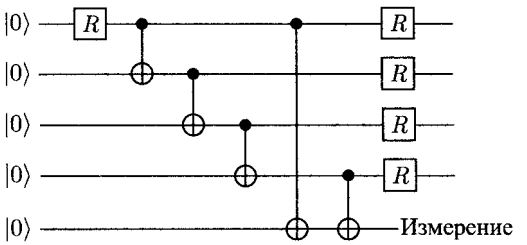


Рис. 11. Конструирование и проверка состояния Шора. Если причиной измерения является 1, то состояние отбрасывается, и готовится новое состояние Шора

Наконец, единственная ошибка в процессе измерения синдрома может стать причиной появления ошибочного синдрома. Таким образом, для подтверждения точности измерение синдрома необходимо повторить (четвертый закон). Когда один и тот же результат получается дважды подряд, он может быть принят с уверенностью, а восстановление продолжено.

Если мы применяем все описанные выше меры предосторожности, то восстановление дает сбой только при появлении двух независимых ошибок, поэтому вероятность возникающей в закодированном блоке ошибки будет порядка ϵ^2. Процедура коррекции ошибок является отказоустойчивой.

Довольно изящная отказоустойчивая процедура измерения синдрома была предложена Стином (Steane 1997). Для измерения синдрома инвертирования бита готовится служебное 7-кубитовое *состояние Стина*

$$|\text{Steane}\rangle_{\text{anc}} = \frac{1}{4}\sum_{v\in H}|v\rangle. \tag{17}$$

(Состояние Стина можно приготовить, применив к состоянию $|0\rangle_{\text{code}}$ побитовый поворот Адамара.) Теперь выполним операции XOR с каждым кубитом блока данных в качестве источника и соответствующим кубитом

служебного состояния в качестве цели, после чего измерим результат действия этих операций на данное служебное состояние. Применяя матрицу контроля четности Хэмминга H к результату *классического* измерения, мы получаем синдром инвертирования бита. (Обратите внимание на соблюдение второго закона: процедура «копирует» данные в служебный блок, состояние которого устроено таким образом, что при его измерении обеспечивается возможность прочтения лишь информации об ошибке.) Аналогичная процедура выполняется в повернутом базисе для определения синдрома обращения фазы. Преимущество процедуры Стина перед процедурой Шора состоит в том, что она требует лишь 14 служебных кубитов и 14 XOR-вентилей. Однако недостаток ее заключается в том, что приготовление служебного состояния более сложное, так что оно само в некоторой степени больше предрасположено к возникновению ошибок.

А что можно сказать об измерении и кодировании? Мы только что отметили, что разрушающее измерение кодового блока надежно, если ошибка содержится только в одном кубите блока. Изображенное на рисунке 9 не разрушающее информацию измерение также не требует модификации. Несмотря на то, что служебное состояние является целью трех последовательных XOR-вентилей, возвращающиеся в блок фазовые ошибки не опасны, так как они не могут поменять $|0\rangle_{\text{code}}$ на $|1\rangle_{\text{code}}$ (или наоборот). Но, поскольку единственная ошибка может оказаться причиной неправильного результата измерения четности, его необходимо повторить (после исправления ошибок), чтобы обеспечить точность порядка ϵ^2 (пример четвертого закона).

В применении к квантовым вычислениям нам потребуется повторно готовить закодированное состояние $|0\rangle_{\text{code}}$. Кодирование может быть выполнено по изображенной на рисунке 8 схеме (за исключением того, что первые два XOR-вентиля можно исключить). Однако ошибки могут распространяться в ходе процедуры кодирования, так что единственной ошибки может быть достаточно для повреждения закодированной информации. Следовательно, важно проверять кодирование, как этого требует третий закон, выполняя неразрушающее измерение блока. На самом деле можно вполне обойтись без этапа кодирования. Каким бы ни было исходное состояние блока, отказоустойчивая коррекция ошибок спроецирует его на пространство, натянутое на $\{|0\rangle_{\text{code}}, |1\rangle_{\text{code}}\}$, а (проверенное) измерение выдаст либо $|0\rangle_{\text{code}}$, либо $|1\rangle_{\text{code}}$. Если получен результат $|1\rangle_{\text{code}}$, то для перевода блока в желаемое состояние $|0\rangle_{\text{code}}$ можно применить (побитовый) оператор NOT.

Кодирование неизвестного квантового состояния выполняется с помощью изображенной на рисунке 8 схемы. Вновь, вследствие распростране-

ния ошибок, единственная ошибка, возникшая в ходе кодирования, может оказаться причиной его сбоя. В этом случае, поскольку никаким измерением нельзя проверить результат кодирования, точность воспроизведения закодированного состояния будет равна $F = 1 - O(\epsilon)$. Тем не менее, смысл в кодировании остается, поскольку закодированное состояние может храниться с требуемой надежностью в течение более продолжительного времени, чем незакодированное.

Обе схемы (Шора и Стина) отказоустойчивого измерения синдрома описаны здесь только для 7-кубитового кода, но их можно приспособить для более сложных кодов, способных корректировать множество ошибок (DiVincenzo & Shor 1996; Steane 1997). По мере возрастания сложности кода схема Стина становится существенно эффективнее схемы Шора.

5. Отказоустойчивые квантовые вентили

Мы убедились, что кодирование может защитить квантовую информацию. Но мы хотим больше, нежели просто *хранить* квантовую информацию с высокой точностью воспроизведения; мы хотим управлять квантовым компьютером, который *обрабатывает* информацию. Конечно, мы могли бы декодировать информацию, выполнить вентиль, а затем кодировать ее заново, но эта процедура на время подвергла бы опасности информацию. Напротив, если мы хотим, чтобы наш компьютер работал надежно, мы должны уметь применять квантовые вентили непосредственно к закодированной информации, а эти вентили должны удовлетворять первому закону отказоустойчивости, если мы хотим избежать катастрофического распространения ошибок.

Действительно, для 7-кубитового кода Стина существует несколько вентилей, которые можно легко применить. Три однокубитовых вентиля могут применяться *побитово*, то есть применение этих вентилей к каждому из семи кубитов в блоке осуществляет тот же вентиль, действующий на закодированный кубит. Мы уже увидели в уравнении (11), что именно так действует поворот Адамара R. То же самое справедливо для NOT-вентиля (поскольку каждое нечетное кодовое слово Хэмминга является дополнением четного кодового слова Хэмминга),[1] а также и вентиля сдвига фазы

$$P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}; \tag{18}$$

нечетные кодовые слова Хэмминга имеют вес $\equiv 3$ (по модулю 4), а четные кодовые слова имеют вес $\equiv 0$ (по модулю 4), поэтому мы фактически при-

[1] Собственно, мы можем выполнить операцию NOT, действующую на закодированный кубит, при помощи лишь трех NOT, применяемых к выбранным кубитам в блоке.

меняем вентиль P^{-1} побитово, чтобы выполнить вентиль P. XOR-вентиль также может выполняться побитово, то есть используя каждый бит блока источника и соответствующий бит блока цели. Это работает, потому что четные кодовые слова образуют субкод, тогда как нечетные кодовые слова представляют его нетривиальный смежный класс.

Таким образом, существуют простые отказоустойчивые процедуры для выполнения NOT-, R-, P- и XOR-вентилей. Но, к сожалению, сами по себе эти вентили не образуют универсального набора. Чтобы уметь выполнять произвольные унитарные преобразования закодированной квантовой информации, нам потребуется сделать подходящее дополнение этого набора. Следуя Шору (Shor 1996), мы добавим 3-кубитовый вентиль Тоффоли, который выполняется с помощью процедуры, изображенной на рисунке 12.[1]

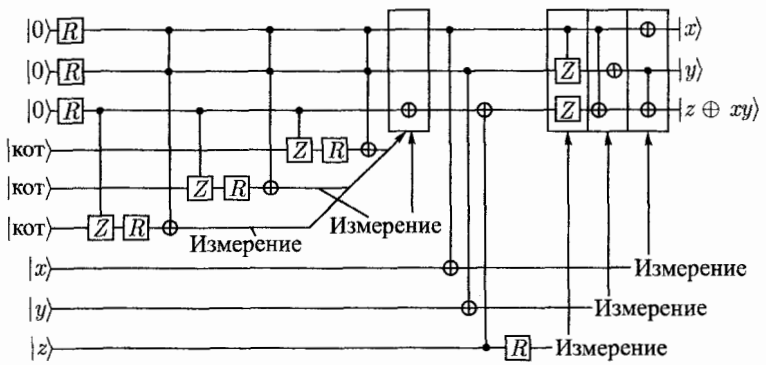


Рис. 12. Отказоустойчивый вентиль Тоффоли. Каждая линия изображает блок из семи кубитов, а вентили применяются трансверсально. Для каждого измерения стрелка указывает на набор вентилей, которые применяются, если результатом измерения является 1, и не действуют, если результатом является 0

Вкратце, процедура работает следующим образом. Во-первых, три закодированных служебных блока приготавливаются в состоянии вида

$$|A\rangle_{\text{anc}} \equiv \sum_{a=0,1} \sum_{b=0,1} |a, b, ab\rangle_{\text{anc}}. \tag{19}$$

[1] Нилл и другие (Knill *et al.* 1996, 1997) описывает альтернативный способ дополнения универсального набора вентилей.

Это приготовление служебного блока выполняется с помощью (проверенного) 7-битового «кот-состояния»

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}\big(|0000000\rangle + |1111111\rangle\big). \tag{20}$$

Выполняется несколько вентилей, включая побитовый вентиль Тоффоли с двумя служебными блоками в качестве управляющих и кот-состоянием в качестве цели. Затем измеряется кот-состояние. (Измерение повторяется, чтобы удостовериться в его точности.) Если результаты измерений четные, значит, желаемое вспомогательное состояние $|A\rangle_{\text{anc}}$ успешно приготовлено; если нечетные, то состояние $|A\rangle_{\text{anc}}$ можно получить, применив NOT-вентиль к третьему служебному блоку.

Тем временем три блока данных терпеливо ожидали готовности служебного блока. Применением трех XOR-вентилей и поворота Адамара состояние данных и служебного блока преобразуется следующим образом:

$$\begin{aligned}&\sum_{a=0,1}\sum_{b=0,1}|a,b,ab\rangle_{\text{anc}}|x,y,z\rangle_{\text{data}} \rightarrow \\ &\quad\rightarrow \sum_{a=0,1}\sum_{b=0,1}\sum_{w=0,1}(-1)^{wz}|a,b,ab\oplus z\rangle_{\text{anc}}|x\oplus a, y\oplus b, w\rangle_{\text{data}}.\end{aligned} \tag{21}$$

Теперь выполняется измерение каждого блока *данных*. Если результатом измерения является 0, никакие действия не предпринимаются, но если результат измерения — 1, то для завершения выполнения вентиля Тоффоли к *служебному* блоку применяется указанный на рисунке 12 набор вентилей. Обратите внимание на то, что исходные блоки данных разрушаются данной процедурой, а также на то, что новыми блоками данных становятся блоки, первоначально бывшие служебными. Важным свойством этой конструкции является то, что все ее этапы организованы так, чтобы удовлетворять требованиям законов отказоустойчивости.

Немного мешает то, что отказоустойчивые вентили формируют дискретный набор, но в то же время это неизбежная черта любой отказоустойчивой схемы. Для отказоустойчивых вентилей нет смысла формировать континуум, ибо как тогда мы сможем избежать совершения ошибки, применяя *ложный* вентиль, который отличается от предназначенного на небольшую величину? В любом случае, поскольку наши отказоустойчивые вентили формируют универсальный набор, их достаточно для приближенного выполнения любого желаемого унитарного преобразования с любой желаемой точностью.

Шор показал, как обобщить этот отказоустойчивый набор вентилей на более сложные коды, способные исправлять большее количество ошибок, а Готтесман (Gottesman 1997ab) описал еще более общую процедуру, которую можно применять к любому из известных квантовых кодов. Таким образом, практически любой корректирующий ошибки квантовый код может использоваться для отказоустойчивых вычислений. В чем же тогда смысл пятого закона? Даже если, в принципе, может использоваться любой код, некоторые из них работают эффективнее. Например, существует 5-кубитовый код, способный корректировать одну ошибку (Bennett *et al.* 1996; Laflamme *et al.* 1996), а Готтесман представил универсальный набор отказоустойчивых вентилей для этого кода. Но реализация этих вентилей достаточно сложна. Для 7-кубитового кода Стина требуется больший блок, но он гораздо удобнее для реализации вычислений; пятый закон отдает предпочтение коду Стина.

6. Порог безошибочности квантовых вычислений

Существуют корректирующие квантовые коды, способные исправить t ошибок, где t может быть сколь угодно большим. Если мы используем такой код и следуем законам отказоустойчивости, тогда непоправимая ошибка возникнет только при появлении $t+1$ *независимых* ошибок в одном блоке до завершения восстановления. Поэтому если вероятность возникновения ошибки в одном квантовом вентиле или отнесенная к единице времени вероятность возникновения ошибки запоминающего устройства имеет порядок ϵ, то вероятность действующей на закодированные данные ошибки в вентиле будет иметь порядок ϵ^{t+1}, что гораздо меньше, чем ϵ, если эта величина достаточно мала. Действительно, может показаться, что при выборе кода со сколь угодно большим t мы можем сделать вероятность ошибки в вентиле сколь угодно малой, но это не факт, по крайней мере для большинства кодов. Проблема состоит в том, что по мере увеличения t сложность кода резко возрастает, соответственно возрастает и сложность процедуры восстановления. В конечном счете мы достигаем момента, когда выполнение восстановления требует так много времени, что становится вероятным накопление $t+1$ ошибок в блоке до завершения восстановительного этапа, и, таким образом, способность кода исправлять ошибки становится сомнительной.

Предположим, что количество вычислительных шагов, необходимых для выполнения измерения синдрома, растет вместе с t, как степень t^b. Тогда вероятность того, что до завершения измерения накопится $t+1$ ошибок, будет вести себя как

$$\text{Block Error Probability} \sim \left(t^b\epsilon\right)^{t+1}, \tag{22}$$

где ϵ — вероятность ошибки на один шаг. Мы можем в таком случае выбрать t, чтобы минимизировать вероятность ошибки ($t \sim e^{-1}\epsilon^{-1/b}$, при условии, что t большое), получая

$$\text{Minimum Block Error Probability} \sim \exp\left(-e^{-1}b\epsilon^{-1/b}\right). \tag{23}$$

Таким образом, если мы рассчитываем безупречно выполнить в совокупности T циклов коррекции ошибок, то наши вентили должны иметь точность

$$\epsilon \sim (\log T)^{-b}. \tag{24}$$

Аналогично, для выполнения квантового вычисления с участием T квантовых вентилей необходимы элементарные вентили заданной точности.

В первоначально описанной Шором (Shor 1996) процедуре степень, характеризующая сложность измерения синдрома, равна $b = 4$; с помощью более оптимизированной процедуры можно достичь скромного улучшения ($b \sim 3$). Размер блока используемого кода растет вместе с t как t^2, поэтому, когда выбирается код для оптимизации вероятности возникновения ошибки, размер блока будет порядка $(\log T)^2$. Конечно, описываемый уравнением (24) скейлинг гораздо предпочтительнее, чем точность $\epsilon \sim T^{-1}$, которая потребовалась бы, если бы кодирование не использовалось вообще. Но при любой заданной точности существует предел продолжительности вычисления, при которой еще можно не опасаться появления ошибок.

Это ограничение можно преодолеть, используя код специального типа, а именно *каскадный* код (Knill & Laflamme 1996; Knill *et al.* 1996, 1997; Aharonov & Ben-Or 1996; Kitaev 1996, 1997). Чтобы понять идею каскадного кода, представьте, что мы используем корректирующий ошибки квантовый код Стина, кодирующий единственный кубит в 7-кубитовом блоке. Но если мы посмотрим внимательнее, с бóльшим разрешением, на один из семи кубитов в блоке, то обнаружим, что на самом деле это не один кубит, а другой 7-кубитовый блок, закодированный, как и ранее, при помощи того же кода Стина. А когда мы изучим один из семи кубитов в *этом* блоке с еще бóльшим разрешением, мы обнаружим, что он тоже в действительности является блоком из семи кубитов, и так далее (см. рис. 13). Если в этой иерархии каскадного соединения всего имеется L уровней, тогда один кубит фактически кодируется в блоке размера 7^L. Каскадное соединение оказывается полезным, поскольку, действуя по принципу «разделяй и властвуй», оно позволяет более эффективно избавляться от ошибок: то есть чаще всего восстановление осуществляется на самом нижнем уровне иерархии, в блоке размера семь, реже — на следующем уровне, в блоке размера $7^2 = 49$, еще реже на следующем уровне, на блоке размера $7^3 = 343$,

и так далее. При таком методе сложность коррекции уже не так стремительно растет с повышением способности квантового кода исправлять ошибки.

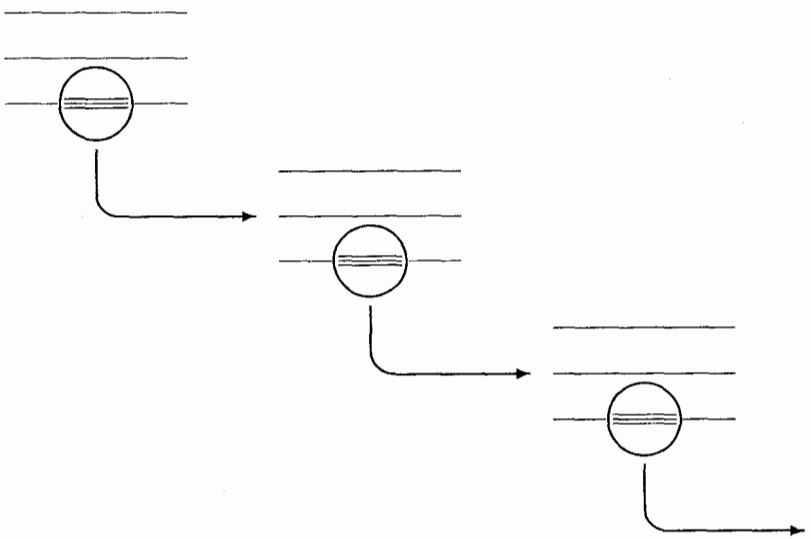


Рис. 13. Каскадное кодирование. Каждый кубит в блоке, рассматриваемый с большим разрешением, сам является закодированным субкодом

Мы видели, что 7-кубитовый код Стина может исправлять одну ошибку. Если вероятность ошибки на кубит равна ϵ, ошибки не коррелированы, а процедура отказоустойчива, то вероятность сбоя при восстановлении будет порядка ϵ^2. Если два кода соединяются в каскад, образуя блок размера 7^2, то ошибка в нем возникает только при повреждении двух из его субблоков размера семь, что происходит с вероятностью порядка $(\epsilon^2)^2$. При добавлении еще одного уровня каскадирования ошибка в блоке размера 7^3 возникает только в случае повреждения двух из его субблоков размера 7^2, а вероятность такого события имеет порядок $\left((\epsilon^2)^2\right)^2$. И так далее, при наличии L уровней каскадного соединения кодов вероятность появления ошибки имеет порядок ϵ^{2^L}, тогда как размер блока равен 7^L. Теперь, если частота появления ошибок для основных вентилей достаточно мала, то отнесенную к одному вентилю вероятность появления ошибки можно уменьшить с помощью каскадного соединения кодов. Если это так, то добавление

следующего уровня каскадирования приведет к дальнейшему понижению вероятности ошибки, и так далее. В этом состоит природа порога безошибочности квантовых вычислений: если кодирование значительно снижает вероятность возникновения ошибки, то, добавляя достаточное количество уровней каскадирования, частоту появления ошибок можно сделать сколь угодно малой. Но если исходная частота появления ошибок слишком высока, то кодирование, напротив, усугубит такое положение вещей.

Чтобы проанализировать эту ситуацию, необходимо принять некоторую конкретную модель ошибок. Я выберу наиболее простую из возможных квазиреалистичную модель: некоррелированные стохастические ошибки.[1] На каждом такте вычисления каждый кубит в устройстве запутывается с окружением. Пусть запутывание описывается уравнением (4) с тем лишь отличием, что теперь четыре состояния окружения предполагаются взаимно ортогональными, а «ошибочные состояния» — имеющими одинаковые нормы. Таким образом, три типа ошибок (инвертирование бита, обращение фазы и обе ошибки одновременно) предполагаются равновероятными. Полная вероятность появления ошибки на каждом шаге обозначается как ϵ_{store}. Кроме этих ошибок запоминающего устройства, поражающих хранящиеся кубиты, существуют ошибки, вносимые самими квантовыми вентилями. Для каждого типа вентиля вероятность появления ошибки при каждом его применении обозначается как ϵ_{gate} (при независимых значениях, приписываемых каждому типу вентилей). Если вентиль действует на более чем один кубит (XOR или Тоффоли), то могут возникать коррелирующие ошибки. Сделаем пессимистическое предположение, что ошибка многокубитового вентиля всегда повреждает все кубиты, на которые он действует; например, неправильный XOR-вентиль вводит ошибки одновременно в кубит источника и кубит цели. Это допущение (среди прочих) делается только для упрощения анализа. При более реалистичных предположениях мы, безусловно, обнаружили бы, что вполне можно пережить и несколько более высокую частоту появления ошибок.

Эффективность каскадного кодирования можно проанализировать, построив систему *потоковых уравнений*, описывающих эволюцию модели ошибки при переходе от одного уровня каскадирования к другому. Пусть, например, мы хотим выполнить XOR-вентиль с последующим этапом исправления ошибок в кубитах, закодированных с помощью каскадного кода Стина, содержащего L уровней (размер блока 7^L). Эти процедуры можно описать с помощью таких же операций, но действующих на субблоки размера 7^{L-1}. Таким образом, вероятность появления ошибки $\epsilon^{(L)}$ для вен-

[1] В разделе 9 я прокомментирую, как изменяется анализ при выборе другой модели ошибок.

тиля, действующего на блок размера L, может быть выражена через вероятность появления ошибки $\epsilon^{(L-1)}$ для вентиля, действующего на блок размера $L-1$. Это соотношение представляет собой одно из потоковых уравнений. В принципе, решая систему потоковых уравнений, можно получить выражение для вероятности появления ошибки «на уровне L» через параметры модели ошибок и исследовать ее поведение с ростом L. Если вероятность появления ошибки в блоке стремится к нулю с ростом его размера L, то это это означает, что вероятности элементарных ошибок лежат «ниже порога». Поскольку вероятности элементарных ошибок могут зависеть от множества параметров, то на самом деле порог представляет собой некоторую гиперповерхность в многомерном параметрическом пространстве рассматриваемой модели.

В общем случае потоковые уравнения довольно сложны; их подробное обсуждение можно найти в работах Готтесмана (Gottesman 1997b), Готтесмана и других (Gottesman *et al.* 1996) или Залки (Zalka 1996). Но для грубой иллюстрации идеи предположим, что ошибки хранения отсутствуют, а единственным источником ошибок являются XOR-вентили, характеризуемые отнесенной к одному вентилю вероятностью появления ошибки $\epsilon_{\rm XOR}$. Допустим, мы хотим применить сеть XOR-вентилей, действующих на закодированные блоки. Оценим, насколько малой должна быть частота появления ошибок $\epsilon_{\rm XOR}$, достигаемая действующими на закодированную информацию вентилями, чтобы эта величина была меньше частоты ошибок для элементарных XOR-вентилей.

Если воспользоваться состоянием Шора в качестве служебного, то для выявления синдромов инвертирования бита и обращения фазы потребуется 12 XOR-вентилей,[1] следовательно, вероятность появления ошибки для каждого измерения синдрома равна

$$\epsilon_{\rm syndrome} \sim 12\epsilon_{\rm XOR}. \tag{25}$$

Чтобы убедиться в надежности синдрома, он повторно измеряется до тех пор, пока дважды подряд не получается один и тот же результат.[2] Из простой комбинаторики можно найти, что вероятность возникновения двух независимых ошибок до успешного завершения измерения синдрома равна $5\epsilon^2_{\rm syndrome}$. Предположим, что обе эти ошибки повреждают данные, тогда они являются причиной появления непоправимой ошибки в блоке. Поскольку нам нужно измерить как синдром инвертирования бита, так и син-

[1] При использовании служебного состояния Стина это число было бы сокращено до 7; однако, поскольку приготовление состояния Стина сложнее, служебное состояние будет в большей степени подвержено ошибкам.

[2] На самом деле это не наилучшая процедура, но здесь мы ее примем для простоты.

дром обращения фазы, мы видим, что вероятность сбоя при каждом выполнении исправления равна

$$\epsilon_{\text{fail}} \sim 2 \cdot 5 \cdot (12\epsilon_{\text{XOR}})^2 = 1440\epsilon_{\text{XOR}}^2. \tag{26}$$

При реализации сети XOR-вентилей мы выполняем исправление в каждом блоке после того, как на него подействовали N вентилей, где N — число, которое будет выбрано для оптимизации количества ошибок на вентиль. Каждый трансверсальный XOR-вентиль, действующий на блок, выполняется путем применения семи элементарных XOR-вентилей. После применения к данному блоку N вентилей и следующего вслед за этим этапа восстановления, отнесенная к одному вентилю вероятность накопления двух ошибок в блоке не превосходит

$$\epsilon_{\text{fail}} \sim \frac{1}{N}\left(\frac{1}{2}7N \cdot 6N \cdot \epsilon_{\text{XOR}}^2 + 7N \cdot \epsilon_{\text{XOR}} \cdot 2 \cdot 2 \cdot \epsilon_{\text{syndrome}} + 1440 \cdot \epsilon_{\text{XOR}}^2\right) =$$

$$= \frac{1}{N}\left(21N^2 + 336N + 1440\right)\epsilon_{\text{XOR}}^2. \tag{27}$$

(Возможны три варианта: либо обе ошибки внесены вентилями, либо одна из них происходит от вентиля, а другая возникла на этапе восстановления, либо наконец, обе они возникли в процессе исправления.[1] Мы пессимистично предполагаем, что все ошибки повреждают данные в блоке и что при возникновении двух ошибок в блоке данных исправление всегда дает сбой.) Минимум в уравнении (27) достигается при $N = 8$, тогда мы получаем $\epsilon_{\text{fail}} = 684\epsilon_{\text{XOR}}^2$, то есть кодирование имеет смысл при $684\epsilon_{\text{XOR}}^2 < \epsilon_{\text{XOR}}$, или $\epsilon_{\text{XOR}} < (684)^{-1} \sim 1.5 \cdot 10^{-3}$. Это наша «пороговая оценка».

Даже если бы все ошибки возникали из-за XOR-вентилей, по ряду причин эта оценка была бы неадекватна. Более того, мы предположили, что служебные состояния (состояния Шора), используемые при измерении синдрома, свободны от ошибок, в то время как для их приготовления и проверки фактически используются те же самые XOR-вентили. К тому же, когда XOR-вентили действуют на субблоки каскадного кода, не всегда возможно применение оптимального числа вентилей перед исправлением. И, конечно, более полный анализ должен включать ошибки запоминающего устройства и отслеживать эволюцию связанных потоков ошибок хранения и ошибок вентилей, обусловленную добавлением очередного уровня каскадирования.

[1] Из двух коэффициентов 2 во втором слагаемом уравнения (27) один возникает вследствие того, что измеряются оба синдрома — инвертирования бита и обращения фазы, второй — потому что ошибка может возникнуть как в ходе первого измерения синдрома, так и при его повторении.

Действительно, когда мы осуществляем исправление на L-уровне, состояние Шора должно быть построено из блоков, закодированных на $(L-1)$-уровне. Важной частью анализа является тщательный контроль того, насколько шумящими являются эти служебные состояния, с целью определить вероятность отказа при исправлении на L-уровне. Блоки данных должны ожидать приготовления служебных состояний; тем временем в них накапливаются ошибки хранения. А так как период ожидания не масштабируется просто уровнем L, потоковые уравнения не являются полностью автомодельными — существует некоторая явная «временная зависимость» потока (то есть L-зависимость). Но по-прежнему качественно верно, что порог определяется требованием, чтобы отказоустойчивая процедура с одним уровнем каскадирования действительно улучшала надежность вентилей, как в приведенном выше примере.

Хотя потоковые уравнения слишком сложны для точного решения, при умеренных предположениях можно получить их приближенные решения. Грубо говоря, выводы следующие (Gottesman *et al.* 1996; Gottesman 1997b). Если ошибки хранения пренебрежимо малы, то пороговая частота появления ошибок на один вентиль имеет порядок 10^{-4}. Если доминируют ошибки хранения, то пороговая частота появления ошибок в течение одного такта имеет порядок 10^{-5} (где один такт представляет собой интервал времени, необходимого для выполнения одного вентиля). От элементарных вентилей Тоффоли не требуется такая точность, как от одно- и двухкубитовых вентилей, — вполне приемлема частота ошибок вентиля Тоффоли порядка 10^{-3}, если достаточно малы другие частоты появления ошибок. (Это приятно, поскольку вентили Тоффоли наиболее сложны в применении, и на практике вероятна их меньшая точность.)

Также следует задать вопрос, какой размер блока необходим для обеспечения некоторой определенной точности. Иначе говоря, если пороговая частота ошибки в вентиле равна ϵ_0, а фактическая частота ошибки в элементарном вентиле $\epsilon < \epsilon_0$, то L-кратное каскадирование кода понизит частоту возникновения ошибок до

$$\epsilon^{(L)} \sim \epsilon_0 \left(\frac{\epsilon}{\epsilon_0}\right)^{2^L}. \tag{28}$$

Таким образом, чтобы быть достаточно уверенными в том, что мы можем завершить вычисление с T вентилями без единой ошибки, мы должны выбрать размер блока 7^L порядка

$$\begin{matrix}\text{block}\\ \text{size}\end{matrix} \sim \left[\frac{\log \epsilon_0 T}{\log \epsilon_0/\epsilon}\right]^{\log_2 7}. \tag{29}$$

Если каскадный код имеет размер блока n и может исправить $t+1$ ошибок, показатель степени $\log_2 7 \sim 2.8$ в уравнении (29) заменяется на $\log n / \log(t+1)$; для семейства кодов, рассмотренных Шором, этот показатель стремится к 2, но для «хороших» кодов, в принципе, может стремиться к 1.

Если частоты ошибок лежат ниже порога безошибочности, также возможно сколь угодно продолжительное хранение *неизвестного* квантового состояния. Однако, как мы уже отмечали в разделе 4, если вероятность ошибки запоминающего устройства на один такт вычисления равна ϵ, то исходное кодирование состояния может быть выполнено с точностью воспроизведения не выше, чем $F = 1 - O(\epsilon)$. При каскадном кодировании мы можем бесконечно долго хранить неизвестную квантовую информацию с достаточно хорошей, но не со сколь угодно высокой, точностью воспроизведения.

Допущения, лежащие в основе этих выводов, будут еще раз рассмотрены в разделе 9.

7. Отказоустойчивая факторизация

Чтобы понять практическое значение скейлингого закона (29), рассмотрим применение каскадного кодирования для реализации квантового алгоритма факторизации Шора (Shor 1994). Алгоритм состоит из двух частей. Для факторизации числа N сначала вычисляется показательная функция по модулю N, чтобы приготовить состояние вида

$$\sum_x |x\rangle_{\text{input}} \otimes |a^x \pmod N\rangle_{\text{output}}, \tag{30}$$

где $a\,(< N)$ — взаимно простое с N. Затем выполняется преобразование Фурье, действующее на входной регистр. Наконец, входной регистр измеряется и выполняется некоторая классическая постобработка, чтобы найти кандидата на роль простого множителя числа N. Приготовление состояния (30) (с применением перечисленных в разделе 5 отказоустойчивых вентилей) описано в работах Бэкмана и др. (Beckman *et al.* 1996) и Ведрала и др. (Vedral *et al.* 1996). Преобразование Фурье требует комментариев. Как показано Гриффитсом и Ниу (Griffits & Niu 1996), преобразование Фурье может быть измерено с помощью однокубитовых вентилей (но при условии, что тип вентиля определяется результатами предыдущих измерений). В частности, для этого необходимо уметь выполнять вентили поворота фазы

$$P(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}. \tag{31}$$

Впрочем, достаточно иметь в распоряжении отказоустойчивую процедуру реализации $P(\theta_0)$ при некотором значении θ_0, иррациональном кратном 2π. Тогда повторное применение $P(\theta_0)$ позволяет сколь угодно точно выполнить $P(\theta)$ при любом θ.

Чтобы построить $P(\theta_0)$, необходимо использовать два служебных кубита и применить два вентиля Тоффоли. Одна из рабочих схем изображена на рисунке 14.[1] Если результатом измерения двух служебных кубитов является $|00\rangle_{\text{ancilla}}$, что происходит с вероятностью $5/8$, то это означает, что схема успешно применила $P(\theta_0)$ к информационному кубиту, где $e^{i\theta_0} = (1+3i)/(3+i)$, или $\cos\theta_0 = 3/5$. Любой другой результат ($|01\rangle$, $|10\rangle$ или $|11\rangle$, каждый из которых возникает с вероятностью $1/8$), означает сбой в работе схемы. Но в этом случае, применяя Z-вентиль, можно восстановить состояние поврежденного кубита и продолжить попытки применения этой схемы вплоть до их успешного завершения.

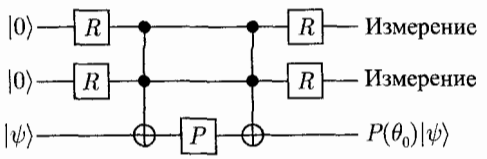


Рис. 14. Схема фазовращателя. Результат измерения $|00\rangle$ означает, что к кубиту данных применен поворот $P(\theta_0)$, где $\cos\theta_0 = 3/5$

В качестве альтернативного метода можно собрать «автономную» библиотеку «угловых кубитов» вида

$$|\theta\rangle = \frac{1}{\sqrt{2}}\bigl(|0\rangle + e^{i\theta}|1\rangle\bigr). \tag{32}$$

Тогда фазовращатель $P(\theta)$ можно реализовать, применяя вентиль CNOT с кубитом данных в качестве управляющего и библиотечным угловым кубитом $|\theta\rangle$ в качестве цели, как это показано на рисунке 15. Затем мы измеряем библиотечный кубит. Результат измерения $|0\rangle$ получается с вероятностью $\frac{1}{2}$, и в этом случае мы успешно применили $P(\theta)$ к информационному кубиту. Но если в результате измерения получено $|1\rangle$, то это означает, что вместо $P(\theta)$ к состоянию $|\psi\rangle$ применено преобразование $P(-\theta)$. В этом случае

[1] Кроме двух вентилей Тоффоли, на этой схеме изображены четыре вентиля Адамара, обозначаемые здесь как R, и фазовращатель $P\,[= P(\pi/4)]$. — *Прим. ред.*

предпринимается еще одна попытка, но на этот раз, чтобы компенсировать ошибку предыдущего шага, применяется схема $P(2\theta)$. Вероятность n сбоев подряд равна всего лишь 2^{-n}.

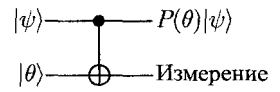


Рис. 15. Схема фазовращателя, использующая в качестве служебного библиотечный кубит. Если результатом измерения является $|0\rangle$, то $P(\theta)$ успешно применен к кубиту данных

Теперь попытаемся факторизовать K-битовое число, используя схему факторизации Бэкмана и др. (Beckman *et al.* 1996) и первый из описанных выше методов выполнения однокубитовых поворотов, требуемых для преобразования Фурье. Чтобы выполнить преобразование Фурье (на регистре с $2K$ кубитами), необходимы $2K$ однокубитовых фазовращателей. Преобразование Фурье должно вычисляться с высокой точностью. Для того чтобы алгоритм факторизации имел достаточные шансы на успех, достаточно выполнение каждого фазового поворота с точностью порядка K^{-1}. Мы можем достичь этой точности, компонуя $P(\theta_0)$ порядка K раз. Поскольку преобразование Фурье требует порядка K^2 вентилей Тоффоли, то для большого K сложность данного алгоритма обусловлена вычислением модулярной показательной функции, которая, согласно Бэкману и др. (Beckman *et al.* 1996)[1], требует $38K^3$ вентилей Тоффоли.

При наличии лучших из известных классических алгоритмов и самых быстродействующих из существующих машин факторизация 130-разрядного числа ($K \sim 430$ бит) займет порядка нескольких месяцев (Lenstra *et al.* 1996). Зададим вопрос, какими возможностями должен обладать квантовый компьютер для выполнения этого задания. От него потребуется способность хранить $5K \sim 2150$ закодированных кубитов, а также выполнить порядка $3 \cdot 10^9$ вентилей Тоффоли. Чтобы достичь достаточно высокой вероятности выполнения безошибочного вычисления, нам бы хотелось, чтобы вероятность появления ошибки на один вентиль Тоффоли была менее 10^{-9}, а вероятность ошибки хранения на полное время выполнения вентиля — менее 10^{-12}. Согласно потоковым уравнениям каскадирования, такие частоты возникновения ошибок могут быть достигнуты для закодиро-

[1]Описанный Бэкманом и др. алгоритм фактически требует $46K^3$ вентилей Тоффоли, но это число может быть снижено до $38K^3$ при использовании предложенного Ричардом Хагесом (Hughes 1997) усовершенствованного алгоритма сравнения.

ванных данных, если частоты ошибок на уровне индивидуальных кубитов равны $\epsilon_{\text{store}} \sim \epsilon_{\text{gate}} \sim 10^{-6}$ и если используются три уровня каскадирования, так что размер кодирующего каждый кубит блока равен $7^3 = 343$. С учетом дополнительных служебных кубитов, необходимых для выполнения вентилей и (параллельного) исправления ошибок, общее число содержащихся в машине кубитов будет порядка 10^6.

При частоте ошибок порядка 10^{-6} для индивидуальных кубитов, каскадное соединение 7-кубитового кода может оказаться наиболее эффективной отказоустойчивой процедурой. Для частот появления ошибок меньшего порядка величины лучше воспользоваться более сложным (некаскадным) кодом, способным исправить несколько ошибок в одном блоке (Shor 1996). При еще более низких частотах появления ошибок можно использовать коды, которые более эффективно используют пространство памяти путем кодирования множества кубитов в одном блоке (Gottesman 1997a). В принципе, когда частота ошибок для индивидуальных кубитов очень низкая для выполнения отказоустойчивого вычисления становится возможным найти хороший код, такой, чтобы отношение числа закодированных кубитов к их общему количеству приближалось к единице.

8. Выявление квантовых утечек

Сейчас я хотел бы более подробно рассмотреть одно из предположений, сделанных в предыдущем анализе. Мы проигнорировали вероятность *утечки*. В нашей модели квантового компьютера каждый из кубитов живет в двумерном гильбертовом пространстве. Мы предположили, что при возникновении ошибки этот кубит либо запутывается с окружением, либо поворачивается в двумерном пространстве в непредсказуемом направлении. Но существует другой возможный тип ошибки, когда кубит из двумерного просачивается в более широкое пространство (Plenio & Knight 1996). Например, в компьютере на ионных ловушках квантовую информацию можно хранить в двумерном пространстве, натянутом на основное состояние иона и некоторое долгоживущее метастабильное состояние (Cirac & Zoller 1995). Но в процессе работы устройства ион может совершить неожиданный переход в другое состояние. Если это состояние быстро распадается в основное состояние, то ошибка может быть обнаружена и исправлена с помощью стандартных на данный момент методов отказоустойчивого исправления квантовых ошибок. Но если ион остается подвешенным в ложном пространстве в течение длительного времени, то эти методы приведут к сбою.

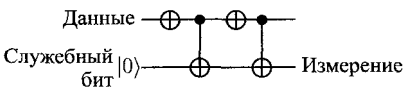


Рис. 16. Схема детектирования квантовой утечки. Если утечка произошла, то результат измерения равен $|0\rangle$, в противном случае — $|1\rangle$

Одна из стратегий решения этой проблемы состоит в определении наиболее вероятных уровней утечки и периодическое возвращение их в основное состояние, но при большом количестве таких уровней этот метод может оказаться слишком громоздким. Более элегантным решением является обнаружение утечки, но без попытки точно диагностировать, что произошло с просочившимся кубитом. Например, в ионной ловушке можно сконструировать CNOT-вентиль, действующий тривиально, если управляющий бит не находится в двумерном пространстве, которому принадлежит кубит. Тогда мы можем выполнить изображенную на рисунке 16 последовательность вентилей. При утечке информационного кубита ничего не происходит, но в случае отсутствия всякой утечки инвертируется служебный бит. Затем мы измеряем служебный бит. Результат измерения $|1\rangle$ проецирует кубит на правильное гильбертово пространство, тогда как результат $|0\rangle$ проецирует кубит на просочившееся состояние.

Если возникла утечка, то поврежденный кубит отбрасывается[1] и заменяется новым кубитом в стандартном состоянии, скажем, в состоянии $|0\rangle$. (Если используется каскадное кодирование, диагностика утечки должна осуществляться только на самых нижних уровнях кодирования.) После этого выполняется обычное измерение синдрома, которое спроецирует кубит на такое состояние, что ошибку можно будет исправить с помощью простого унитарного преобразования. Поскольку еще до измерения синдрома известно, что поврежденный кубит находится в конкретной позиции внутри блока, можно применить модернизированную версию исправления ошибок, разработанную с целью диагностики и исправления ошибок, находящихся в известных позициях (Grassl *и др.* 1996).

9. Машина мечты

Вспомним некоторые важные допущения, сделанные нами при оценке порога безошибочности:

[1] Конечно, позднее, после возвращения его в основное состояние, этим кубитом можно пользоваться снова.

- **Случайные ошибки.** Мы предположили, что ошибки не имеют никакой систематической составляющей.[1] Это допущение позволяет складывать вероятности, а не их амплитуды, для оценки того, как с течением времени растет вероятность появления ошибки. Систематические ошибки будут накапливаться гораздо быстрее, и, следовательно, допустимая частота ошибок будет намного меньше. Иначе говоря, если порог безошибочности для случайных ошибок равен ϵ_0, то для максимально законспирированных систематических ошибок он будет иметь порядок ϵ_0^2. Моя позиция такова: (1) даже если наши аппаратные средства предрасположены к совершению ошибок с систематическими фазами, эти ошибки будут стремиться к взаимному уничтожению в процессе достаточно продолжительного вычисления (Obenland & Despain 1996ab; Miquel *et al* 1997), и (2) поскольку систематические ошибки в принципе можно понять и устранить, с фундаментальной точки зрения более важным является понимание ограничений, накладываемых на работу машины случайными ошибками.
- **Некоррелированные ошибки.** Мы предположили, что ошибки являются некоррелированными как в пространстве, так и во времени. Таким образом, когда мы говорим, что вероятность ошибки на один кубит равна $\epsilon \sim 10^{-5}$, мы фактически имеем в виду, что для любой пары кубитов вероятность их одновременного повреждения ошибками имеет порядок $\epsilon^2 \sim 10^{-10}$. Это сильное и важное предположение, поскольку при возникновении нескольких ошибок в одном и том же блоке кода наши схемы кодирования дают сбой. Будущие квантовые инженеры столкнутся с проблемой обеспечения того, чтобы это предположение достаточно хорошо работало в конструируемых ими устройствах.
- **Максимальный параллелизм.** Мы предположили, что многие квантовые вентили могут выполняться параллельно на каждом такте вычислений. Это допущение дает возможность одновременного осуществления исправления ошибок во всех кодовых блоках, поэтому оно важно для контроля ошибок хранения кубитов. (Другими словами, добавление к коду следующего уровня каскадирования привело бы к увеличению вероятности сбоя, поскольку каждому не занятому в процессе отдельному кубиту пришлось бы дольше ожидать своей очереди исправления ошибок.) Если мы игнорируем ошибки запоминающего устройства, то в анализе порога безошибочности параллельная работа не является необходимой, но она, конечно, желательна для ускорения процесса вычисления.

[1] Нилл и др. (Knill *et al.* 1996, 1997) доказали существование порога безошибочности для гораздо более общих моделей ошибки.

- **Независимая от количества кубитов частота возникновения ошибок.** Мы предположили, что частота ошибок не зависит от количества хранящихся в устройстве кубитов. Неявно это предположение относится к природе аппаратных средств. Например, оно было бы необоснованным, если бы все кубиты хранились в единственной ионной ловушке и делили бы один фононный канал передачи информации. (Cirac & Zoller 1995).
- **Вентили могут действовать на любую пару кубитов.** Мы предположили, что наша машина оборудована набором фундаментальных вентилей, которые могут применяться к любой паре хранящихся кубитов (или к тройке кубитов в случае вентиля Тоффоли), независимо от степени их близости друг к другу. На практике возможны издержки как по времени выполнения, так и по частоте появления ошибок, связанные с перемещением кубитов для того, чтобы вентиль мог действовать эффективно на определенную пару. Оставим проблему выбора архитектуры, минимизирующей эти затраты, будущим конструкторам машин.
- **Новые служебные кубиты.** Мы предположили, что наш компьютер имеет доступ к достаточному запасу свежих служебных кубитов. Служебные кубиты используются как для выполнения вентилей (Тоффоли), так и для осуществления исправления ошибок. С накоплением эффектов случайных ошибок генерируется энтропия, а процесс исправления ошибок выбрасывает ее из вычислительного устройства в служебные регистры. В принципе, вычисление может продолжаться независимо по мере поступления новых служебных кубитов, но на практике мы захотим очистить служебный кубит и использовать его снова. Стирание служебного кубита неизбежно вызовет рассеяние мощности и образование тепла; таким образом, потребуется охлаждение устройства.
- **Никаких ошибок утечки.** Ошибки утечки игнорировались. Но, как отмечено в разделе 8, их учет не существенно влияет на выводы.

Наши предположения были достаточно реалистичными, поэтому можно смело утверждать, что квантовый компьютер с количеством кубитов порядка миллиона и частотой ошибок на вентиль порядка одной на миллион будет мощным и полезным устройством (при условии достаточной скорости обработки данных). С точки зрения текущего состояния технологии (Monroe *et al* 1995; Turchette *et al* 1995; Cory *et al* 1996; Gershenfeld & Chuang 1997), эти числа выглядят пугающе. Но на самом деле даже машина, удовлетворяющая гораздо менее жестким техническим требованиям, все же может быть очень полезна (Preskill 1997). Прежде всего, поми-

мо факторизации, квантовые компьютеры могут выполнять другие задачи, и некоторые из них (в частности, моделирование квантовых систем (Lloyd 1996)) могут быть совершены менее надежным или менее крупным устройством. Более того, по ряду причин наша оценка порога безошибочности может оказаться слишком консервативной. Например, она была получена при допущении, что фазовые и амплитудные ошибки в кубитах одинаково вероятны. Располагая более реалистичной моделью, лучше описывающей вероятности ошибок в действующем устройстве, можно было бы осуществить более эффективную схему исправления ошибок, позволяющую пережить более высокую частоту их появления. Кроме того, даже в сформулированных предположениях приведенный выше анализ отказоустойчивой схемы не вполне точен; при более тонком анализе можно ожидать несколько более высокого порога безошибочности, возможно, даже значительно более высокого. Существенного улучшения можно добиться и путем модификации отказоустойчивых схем либо в результате обнаружения более эффективного способа реализации универсального набора отказоустойчивых вентилей, либо в результате открытия более эффективных методов выполнения измерения синдрома ошибки. Я не буду удивлен, если окажется, что квантовый компьютер со всевозможными усовершенствованиями может работать эффективно с вероятностью ошибки на вентиль, скажем, порядка 10^{-4} (число 10^{-4} вполне может оказаться *ниже* порога безошибочности; более оптимистичные оценки порога безошибочности были выдвинуты Залкой (Zalka 1996)).

В любом случае, принимая эти оценки в качестве номинальных, мы получаем приблизительную цель, к которой должны стремиться: 10^6 кубитов с частотой ошибок 10^{-6}. Это звучит достаточно жестко, но, конечно, могло быть и хуже. Если бы мы пришли к выводу, что нам необходима частота ошибок, скажем, порядка 10^{-20}, тогда будущие перспективы квантовых вычислений были бы действительно неясными. Частота ошибок 10^{-6} несомненно претенциозна, но, возможно, не находится за рамками того, что может быть достигнуто в будущем. В любом случае, сейчас мы имеем четкое представление о том, насколько хорошей должна быть работа квантового компьютера. И это по сути уже является небывалым прогрессом по сравнению с прошлым годом.

Эта работа выполнена при частичной поддержке Департамента Энергетики, грант DE-FG03-92-ER40701, а также Управления перспективного планирования оборонных научно-исследовательских работ (DARPA), грант DAAH04-96-1-0386 управляемый Военным исследовательским центром. Я признателен Дэвиду ДиВинченцо и Войцеху Зуреку за организа-

цию стимулирующих встреч, я благодарю Эндрю Стина и Кристофа Залку за полезные замечания к рукописи этой статьи. Я также хочу поблагодарить своих сотрудников Дэвида Бэкмана, Джара Эвслина, Шэма Какадэ и особенно Дэниела Готтесмана за многочисленные продуктивные дискуссии по проблемам отказоустойчивых вычислений.

Литература

Aharonov D. & Ben-Or M. 1996a Fault tolerant quantum computation with constant error. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing* ACM. New York, 1998, p. 176. (Online preprint quant-ph/9611025.)

Beckman D., Chari A., Devabhaktuni S. & Preskill J. 1996 Efficient networks for quantum factoring. *Phys. Rev.* A **54**, 1034–1063.

Bennett C., DiVincenzo D., Smolin J. & Wootters W. 1996 Mixed state entanglement and quantum error correction. *Phys. Rev.* A **54**, 3824–3851.

Calderbank A. R. & Shor P. W. 1996 Good quantum error-correcting codes exist. *Phys. Rev.* A **54**, 1098–1105.

Calderbank A. R., Rains E. M., Shor P. W. & Sloane N. J. A. 1996 Quantum error correction via codes over GF(4). IEEE Trans. Inf. Theory, 44(4), pp. 1369–1387 (Online preprint quant-ph/9608006.)

Calderbank A. R., Rains E. M., Shor P. W. & Sloane N. J. A. 1997 Quantum error correction and orthogonal geometry. *Phys. Rev. Lett.* **78**, 405–408.

Cirac J. I. & Zoller P. 1995 Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091–4094.

Cory D. G., Fahmy A. F. & Havel T. F. 1996 Nuclear magnetic resonance spectroscopy: an experimentally accessible paradigm for quantum computing. In *Proceedings of the 4th Workshop on Physics and Computation*, T. Toffoli, M. Biafore, and J. Leao (eds.), Boston: New England Complex Systems Institute, pp. 87–91.

Dieks D. 1982 Communication by electron-paramagnetic-resonance devices. *Phys. Lett.* A **92**, 271–272.

DiVincenzo D. & Shor P. 1996 Fault-tolerant error correction with efficient quantum codes. *Phys. Rev. Lett.* **77**, 3260–3263.

Gershenfeld N. & Chuang I. 1997 Bulk spin resonance quantum computation. *Science* **275**, 350–356.

Griffiths R. B. & Niu C. 1996 Semiclassical Fourier transform for quantum computation. *Phys. Rev. Lett.* **76**, 3228–3231.

Gottesman D. 1996 Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev.* A **54**, 1862–1868.

Gottesman D. 1997a A theory of fault-tolerant quantum computation. *Phys. Rev.* A **57**, 127–137 (1998). (Online preprint quant-ph/9702029.)

Gottesman D. 1997b Stabilizer codes and quantum error correction. Ph. D. thesis, California Institute of Technology. (Online preprint quant-ph/9705052.)

Gottesman D., Evslin J. Kakade S. & Preskill J. 1996, to be published.

Grassl M., Beth Th. & Pellizzari T. 1996 Codes for the quantum erasure channel. *Phys. Rev.* A **56**, 33–38 (1997). (Online preprint quant-ph/9610042.)

Hughes R. 1997 (private communication).

Kitaev A. Yu. 1996a Quantum error correction with imperfect gates, in *Proceedings of the Third International Conference on Quantum Communication and Measurement*, Ed O. Hirota, A/S/ Holevo, and C.M. Caves, pp. 181–188 (New York, Plenum, 1997).

Kitaev A. Yu. 1996b. Китаев А.Ю. Квантовые вычисления: алгоритмы и исправления ошибок, Успехи мат. наук, **52**, стр. 53–112 (1997).

Knill E. & Laflamme R. 1996 Concatenated quantum codes. Techn. Report LAOR-96-2808. (Online preprint quant-ph/9608012.)

Knill E. & Laflamme R. 1997. A theory of quantum error-correcting codes. *Phys. Rev.* A **55**, 900–911.

Knill E., Laflamme R. & Zurek W. 1996 Accuracy threshold for quantum computation. (Online preprint quant-ph/9610011.)

Knill E., Laflamme R. & Zurek W. 1997 Resilient quantum computation: error models and thresholds. *Proc. Roy. Soc. Lond.* A **454**, 365–384 (1998) (Online preprint quant-ph/9702058.)

Laflamme R., Miquel C., Paz J. P., & Zurek W. 1996 Perfect quantum error correction code. *Phys. Rev. Lett.* **77**, 198–201.

Landauer R. 1995 Is quantum mechanics useful? *Phil. Tran. R. Soc. Lond.* **353**, 367–376.

Landauer R. 1996 The physical nature of information. *Phys. Lett.* A **217**, 188–193.

Landauer R. 1997 Is quantum mechanically coherent computation useful? In *Proc. Drexel-4 Symposium on Quantum Nonintegrability-Quantum-Classical Correspondence*, Philadelphia, PA, 8 September 1994 (ed. D. H. Feng and B.-L. Hu), Boston: International Press.

Lenstra A. K., Cowie J., Elkenbracht-Huizing M., Furmanski W., Montgome-

ry P. L., Weber D. & Zayer J. 1996 RSA factoring-by-web: the world-wide status. (Online document http://www.npac.syr.edu/factoring/status.html.)

Lloyd S. 1996 Universal quantum simulators. *Science* **273**, 1073–1078; correction in *Science* **279**, 1113–1117 (1998).

Lloyd, S. 1997 The capacity of a noisy quantum channel. *Phys. Rev.* A **55**, 1613–1622.

MacWilliams F. J. & Sloane N. J. A. 1977 *The Theory of Error-Correcting Codes*. New York: North-Holland Publishing Company. Русский перевод: Ф.Дж. Мак-Вильямс, Н.Дж. Слоэн, *Теория кодов, исправляющих ошибки*, Связь, М., 1979.

Miquel C., Paz J. P. & Zurek W. H. 1997 Quantum computation with phase drift errors, *Phys. Rev. Lett.*, **78**, 3971–3974 (1997); (Online preprint quant-ph/9704003.)

Monroe C., Meekhof D. M., King B. E., Itano W. M. & Wineland D. J. 1995 Demonstration of a fundamental quantum logic gate. *Phys. Rev. Lett.* **75**, 4714–4717.

Obenland K. & Despain A. M. 1996a Simulation of factoring on a quantum computer architecture. In *Proceedings of the 4th Workshop on Physics and Computation*, Boston, November 22–24, 1996, Boston: New England Complex Systems Institute.

Obenland K. & Despain A. M. 1996b Impact of errors on a quantum computer architecture. Technical Report, Information Science Institute, University of Southhern California, Oct 1, 1996;
(online preprint http:// www.isi.eu/acal/quantum/quantumoperrors.ps, 1996).

Plenio M. B. & Knight P. L. 1996 Decoherence limits to quantum computation using trapped ions. *Proc. Roy. Soc. Lond.* A **453**, 2017–2041 (1997) (Online preprint quant-ph/9610015.)

Preskill J. 1997 Quantum computing: pro and con. *Proc. Roy. Soc. Lond.* A **454**, 469–486 (1998) (Online preprint quant-ph/9705032.); перевод в *Квантовые вычисления: за и против*. — РХД, Ижевск (1999).

Shor P. 1994 Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Fundamentals of Computer Science*. Los Alamitos, CA: IEEE Press, pp. 124–134. Расширенная версия: SIAM J. Comp. **26**, 1484–1509 (1997), online preprint quant-ph/9508027; перевод: П. Шор, Полиномиальные по времени алгоритмы разложения числа на простые множители и нахождения дискретного логарифма для квантовых компьютеров. *Квантовый компьютер и квантовые вычисления*. — Ижевск, РХД (1999).

Shor P. 1995 Scheme for reducing decoherence in quantum memory. *Phys. Rev.* A **52**, R2493–R2496.

Shor P. 1996 Fault-tolerant quantum computation. In *Proceedings of 37 Annual Symposium on the Foundations of Computer Science.* pp. 56–65, Los Alamitos, CA: IEEE Press (Online preprint quant-ph/9605011).

Shor P. 1997, these proceedings.

Shor P. & and Smolin J. 1996 Quantum error-correcting codes need not completely reveal the error syndrome. (Online preprint quant-ph/9604006.)

Steane A. M. 1996a Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77**, 793–797.

Steane A. M. 1996b Multiparticle interference and quantum error correction. *Proc. Roy. Soc. Lond.* A **452**, 2551–2577.

Steane A. M. 1997 Active stabilization, quantum computation and quantum state synthesis. *Phys. Rev. Lett.* **78**, 2252–2255.

Turchette Q. A., Hood C. J., Lange W., Mabuchi H. & Kimble H. J. 1995 Measurement of conditional phase shifts for quantum logic. *Phys. Rev. Lett.* **75**, 4710–4713 (1995).

Vedral V., Barenco A. & Ekert A. 1996 Quantum networks for elementary arithmetic operations. *Phys. Rev.* A **54**, 147–153

Unruh W. G. 1995 Maintaining coherence in quantum computers. *Phys. Rev.* A **51**, 992–997.

Wootters W. K. & Zurek W. H. 1982 A single quantum cannot be cloned. *Nature* **299**, 802–803; *Nature* **304**, 188-189 (1983).

Zalka C. 1996 Threshold estimate for fault tolerant quantum computing. (Online preprint quantph/9612028.)

Отказоустойчивые квантовые вычисления[1]

Джон Прескилл[2]

Открытие *коррекции квантовых ошибок* существенно улучшило долгосрочные перспективы технологии квантовых вычислений. Закодированную квантовую информацию можно защитить от ошибок, возникающих вследствие неконтролируемых взаимодействий с окружением или неидельного выполнения квантовых логических операций. Исправление ошибок может быть эффективным даже при возникновении случайных ошибок в ходе самой процедуры восстановления. Более того, закодированную квантовую информацию можно обрабатывать без серьезного распространения ошибок. В принципе, сколь угодно продолжительное квантовое вычисление может быть надежно выполнено при условии, что средняя вероятность появления ошибки на квантовый вентиль меньше определенной критической величины, называемой *порогом безошибочности*. Возможна разработка средств аппаратного обеспечения квантовых вычислений, обладающих собственной внутренней отказоустойчивостью, с привлечением топологических взаимодействий Аронова–Бома для обработки квантовой информации.

1. Потребность в отказоустойчивости

Квантовые компьютеры кажутся способными, во всяком случае, в принципе, решать определенные задачи намного быстрее, чем любой мыслимый классический компьютер [1–3]. Однако на практике технология квантовых вычислений все еще находится в зачаточном состоянии. Несмотря на то, что практичный и полезный квантовый компьютер, возможно, будет сконструирован уже в обозримом будущем, мы до сих пор не можем ясно представить, как будут выглядеть аппаратные средства этой машины. Тем не менее, мы можем быть абсолютно уверены в том, что любой практичный квантовый компьютер будет включать в свою работу определенный тип

[1] J. Preskill, Fault-tolerant Quantum Computation, *Introduction to Quantum Computation*, H.-K. Lo, S. Popesku, T. P. Spiller (Eds.), World Scientific, Singapore et al. (1998).

[2]Калифорнийский технологический институт, Пасадена, CA 91125, США.

исправления ошибок. По сравнению с обычными цифровыми, квантовые компьютеры гораздо более подвержены ошибкам, поэтому для предотвращения аварийных отказов в их работе потребуются определенные методы контроля и исправления этих ошибок.

Декогерентизация — самый грозный враг квантового компьютера [4–8]. Мы знаем, как приготовить квантовое кот-состояние, то есть суперпозицию живого и мертвого кота, но мы никогда не сможем наблюдать такие макроскопические суперпозиции вследствие их крайней нестабильности. Ни один реально существующий кот не может быть изолирован от окружающей его среды: окружение «измеряет» кота, незамедлительно проецируя его на состояние — совершенно живой или совершенно мертвый [9]. Квантовый компьютер, возможно, не так сложен, как кот, но это сложная квантовая система, и, подобно коту, она взаимодействует с окружением. Хранящаяся в компьютере информация разрушается, что ведет к ошибкам и сбою в вычислениях. Можем ли мы защитить квантовый компьютер от подрывающего влияния декогерентизации?

Декогерентизация — не единственный наш враг [4–6]. Даже если бы мы были способны добиться превосходной изоляции нашего компьютера от окружающей среды, мы не могли бы рассчитывать на безупречно точное выполнение квантовых логических вентилей. Как и в классическом аналоговом компьютере, ошибки в квантовых вентилях образуют континуум. В ходе вычисления мелкие ошибки в вентилях могут накапливаться и, в конечном счете, вызвать сбой, а способ их исправления неочевиден. Можем ли мы предотвратить катастрофическое накопление этих малых ошибок?

Перспективы на будущее квантовых вычислений получили огромную поддержку благодаря открытию, что коррекция квантовых ошибок, в принципе, действительно возможна [10–12]. Но самого по себе этого открытия недостаточно, чтобы гарантировать надежную работу шумящего квантового компьютера. Чтобы выполнить протокол квантовой коррекции ошибок, мы должны сначала закодировать квантовую информацию, которую хотим защитить, а затем многократно выполнить операции по исправлению, которые обращают накапливаемые ошибки. Но кодирование и исправление сами по себе являются сложными квантовыми вычислениями, и в ходе их выполнения неизбежно будут возникать новые ошибки. Следовательно, чтобы достичь высокой надежности даже при появлении некоторых ошибок на этапе восстановления, нам необходимо найти достаточно сильные методы их исправления.

Более того, чтобы оперировать квантовым компьютером, мы должны уметь больше, чем просто *хранить* квантовую информацию; мы должны уметь *обрабатывать* ее. Мы должны быть в состоянии выполнять кван-

товые вентили, в которых два или более закодированных кубитов взаимодействуют друг с другом. Если в одном кубите возникает ошибка, то существует вероятность ее распространения на другие кубиты, с которыми зараженный кубит будет взаимодействовать в ходе выполняемых вслед за этим квантовых вентилей. Мы должны сконструировать наши вентили таким образом, чтобы минимизировать распространение ошибок.

Включение коррекции квантовых ошибок несомненно усложнит работу квантового компьютера. Создание необходимого для защиты от ошибок резерва потребует увеличить количество элементарных кубитов. Применение вентилей к закодированной информации и периодическое включение этапов исправления ошибок замедлит процесс вычисления. Ввиду этого необходимого усложнения устройства *a priori* не очевидно, что коррекция ошибок приведет к улучшению его работы.

Устройство, работающее эффективно, даже если его элементарные компоненты неидеальны, называют *отказоустойчивым*. Настоящий раздел посвящается теории отказоустойчивых квантовых вычислений. В ней мы исследуем сформулированные выше проблемы и вопросы.

В самом деле, подобные вопросы возникают и в теории отказоустойчивых *классических* вычислений. Но поскольку существующая, базирующаяся на кремнии, полупроводниковая схемотехника в высшей степени надежна, отказоустойчивость не является важнейшим условием работоспособности современных цифровых компьютеров. Несмотря на это, изучение отказоустойчивых классических вычислений имеет выдающуюся историю. В 1952 году фон Нейман [13] предложил повысить надежность работы схемы с шумящими вентилями путем многократного выполнения каждого вентиля и применения мажоритарной схемы. Он сделал вывод, что если отказы вентилей статистически независимы, а вероятность отказа в расчете на один вентиль достаточно мала, то любое вычисление можно выполнить с достаточной надежностью. Единственным недостатком анализа фон Неймана было то, что он предположил идеальную передачу битов по соединяющим вентили «проводникам».[1] Выход за рамки этого предположения оказался сложным, но, в конечном счете, успех был достигнут в 1983 году Гаксом [14], который описал универсальный клеточный автомат с иерархической структурой, которая может поддерживаться локальными операциями в присутствии шума, не нуждаясь в непосредственной нелокальной связи между его компонентами.

[1]Это достаточно серьезная проблема, ибо схема фон Неймана не может быть реализована в трехмерном пространстве с ограниченными по длине проводниками; предполагается, что вероятность появления ошибки передачи стремится к единице при стремящейся к бесконечности длине линий передачи.

Интересен вопрос: способна ли квантовая система подобным образом обеспечивать функционирование сложной иерархической структуры? Однако мы не будем столь амбициозны, чтобы браться здесь за эту проблему. Поскольку нас интересуют ограничения, накладываемые шумом на обработку *квантовой* информации, мы разделим наши вентили на классические и квантовые и будем считать, что классические вентили могут выполняться с идеальной точностью и так быстро, как это необходимо.[1] Это предположение будет хорошо обоснованным до тех пор, пока тактовая частота и точность классического компьютера значительно превосходят соответствующие характеристики квантового компьютера.

Обсудив во втором разделе свойства конкретного корректирующего кода (семикубитовый код Стина [12]), в следующем разделе мы сделаем краткий обзор основных принципов отказоустойчивого исправления квантовых ошибок. Ошибки, возникающие в ходе самого процесса исправления, впоследствии могут повредить закодированную квантовую информацию; следовательно, эффективность коррекции требует особой тщательности ее выполнения. Для измерения *синдромов*, диагностирующих ошибки в блоках закодированных данных, используются служебные кубиты. Последние могут содержать ошибки, поэтому необходимо минимизировать вероятность их распространения на закодированную информацию. В третьем разделе описываются предложенные Питером Шором [13] и Эндрю Стином [16] методы управления распространением ошибок в процессе исправления.

Предмет четвертого раздела составляет отказоустойчивая обработка квантовой информации. Здесь главная проблема заключается в создании универсального набора квантовых вентилей, способных действовать на блоки закодированных данных, не внося непомерного количества ошибок. В этом разделе сделан обзор некоторых схем универсальных вычислений (Питер Шор [15] и Дэниел Готтесман [17]).

Коль скоро элементарные вентили нашего квантового компьютера достаточно надежны, мы можем применять их к закодированной информации, а также в процессах отказоустойчивого исправления ошибок, повышая таким образом надежность нашего устройства. Но для любого заданного квантового кода или даже большинства бесконечных классов, содержащих коды с блоками сколь угодно больших размеров, при выполнении достаточно продолжительного вычисления рано или поздно эти процедуры дадут сбой. Однако, как показано в разделе 5, существует особый класс кодов (*каскадные коды*), позволяющих надежно выполнять все более и бо-

[1] Однако, когда мы будем рассматривать коррекцию ошибок с помощью кодов со сколь угодно большими размерами блоков, мы будем требовать, чтобы объем выполняемых классических операций был полиномиально ограничен по размеру блоков.

лее продолжительные квантовые вычисления, требуя для этого умеренной скорости увеличения размера блока [18–24]. Применяя каскадные коды, мы можем установить *порог безошибочности* квантовых вычислений; как только наше «железо» будет удовлетворять установленным критерием точности, корректирующие ошибки квантовые коды и отказоустойчивые процедуры позволят осуществлять сколь угодно длинные квантовые вычисления со сколь угодно высокой надежностью. Этот результат приблизительно аналогичен выводу фон Неймана относительно классической отказоустойчивости, тогда как иерархическая структура каскадного кодирования напоминает конструкцию Гакса. Приводится набросок оценки порога безошибочности в предположениях о характере ошибок, перечисленных в разделе 6.

С развитием отказоустойчивых методов стало ясно, что, активно вмешиваясь в работу квантового компьютера, оператор, в принципе, может защитить его от ошибок в шумящем (но не *слишком сильно*) окружении. Хотя, в более далекой перспективе, надежность практичного квантового компьютера может быть достигнута совершенно иным способом — с помощью внутренне отказоустойчивого аппаратного обеспечения. Такое «железо», устойчивое к локальным возмущениям, могло бы оперировать сравнительно небрежно организованными процедурами и, тем не менее, надежно хранить и обрабатывать квантовую информацию. Предмет раздела 7 составляет предложенный Алексеем Китаевым [25] проект отказоустойчивых аппаратных средств, в которых квантовые вентили используют неабелевы взаимодействия Аароно́ва–Бо́ма между пространственно разделенными квазичастицами в гипотетически возможной двумерной спиновой системе. Хотя лабораторная реализация идеи Китаева, возможно, является делом отдаленного будущего, его работа открывает новый подход к проблеме квантовой отказоустойчивости, позволяющий отказаться от анализа абстрактных квантовых схем в пользу поиска новых физических принципов, которые можно было бы использовать для надежной обработки квантовой информации.

С точки зрения современной технологии, сделанные в этой главе заявления относительно потенциально возможного отказоустойчивого управления сложными квантовыми состояниями могут показаться претенциозными. Конечно, нам предстоит еще долгий путь, прежде чем будут созданы устройства, которые смогут, скажем, использовать порог безошибочности квантовых вычислений. Тем не менее, я уверен в том, что сегодняшняя работа по коррекции квантовых ошибок будет иметь продолжение. Последние три года теория квантовых вычислений развивалась с захватывающей скоростью. Если квантовая классификация вычислительной сложности отличается от классической (что очень похоже на правду), то ни один мыслимый классический компьютер не сможет точно предсказать поведение

даже умеренного числа кубитов (порядка 100). Тогда, возможно, относительно маленькие квантовые системы будут иметь намного больший потенциал, чем тот, что мы можем сегодня представить. Этот потенциал до сих пор не реализован, так как пока мы не в состоянии защитить такие системы от разрушительного влияния шума и декогерентизации. Таким образом, открытие отказоустойчивых методов коррекции квантовых ошибок и квантовых вычислений имеет исключительно глубокие последствия, как для будущего экспериментальной физики, так и для будущего технологии. Теоретические достижения осветили дорогу в будущее, в котором сложные квантовые системы, вероятно, окажутся более покладистыми.

2. Коррекция квантовых ошибок: 7-кубитовый код

Чтобы понять, почему возможна коррекция квантовых ошибок, очень поучительно изучить конкретный код. Простым и важным примером кода, исправляющего квантовые ошибки, является 7-кубитовый код, разработанный Эндрю Стином [11,12]. Этот код позволяет хранить один кубит квантовой информации (произвольное состояние в двумерном гильбертовом пространстве), используя всего семь кубитов (погружая двумерное гильбертово пространство в пространство размерности 2^7). Фактически код Стина тесно связан с известным классическим корректирующим ошибки [7,4,3]-кодом Хэмминга [26]. Чтобы понять принцип работы кода Стина, важно сначала понять классический код Хэмминга.

Код Хэмминга использует блок из семи битов для кодирования четырех битов классической информации: имеется $16 = 2^4$ строк длины семь, представляющих собой истинные кодовые слова. Кодовые слова можно характеризовать матрицей проверки четности:

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}. \tag{1}$$

Каждое истинное кодовое слово представляет собой 7-битовый вектор (-столбец) v_{code}, удовлетворяющий уравнению

$$\sum_k H_{jk}(v_{\text{code}})_k = 0 \pmod 2; \tag{2}$$

то есть в арифметике по модулю 2 матрица H уничтожает каждое кодовое слово. Поскольку $Z_2 = \{0, 1\}$ является (конечным) полем, в нем применимы результаты линейной алгебры. H имеет три линейно независимых стро-

ки, а ее ядро[1] натянуто на четыре линейно независимых вектора-столбца. 16 истинных кодовых слов представляют собой все возможные линейные комбинации этих четырех векторов с коэффициентами, выбираемыми из $\{0, 1\}$.

Теперь предположим, что v_{code} — (неизвестное) истинное кодовое слово, в котором возникла единственная (неизвестная) ошибка: один из семи битов инвертировался. Наша задача — определить поврежденный бит и исправить ошибку. Этот трюк можно выполнить с помощью матрицы проверки четности. Пусть e_i обозначает вектор с единицей в i-ой позиции и нулями в остальных. Тогда при инвертировании i-го бита v_{code} становится равным $v_{\text{code}} + e_i$. Если мы подействуем на этот вектор матрицей H, то получим

$$H(v_{\text{code}} + e_i) = He_i \tag{3}$$

(поскольку матрица H уничтожает кодовое слово v_{code}), что является именно i-м столбцом матрицы H. Поскольку все столбцы матрицы H различны, то этим однозначно определяется значение i. Выяснив местоположение ошибки, ее можно исправить путем повторного инвертирования i-го бита. Таким образом, если инвертируется только один бит, то мы можем однозначно восстановить закодированную информацию; но при инвертировании двух или большего числа разных битов закодированные данные будут повреждены. Примечательно то, что величина He_i выявляет позицию ошибки, ничего не сообщая о кодовом слове v_{code}, то есть не раскрывая закодированную информацию.

Код Стина обобщает этот вид классического кода коррекции ошибок на *квантовый* код, корректирующий ошибки. Он использует 7-кубитовый «блок» для кодирования одного кубита квантовой информации, то есть произвольного состояния в двумерном гильбертовом пространстве, натянутом на два состояния: «логический нуль» $|0\rangle_{\text{code}}$ и «логическая единица» $|1\rangle_{\text{code}}$. Код сконструирован так, что он позволяет исправить одну произвольную ошибку, возникающую в любом из семи кубитов в блоке.

Что мы подразумеваем под произвольной ошибкой? Рассматриваемый кубит может подвергнуться случайному *унитарному* преобразованию или *потерять когерентность*, запутавшись с состояниями окружающей среды. Предположим, что неповрежденный кубит должен находиться в состоянии $a|0\rangle + b|1\rangle$. (Конечно, он может быть запутан с остальными, так что коэффициенты a и b не обязательно являются комплексными числами; это могут быть состояния, одновременно ортогональные $|0\rangle$ и $|1\rangle$, которые, как мы

[1] Ядро матрицы или ее *нуль-пространство* — пространство векторов, удовлетворяющих уравнению (2). — *Прим. ред.*

предполагаем, не подвержены ошибкам.) Теперь, если этот кубит поврежден произвольной ошибкой, результирующее состояние можно разложить следующим образом:

$$\begin{aligned} a|0\rangle + b|1\rangle \rightarrow & (a|0\rangle + b|1\rangle) \otimes |A_{\text{no error}}\rangle_{\text{env}} + \\ & + (a|1\rangle + b|0\rangle) \otimes |A_{\text{bit-flip}}\rangle_{\text{env}} + \\ & + (a|0\rangle - b|1\rangle) \otimes |A_{\text{phase-flip}}\rangle_{\text{env}} + \\ & + (a|1\rangle - b|0\rangle) \otimes |A_{\text{both error}}\rangle_{\text{env}}, \end{aligned} \tag{4}$$

где каждое $|A\rangle_{\text{env}}$ обозначает состояние окружающей среды. Мы не делаем каких-то определенных предположений относительно ортогональности или нормировки состояний $|A\rangle_{\text{env}}$,[1] поэтому уравнение (4) не влечет за собой потери общности. Мы приходим к выводу, что эволюцию кубита можно выразить как линейную комбинацию четырех возможностей: (1) не возникает никакой ошибки, (2) возникает инвертирование бита $|0\rangle \longleftrightarrow |1\rangle$, (3) возникает обращение относительной фазы $|0\rangle$ и $|1\rangle$, (4) происходит одновременное инвертирование бита и обращение фазы.

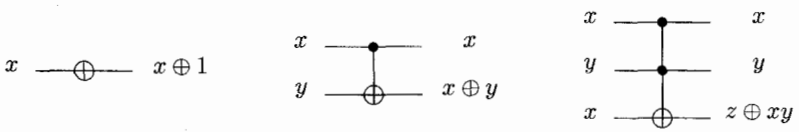


Рис. 1. Схематическое изображение NOT-вентиля, XOR-вентиля (контролируемое NOT) и вентиля Тоффоли (дважды контролируемое NOT)

Теперь понятно, как должен работать квантовый код, корректирующий ошибки [12, 27]. Выполняя подходящее измерение, мы хотим диагностировать, какая из этих четырех возможностей фактически осуществилась. Конечно, в общем случае состоянием кубита будет линейная комбинация этих четырех состояний, но измерение должно спроецировать его на базис, используемый в уравнении (4). После этого мы можем приступить к исправлению ошибки с помощью одного из четырех унитарных преобразований:

$$(1)\quad I, \quad (2)\quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (3)\quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (4)\quad Y = X \cdot Z \tag{5}$$

(какое из них применить, указывает результат измерения). Применив это преобразование, мы возвращаем кубит в его исходное состояние и полностью распутываем квантовые состояния кубита и окружающей среды.

[1] Хотя, конечно, совместная эволюция кубита и окружения должна быть унитарной.

Существенно, что при диагностике ошибки мы ничего не узнаем о закодированной квантовой информации, поскольку получение любой информации о коэффициентах a и b в уравнении (4) неизбежно приведет к разрушению когерентности кубита.

Если мы используем код Стина, отвечающее этим критериям измерение возможно. Логический нуль является равновзвешенной суперпозицией всех кодовых слов кода Хэмминга (H) с четным весом (слов с четным количеством единиц),

$$\begin{aligned}|0\rangle_{\text{code}} &= \frac{1}{\sqrt{8}} \sum_{v_{\text{even}} \in \text{H}} |v_{\text{even}}\rangle = \\ &= \frac{1}{\sqrt{8}} (|0000000\rangle + |0001111\rangle + |0110011\rangle + |0111100\rangle + \\ &+ |1010101\rangle + |1011010\rangle + |1100110\rangle + |1101001\rangle),\end{aligned} \tag{6}$$

а логическая единица является равновзвешенной суперпозицией всех кодовых слов кода Хэмминга с нечетным весом (слов с нечетным количеством единиц),

$$\begin{aligned}|1\rangle_{\text{code}} &= \frac{1}{\sqrt{8}} \sum_{v_{\text{even}} \in \text{H}} |v_{\text{even}}\rangle = \\ &= \frac{1}{\sqrt{8}} (|1111111\rangle + |1110000\rangle + |1001100\rangle + |1000011\rangle + \\ &+ |0101010\rangle + |0100101\rangle + |0011001\rangle + |0010110\rangle).\end{aligned} \tag{7}$$

Поскольку все фигурирующие в уравнениях (6) и (7) состояния являются кодовыми словами Хэмминга, инвертирование единственного бита в блоке несложно обнаружить, выполняя простое квантовое вычисление, как это показано на рисунке 2 (используются приведенные на рис. 1 обозначения). Мы расширяем блок из семи кубитов тремя служебными кубитами[1] и выполняем унитарную операцию:

$$|v\rangle \otimes |0\rangle_{\text{anc}} \rightarrow |v\rangle \otimes |Hv\rangle_{\text{anc}}, \tag{8}$$

где H — матрица проверки четности Хэмминга, а $|\cdot\rangle_{\text{anc}}$ обозначает состояние трех служебных битов. Если ошибка содержится лишь в одном из семи кубитов, то измерение служебного кубита спроецирует его либо на инвертированное состояние, либо на исходное неповрежденное (но не на любую

[1] Чтобы сделать процедуру отказоустойчивой, нам потребуется увеличить число служебных кубитов, как это обсуждается в разделе 3.

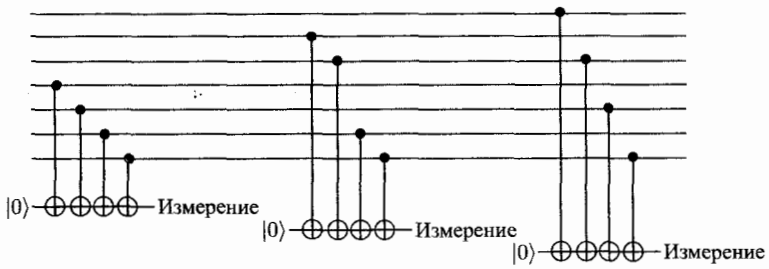


Рис. 2. Вычисление синдрома инвертирования бита для 7-кубитового кода Стина. Повторение вычисления в повернутом базисе диагностирует ошибки обращения фазы. Чтобы сделать процедуру отказоустойчивой, каждый служебный кубит нужно заменить четырьмя кубитами в надлежащем состоянии

нетривиальную суперпозицию этих двух состояний). В первом случае результат измерения диагностирует поврежденный бит, ничего не сообщая о закодированной в блоке квантовой информации.

Но чтобы осуществить коррекцию квантовых ошибок, нам понадобится наряду с ошибками инвертирования бита диагностировать и фазовые ошибки. Для достижения этой цели мы можем (следуя Стину [11, 12]) изменить базис для каждого кубита, применив поворот Адамара:

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{9}$$

Тогда ошибки, бывшие фазовыми в базисе $|0\rangle, |1\rangle$, в повернутом базисе

$$|\tilde{0}\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |\tilde{1}\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{10}$$

становятся ошибками инвертирования бита. Следовательно, достаточно того, что наш код способен диагностировать ошибки инвертирования бита в этом повернутом базисе. Но если мы применяем поворот Адамара к каждому из семи кубитов, логический нуль и логическая единица Стина в повернутом базисе принимают вид

$$\begin{aligned} |\tilde{0}\rangle_{\text{code}} &= \frac{1}{4} \sum_{v \in H} |v\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{\text{code}} + |1\rangle_{\text{code}}), \\ |\tilde{1}\rangle_{\text{code}} &= \frac{1}{4} \sum_{v \in H} (-1)^{\text{wt}(v)} |v\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{\text{code}} - |1\rangle_{\text{code}}) \end{aligned} \tag{11}$$

(где $\mathrm{wt}(v)$ обозначает вес v). Ключевым моментом является то, что $|\tilde{0}\rangle_{\mathrm{code}}$ и $|\tilde{1}\rangle_{\mathrm{code}}$, подобно $|0\rangle_{\mathrm{code}}$ и $|1\rangle_{\mathrm{code}}$ являются суперпозициями кодовых слов Хэмминга. Следовательно, в повернутом базисе, как и в исходном, мы можем выполнить проверку четности Хэмминга для диагностики инвертирования битов, которое в исходном базисе представляло собой обращение фазы. При условии, что поврежден лишь один кубит, выполнение проверки четности в обоих базисах полностью диагностирует ошибку и дает возможность ее исправить.

В описанной выше схеме исправления ошибок я предполагал, что ошибка поражает только один кубит в блоке. Очевидно, что это допущение нереалистично; обычно все кубиты в определенной степени запутываются с окружением. Однако, как мы уже видели, процедура определения синдрома ошибки, как правило, проецирует каждый кубит на исходное неповрежденное состояние. Для каждого кубита появление ошибки происходит с отличной от нуля, но малой по предположению, вероятностью, которую мы будем обозначать ϵ. Сейчас мы сделаем очень важное предположение: ошибки, действующие на разные кубиты в одном блоке, полностью некоррелированы между собой. При таком допущении вероятность появления двух ошибок имеет порядок ϵ^2 и, при достаточно малой величине ϵ, гораздо меньше вероятности возникновения одной ошибки. Поэтому, с точностью порядка ϵ, мы можем уверенно ограничить наше внимание случаем, когда ошибка содержится максимум в одном кубите блока. (На самом деле, для этого вывода не требуется, чтобы действующие на разные кубиты ошибки были *полностью* некоррелированными. Если все кубиты подвергаются воздействию одного и того же слабого магнитного поля, так что вероятность инвертирования каждого из них равна ϵ, все будет в порядке, поскольку вероятность переворота спинов имеет порядок ϵ^2. Проблему вызвал бы возникающий с вероятностью порядка ϵ процесс, при котором происходит инвертирование двух спинов одновременно.)

Но в (маловероятном) случае, когда в одном и том же блоке кода возникает две ошибки, наша процедура исправления обычно дает сбой. Если в одном и том же блоке инвертируются два бита, то проверка четности Хэмминга неправильно диагностирует ошибку. Восстановление вернет квантовое состояние в кодовое подпространство, но в *закодированной* в блоке информации произойдет инвертирование бита

$$|0\rangle_{\mathrm{code}} \rightarrow |1\rangle_{\mathrm{code}}, \quad |1\rangle_{\mathrm{code}} \rightarrow |0\rangle_{\mathrm{code}}. \tag{12}$$

Аналогично, если в одном и том же блоке возникают две фазовые ошибки, то есть две ошибки инвертирования битов в повернутом базисе, то после

восстановления в нем произойдет инвертирование бита в повернутом базисе или обращение фазы в исходном базисе

$$|0\rangle_{\text{code}} \to |0\rangle_{\text{code}}, \quad |1\rangle_{\text{code}} \to -|1\rangle_{\text{code}} \tag{13}$$

(если один кубит в блоке содержит фазовую ошибку, а другой — ошибку инвертирования бита, то исправление будет успешным).

Таким образом, код Стина способен повысить надежность хранения квантовой информации. Предположим, мы хотим сохранить один кубит в неизвестном чистом состоянии $|\psi\rangle$. Из-за несовершенства запоминающего устройства состояние ρ_{out}, которое мы мы восстановим, будет иметь точность воспроизведения

$$F \equiv \langle\psi|\rho_{\text{out}}|\psi\rangle = 1 - \epsilon. \tag{14}$$

Но если мы храним кубит с использованием 7-кубитового блочного кода Стина, если каждый из семи кубитов сохраняется с точностью воспроизведения $F = 1 - \epsilon$, если ошибки кубитов некоррелированны; и, наконец, если мы можем безукоризненно выполнять коррекцию ошибок, кодирование и декодирование (подробнее об этом ниже), то закодированная информация может храниться с улучшенной точностью воспроизведения $F = 1 - O(\epsilon^2)$.

Кубит в неизвестном состоянии можно закодировать, используя изображенную на рисунке 3 схему. Принцип работы кодирующего устройства проще понять, если использовать альтернативное выражение для матрицы проверки четности Хэмминга,

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{15}$$

(Эта форма матрицы H получается из (1) перестановкой столбцов, то есть всего лишь изменением нумерации битов в блоке.) Четный субкод кода Хэмминга фактически является пространством, натянутым на строки матрицы H; поэтому мы видим, что (в этом представлении H) первые три бита строки полностью характеризуют представленные в субкоде данные. Оставшиеся четыре бита — биты четности, обеспечивающие необходимую для защиты от ошибок избыточность. При кодировании неизвестного состояния $a|0\rangle + b|1\rangle$ кодирующее устройство сначала использует два XOR-вентиля, чтобы приготовить состояние $a|0000000\rangle + b|0000111\rangle$ — суперпозицию четного и нечетного кодовых слов Хэмминга. Оставшаяся часть схемы добавляет к этому состоянию $|0\rangle_{\text{code}}$: повороты Адамара (R) готовят равновзвешенную суперпозицию всех восьми возможных значений для

первых трех битов в блоке, а оставшиеся XOR-вентили включают предписанные матрицей H биты проверки четности.

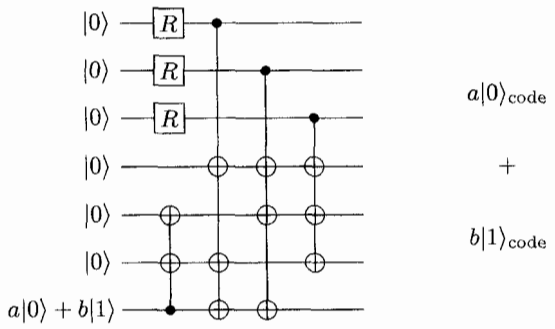


Рис. 3. Кодирующая схема для 7-кубитового кода Стина

Мы также хотим иметь возможность измерять закодированный кубит, скажем, проецируя его на ортогональный базис $\{|0\rangle_{\text{code}}, |1\rangle_{\text{code}}\}$. Если мы не возражаем против разрушения закодированного блока в процессе измерения, то достаточно измерить каждый из семи кубитов в блоке, проецируя на базис $\{|0\rangle, |1\rangle\}$; затем, чтобы получить кодовое слово Хэмминга, мы выполняем классическую коррекцию ошибок в результатах измерения. Четность этого кодового слова является значением логического кубита. (Этап классической коррекции ошибок обеспечивает защиту от ошибок измерения. Например, если блок находится в состоянии $|0\rangle_{\text{code}}$, то при измерении элементарных кубитов для определения логического кубита могут возникнуть две независимые ошибки, в результате чего будет получено неверное значение $|1\rangle_{\text{code}}$.)

В применении к квантовым вычислениям нам потребуется выполнять измерение, проецирующее на базис $\{|0\rangle_{\text{code}}, |1\rangle_{\text{code}}\}$ без разрушения блока. Эта выполняется путем копирования четности блока на служебный кубит с последующим его измерением. Схема, представляющая неразрушающее измерение кодового блока [в случае, когда матрица проверки четности аналогична описанной в уравнении (15)], показана на рисунке 4. Измерение является неразрушающим в том смысле, что оно сохраняет кодовое подпространство; конечно, оно «разрушает» когерентную суперпозицию $a|0\rangle_{\text{code}} + b|1\rangle_{\text{code}}$, коллапсируя это состояние либо к $|0\rangle_{\text{code}}$ (с вероятностью $|a|^2$), либо к $|1\rangle_{\text{code}}$ (с вероятностью $|b|^2$).

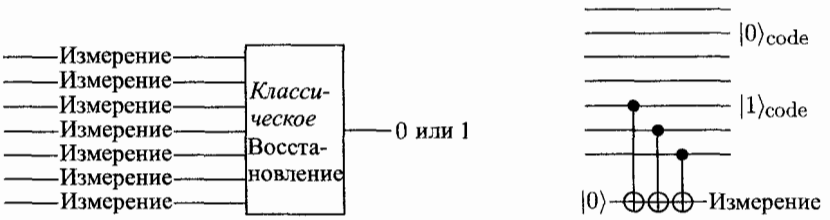


Рис. 4. Разрушающее и неразрушающее измерения логического кубита

7-кубитовый код Стина может исправить только одну ошибку в кодовом блоке, но можно построить лучшие коды [12, 28–31], способные защитить информацию от t ошибок в одном блоке, так что закодированная информация может сохраняться с точностью воспроизведения $F = 1 - O(\epsilon^{t+1})$.

Ключевая идея, сделавшая возможной коррекцию квантовых ошибок, состоит в том, что *с запутыванием можно бороться с помощью запутывания*. Запутывание может быть нашим врагом, поскольку запутывание устройства с окружением может скрыть от нас квантовую информацию и таким образом послужить причиной появления ошибок. Но запутывание может оказаться и нашим союзником — информацию, которую мы хотим защитить, можно закодировать в запутанном состоянии, то есть в корреляциях, охватывающих большое число кубитов. Тогда она остается недоступной при измерении лишь нескольких кубитов. По тем же причинам эта информация не может быть *повреждена*, если окружение взаимодействует лишь с несколькими кубитами.

Более того, мы узнали, что хотя квантовый компьютер в определенном смысле является аналоговым устройством, совершаемые им ошибки можно *оцифровать*. Мы боремся с малыми ошибками, выполняя подходящие измерения, проецирующие состояние квантового компьютера либо на исходное неповрежденное, либо на состояние с большой ошибкой, которую затем можно исправить известными методами. Мы также увидели, что можно измерять ошибки, не измеряя при этом данные, — можно получить точную информацию о природе ошибки, ничего не узнав о закодированной в нашем устройстве квантовой информации (что могло бы повлечь за собой декогерентизацию и сбой в вычислении).

Все квантовые коды коррекции ошибок используют одну и ту же фундаментальную стратегию: маленькое подпространство гильбертова пространства устройства определяется в качестве кодового подпространства. Это пространство тщательно выбирается, так чтобы все ошибки, которые

мы хотим исправить, перемещали его во взаимно ортогональные *подпространства ошибок*. После того, как наша система провзаимодействовала с окружающей средой, можно выполнить измерение, которое сообщит, в каком из этих взаимно ортогональных подпространств оказалась система и, следовательно, однозначно определить тип возникшей ошибки. После этого ошибку можно исправить, применив подходящее унитарное преобразование.

3. Отказоустойчивое исправление

До сих пор в наших обсуждениях предполагалось, что мы можем безошибочно кодировать квантовую информацию и выполнять исправление ошибок. Но, конечно, эти процедуры не будут выполняться безупречно. Восстановление само по себе является предрасположенным к появлению ошибок квантовым вычислением. Если вероятность появления ошибки для каждого бита в кодовом блоке равна ϵ, то естественно предположить, что каждый применяемый в ходе восстановления квантовый вентиль с вероятностью ϵ вызывает появление ошибки (или, что «ошибки запоминающего устройства» возникают в ходе исправления с вероятностью порядка ϵ). Если процедура исправления сконструирована небрежно, то вероятность ее отказа (например, вследствие возникновения двух ошибок в одном блоке) может иметь порядок ϵ. Тогда мы не извлечем пользы от применения квантовых кодов коррекции ошибок; в действительности вероятность появления ошибки на один кубит информации будет даже выше, чем при отсутствии какого-либо кодирования. Поэтому мы должны проанализировать все возможные причины, по которым процедуры исправления могут дать сбой с вероятностью порядка ϵ, и убедиться, что они устранены. Только тогда они будут *отказоустойчивыми*, и только тогда кодирование окупит себя при достаточно малой величине ϵ.

3.1. Проблема обратного действия

Распространение ошибок является серьезной проблемой. Если в одном кубите возникает ошибка, а затем мы применяем вентиль, в котором этот кубит взаимодействует с другим, то существует конечная вероятность распространения ошибки на второй кубит. Следует соблюдать осторожность, чтобы сдержать инфекцию, или, по крайней мере, нужно стремиться предотвратить возникновение двух ошибок в одном блоке.

При выполнении исправления ошибок мы многократно используем двухкубитовый XOR-вентиль. Этот вентиль способен распространять ошибки в двух различных направлениях. Во-первых, очевидно, что если

в одном кубите возникает ошибка инвертирования бита, а затем он используется в качестве источника XOR-вентиля, то инвертирование бита перейдет «в прямом направлении» на кубит цели. Более тонким является второй тип распространения ошибок; его можно понять, используя тождество, представленное на рисунке 5: при повороте Адамара базисов обоих кубитов источник и цель XOR-вентиля меняются ролями. Поскольку мы помним, что эта замена базиса также заменяет ошибку инвертирования бита на фазовую ошибку, мы делаем вывод, что при возникновении в одном кубите фазовой ошибки и последующем использовании его в качестве целевого кубита XOR-вентиля ошибка переходит на кубит источника.

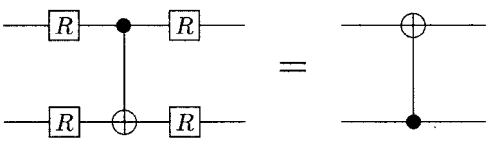


Рис. 5. Полезное тождество. Источник и цель XOR-вентиля меняются местами, если мы с помощью поворота Адамара выполняем замену базисов

Рис. 6. Плохой и хороший варианты измерения синдрома. Плохая схема использует один и тот же служебный бит несколько раз; хорошая схема использует служебный бит лишь однократно

Сейчас мы можем понять, что изображенная на рисунке 2 схема не является отказоустойчивой. Проблема состоит в том, что один и тот же служебный кубит используется в качестве цели четырех последовательных XOR-вентилей. Если на определенной стадии в служебном кубите возникает хотя бы одна фазовая ошибка, эта единственная ошибка может вернуть-

ся обратно к двум или более кубитам кодового блока данных. В результате с вероятностью ϵ в блоке может возникнуть фазовая ошибка, что недопустимо.

Чтобы понизить вероятность отказа до величины порядка ϵ^2, мы должны так изменить схему исправления, чтобы каждый служебный кубит взаимодействовал не более чем с одним кубитом в блоке. Один из способов решения этой задачи состоит в увеличении количества служебных кубитов от одного до четырех, чтобы каждый из них являлся целью одного XOR-вентиля, как это показано на рисунке 6. Затем мы можем измерить все четыре служебных кубита. Бит искомого нами синдрома является битом *четности* четырех измеряемых кубитов. По сути, мы скопировали из блока данных в служебный кубит некоторую информацию о возникшей ошибке, а при измерении служебного кубита мы считываем эту информацию.

Тем не менее, эта процедура по-прежнему неадекватна поставленной цели, поскольку мы скопировали *слишком много* информации. Схема запутывает служебный кубит с возникшей в данных ошибкой, что хорошо, но она запутывает его и с самими данными, а это плохо. Измерение служебного кубита разрушает тщательно приготовленную суперпозицию базисных состояний (6) и (7) для $|0\rangle_{\text{code}}$ и $|1\rangle_{\text{code}}$. Пусть, например, мы измеряем первый бит синдрома, как это показано на рисунке 2, но с увеличенным от одного до четырех количеством служебных кубитов. Следовательно, на самом деле мы измеряем последние четыре бита в блоке. Если в результате измерения мы получаем, скажем, $|0000\rangle_{\text{anc}}$, то это означает, что мы спроецировали $|0\rangle_{\text{code}}$ на $|0000000\rangle$, а $|1\rangle_{\text{code}}$ на $|1110000\rangle$; кодовые слова теряют всякую защиту от фазовых ошибок.

3.2. Приготовление служебного состояния

Продолжим модификацию процедуры исправления, сохраняя ее хорошие черты и исключая плохие. Мы хотим скопировать на служебные кубиты информацию об ошибках, возникающих в блоке данных, не внося в него многочисленных фазовых ошибок и не разрушая когерентности закодированных в нем данных. Для достижения этой цели необходимо, прежде чем приступать к вычислению синдрома ошибки, приготовить подходящее состояние служебных кубитов. Это состояние строится таким образом, чтобы результат его измерения открывал информацию об ошибках, ничего не сообщая о состоянии данных.

Один из способов удовлетворить этому критерию был найден Питером Шором. Он предложил использовать в качестве состояний четырех служебных кубитов равновзвешенную суперпозицию всех кодовых слов четного

веса (*состояние Шора*)

$$|\mathrm{Shor}\rangle_{\mathrm{anc}} = \frac{1}{\sqrt{8}} \sum_{v_{\mathrm{even}}} |v_{\mathrm{even}}\rangle_{\mathrm{anc}}. \tag{16}$$

Чтобы вычислить каждый бит синдрома, мы готовим служебные кубиты в состояниях Шора, выполняем четыре XOR-вентиля (с подходящими кубитами блока данных в качестве источников и четырьмя кубитами в состояниях Шора в качестве целей), а затем измеряем служебное состояние.

Если вычисляемый бит тривиален, то к кодовым словам v в (16) добавляется строка четного веса, что оставляет состояние Шора неизменным; если бит синдрома нетривиален, тогда состояние Шора преобразуется в равновзвешенную суперпозицию кодовых слов с нечетным весом. Таким образом, четность результата измерения выявляет значение бита синдрома, но никакую другую информацию о состоянии блока данных из этого измерения извлечь нельзя — мы нашли способ выделить синдром, не повреждая кодовые слова. (Определяемая нами при измерении отдельная строка с заданной четностью выбирается случайным образом и не имеет ничего общего с состоянием блока данных.)

Всего существует шесть битов синдрома (по три для диагностики ошибок инвертирования бита и для диагностики ошибок обращения фазы), поэтому измерение синдрома использует 24 служебных бита, приготовленных в шести состояниях Шора, и 24 XOR-вентиля.

Одним из способов получения синдрома обращения фазы может быть следующий: сначала применить к блоку данных семь параллельных R-вентилей для поворота базиса, затем, как показано на рисунке 2, применить XOR-вентили (но с приготовленными в состояниях Шора служебными кубитами) и, наконец, применить семь R-вентилей для поворота данных в исходный базис. Однако с целью усовершенствования этой процедуры мы можем использовать представленное на рисунке 5 тождество. Обращая направление XOR-вентилей (то есть используя служебные биты в качестве источников, а данные — в качестве цели), мы можем избежать применения R-вентилей к данным и, следовательно, понизить вероятность их повреждения [16, 24], как это показано на рисунке 7.

Другой способ приготовления служебного кубита был предложен Эндрю Стином. Его 7-кубитовое служебное состояние является равновзвешенной суперпозицией кодовых слов Хэмминга:

$$|\mathrm{Steane}\rangle_{\mathrm{anc}} = \frac{1}{4} \sum_{v \in \mathrm{H}} |v\rangle. \tag{17}$$

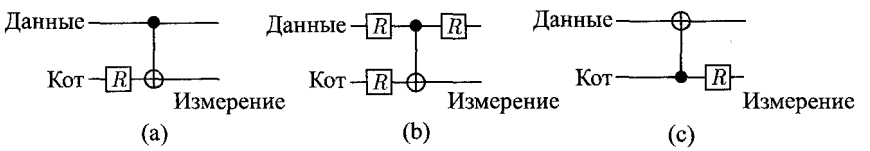


Рис. 7. (a) Схематическое изображение процедуры вычисления одного бита синдрома ошибки инвертирования бита. Применяемый к «кот-состоянию» вентиль Адамара завершает приготовление состояния Шора (см. раздел 3.3). Как XOR-вентиль, так и вентиль Адамара на диаграмме фактически представляют собой четыре параллельно выполняемых вентиля. (b) Схематическое изображение процедуры вычисления одного бита синдрома ошибки обращения фазы. Она аналогична (a), но применяется к данным в базисе, повернутом преобразованием Адамара. (c) Эквивалентная (b) схема, упрощенная с помощью тождества, изображенного на рисунке 5

(Это состояние можно также записать в виде $(|0\rangle_{\text{code}} + |1\rangle_{\text{code}})/\sqrt{2}$; его можно получить, применив к состоянию $|0\rangle_{\text{code}}$ побитовый поворот Адамара.) Чтобы вычислить синдром инвертирования бита, мы с помощью XOR-вентиля обращаем каждый кубит блока данных в соответствующий кубит служебного состояния, а затем измеряем его. Применяя к результату *классического* измерения матрицу проверки четности Хэмминга H, мы выделяем синдром инвертирования бита. Как и в методе Шора, эта процедура «копирует» данные в служебный кубит, состояние которого выбрано таким образом, чтобы при его измерении обеспечить прочтение информации только об ошибке. Например, если ошибка отсутствует, найденная в измерении конкретная строка является случайно выбранным кодовым словом Хэмминга и ничего не сообщает о состоянии данных. При определении синдрома обращения фазы аналогичная процедура выполняется в повернутом базисе. Преимущество метода Стина перед процедурой Шора состоит в том, что он требует лишь 14 служебных битов и 14 XOR-вентилей. Но он также имеет и недостаток: приготовление состояния Стина сложнее, чем приготовление состояния Шора, так что оно отчасти более подвержено возникновению ошибок.

3.3. Проверка служебного состояния

Продолжая изучать все ситуации, в которых сбой процесса восстановления может произойти из-за единственной ошибки, мы обнаруживаем другую потенциальную проблему. Вследствие распространения ошибок, единственная ошибка, возникающая в процессе приготовления состояния Шора

или состояния Стина, может послужить причиной появления в них двух фазовых ошибок, а они, в свою очередь, могут распространиться на закодированные данные, если для измерения синдрома будет использован дефектный служебный кубит. Наша процедура по-прежнему не является отказоустойчивой.

Следовательно, перед его использованием состояние служебных кубитов должно быть протестировано на многократные фазовые ошибки. Если оно не выдерживает тест, оно должно быть забраковано, а служебное состояние построено заново.

Один из способов приготовления и проверки состояния Шора изображен на рисунке 8. Первый вентиль Адамара и первые три XOR-вентиля в этой схеме готовят «кот-состояние» ($|0000\rangle + |1111\rangle$), максимально запутанное состояние четырех служебных кубитов; последние четыре вентиля Адамара преобразуют кот-состояние в состояние Шора. Но единственная ошибка, возникающая в ходе выполнения второго или третьего XOR-вентиля, может повлечь за собой появление двух ошибок в кот-состоянии (оно может принять вид ($|0011\rangle + |1100\rangle$)). Эти две ошибки инвертирования бита в кот-состоянии превращаются в две фазовые ошибки в состоянии Шора, что по возвращении в исходное состояние во время измерения синдрома приведет к фазовой ошибке в блоке.

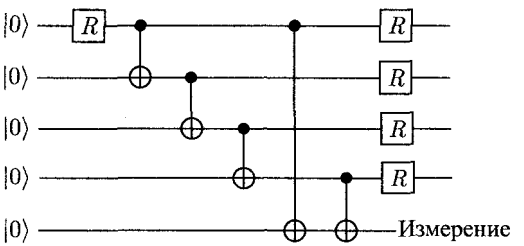


Рис. 8. Приготовление и проверка состояния Шора. Если результат измерения равен 1, тогда это состояние бракуется и готовится новое

Заметим, однако, что во всех случаях, когда один плохой вентиль может вызвать в кот-состоянии две ошибки инвертирования бита, разные значения будут иметь первый и четвертый биты кот-состояния. Следовательно, мы добавляем к схеме два последних XOR-вентиля (сопровождаемые измерением), чтобы проверить соответствие этих двух битов кот-состояния. Если проверка успешна, мы можем перейти к измерению синдрома, уверенные в том, что вероятность возникновения двух фазовых ошибок в состоя-

нии Шора имеет порядок ϵ^2. Если проверка дает сбой, мы можем выкинуть кот-состояние и совершить новую попытку.

Конечно, единственная ошибка в схеме приготовления также может привести к появления двух *фазовых* ошибок в кот-состоянии и, следовательно, двух ошибок инвертирования бита в состоянии Шора; мы не пытались проверять состояние Шора на наличие ошибок инвертирования битов. Но, по сравнению с фазовыми ошибками, они нас беспокоят гораздо меньше. Ошибки инвертирования бита становятся причиной ошибочного измерения синдрома, но они не распространяются в обратном направлении и не повреждают закодированную информацию.

Если для измерения синдрома используется метод Стина, то сначала с помощью изображенной на рисунке 3 кодирующей схемы (но без первых двух XOR-вентилей) строится состояние $|0\rangle_{\text{code}}$, а затем, чтобы завершить приготовление состояния Стина, к каждому кубиту $|0\rangle_{\text{code}}$ применяется вентиль Адамара. Поскольку вновь единственная ошибка, возникающая в процессе кодирования, может привести к появлению в $|0\rangle_{\text{code}}$ двух ошибок инвертирования бита, которые в состоянии Стина преобразуются в две фазовые ошибки, то на этом этапе также необходима проверка. Ее можно осуществить, выполняя неразрушающее измерение, которое гарантирует, что (с точностью до одного инвертирования бита) приготовленным состоянием является $|0\rangle_{\text{code}}$, а не $|1\rangle_{\text{code}}$. С этой целью готовятся два блока в состоянии $|0\rangle_{\text{code}}$, затем производится побитовое выполнение XOR-вентиля с первым блоком в качестве источника и вторым — в качестве цели и, наконец, выполняется разрушающее измерение второго блока. Применяя к результатам этого измерения классический корректирующий код Хэмминга, можно исправить одну возможную ошибку инвертирования бита и идентифицировать измеренный блок как $|0\rangle_{\text{code}}$ или $|1\rangle_{\text{code}}$. Если результатом последнего измерения является $|0\rangle_{\text{code}}$, то это означает, что другой блок прошел проверку. В противном случае считается, что он поврежден и должен быть восстановлен путем инвертирования битов.

Однако эта процедура проверки все еще ненадежна, поскольку на самом деле поврежденным может оказаться измеряемый нами блок, а не тот, который мы пытаемся проверить. Следовательно, необходимо повторить этап проверки. Если измеряемый блок выдает один и тот же результат дважды подряд, то проверка может считаться надежной. А что, если во второй раз будет получен другой результат? Тогда непонятно, что делать, инвертировать биты проверяемого блока или оставить его в покое. Можно сделать еще одну попытку, чтобы разорвать этот узел, но на самом деле в этом нет необходимости; в действительности, если две попытки исправления дают противоречивые результаты, то лучше ничего не предпринимать. Поскольку

результаты противоречат друг другу, то понятно, что один из двух измеряемых блоков поврежден. Следовательно, вероятность того, что проверяемый блок тоже ошибочный, имеет порядок ϵ^2, и ею можно пренебречь. Итак, с помощью этой процедуры проверки удается построить такое состояние Стина, в котором вероятность появления многократных фазовых ошибок (способных распространяться на закодированные данные при измерении синдрома) имеет порядок ϵ^2.

3.4. Проверка синдрома

Одна ошибка инвертирования бита в служебном состоянии может привести к повреждению синдрома. Ошибка может возникнуть из-за неправильного приготовления служебного состояния или вследствие ошибки, появившейся в процессе вычисления синдрома. Последний случай особенно опасен, потому что возникающая с вероятностью порядка ϵ единственная ошибка может повредить как блок данных, так и служебное состояние. Это может случиться, например, вследствие того, что плохой XOR-вентиль одновременно вносит ошибки в источник и цель, или из-за того, что в блоке данных в ходе измерения синдрома возникает ошибка, которая затем распространяется XOR-вентилем на состояние служебных кубитов.

В таком случае, если бы мы воспользовались поврежденным синдромом для исправления ошибки, то на самом деле ввели бы в блок закодированных данных дополнительную ошибку. Поэтому наша процедура остается не полностью отказоустойчивой; возникающая с вероятностью порядка ϵ ошибка может привести к фатальному повреждению данных.

Следовательно, необходимо найти возможность обеспечить более надежный синдром. Очевидным способом является повторение измерения синдрома. Оно не обязательно, если результат измерения синдрома тривиален (не сообщает об ошибке); даже если в блоке данных имеется ошибка, которая оказалась не замеченной, не стоит беспокоиться, что это усугубит ситуацию, просто потому, что в этом случае не совершается никаких действий. С другой стороны, если синдром сообщает об ошибке, то его измерение повторяется. В этом случае, если вновь получается тот же самый результат, можно с уверенностью принимать синдром и переходить к исправлению, поскольку невозможно получение с вероятностью ϵ одного и того же (нетривиального) ошибочного синдрома.

Если первые два измерения синдрома не согласуются, тогда можно продолжать его измерение до тех пор, пока один и тот же результат не будет получен два раза подряд — результат, которому можно доверять. В качестве альтернативы, можно ничего не предпринимать, пока ошибка не

будет надежно детектирована в следующем раунде коррекции. По крайней мере, это бездействие не усугубит ситуацию, а ошибка, если она на самом деле присутствует в закодированных данных, возможно, будет обнаружена в следующий раз.

Существуют и другие способы повысить нашу уверенность в синдроме. Например, вместо повторения измерения целого синдрома можно было бы вычислить некоторые дополнительные избыточные биты синдрома и подвергнуть их проверке на четность. Если в синдроме содержится ошибка, то этот метод, как правило, ее обнаружит; таким образом, если проверка на четность проходит успешно, то корректность синдрома является достаточно вероятной [24, 32].

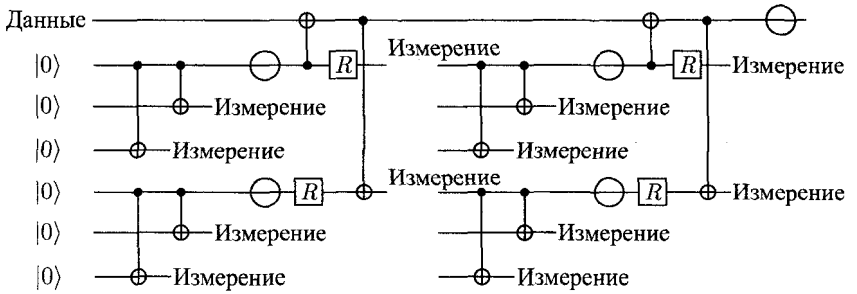


Рис. 9. Полная схема исправления ошибок согласно Стину. Готовятся и затем проверяются закодированные состояния $|0\rangle$. Проверенные $|0\rangle$ используются в качестве служебных кубитов для вычисления синдромов как инвертирования битов, так и обращения фазы каждый; измеряется дважды. Большие круги обозначают действия (обусловленные результатами измерения), предпринимаемые для восстановления служебных состояний, или действия по восстановлению блока данных на финальном этапе

Итак, мы собрали вместе все элементы отказоустойчивой процедуры исправления ошибок. Если мы предпримем все описанные выше меры предосторожности, то исправление даст сбой только при появлении двух независимых ошибок, поэтому вероятность необратимого повреждения закодированного блока имеет порядок ϵ^2.

Полная квантовая схема исправления ошибок согласно Стину показана на рисунке 9. Отметим, что исправление ошибок как инвертирования бита, так и обращения фазы повторяются дважды. На схеме показана и проверка состояний Стина, но этап кодирования этих состояний опущен.

3.5. Измерение и кодирование

Конечно, мы хотим уметь надежно измерять наши закодированные кубиты. Но как уже отмечалось в разделе 2, *разрушающее* измерение кодового блока надежно, если только один кубит в блоке имеет ошибку инвертирования бита. Если в случае одного кубита вероятность ошибочного измерения имеет порядок ϵ, то ошибочные измерения кодового блока возникают с вероятностью порядка ϵ^2. Отказоустойчивое неразрушающее измерение также может быть выполнено, как уже отмечалось при обсуждении проверки состояния Стина (раздел 3.3). В качестве альтернативной процедуры можно было бы без каких-либо модификаций использовать изображенное на рис. 4 неразрушающее измерение. Хотя служебный бит является целью трех последовательных XOR-вентилей, возвращающиеся в блок фазовые ошибки не так губительны, поскольку они не могут заменить $|0\rangle_{\text{code}}$ на $|1\rangle_{\text{code}}$ (или наоборот). Но, поскольку единственная ошибка инвертирования бита (как в блоке данных, так и в служебном кубите) может вызвать ошибку измерения четности, для обеспечения точности порядка ϵ^2 измерение необходимо повторить (после исправления ошибки инвертирования бита). (Мы умолчали об этой процедуре в описании проверки состояния Стина, чтобы избежать разочарования от необходимости коррекции ошибок при подготовке служебных кубитов, предназначенных для коррекции ошибок!)

Часто нам будет необходимо приготовить известное закодированное квантовое состояние, например, $|0\rangle_{\text{code}}$. Мы только что обсудили (в разделе 3.3 в связи с приготовлением состояния Стина), как можно надежно выполнить это кодирование. В сущности, кодирующая схема не нужна. Каким бы ни было исходное состояние блока, (отказоустойчивое) исправление ошибок спроецирует его на натянутое на $\{|0\rangle_{\text{code}}, |1\rangle_{\text{code}}\}$ пространство, а (проверенное) измерение выдаст либо $|0\rangle_{\text{code}}$, либо $|1\rangle_{\text{code}}$. Если получен результат $|1\rangle_{\text{code}}$, то для обращения блока в желаемое состояние $|0\rangle_{\text{code}}$ можно применить (побитовый) NOT-оператор.

Если мы хотим закодировать неизвестное квантовое состояние, то мы используем кодирующую схему, изображенную на рисунке 3. Снова, вследствие распространения ошибок, единственная ошибка в процессе кодирования может привести к сбою. В данном случае, поскольку никакое измерение не может проверить кодирование, точность воспроизведения закодированного состояния будет неизбежно равна $F = 1 - O(\epsilon)$. Тем не менее, смысл в кодировании не теряется, поскольку оно позволяет хранить состояние с достаточной точностью воспроизведения в течение более продолжительного времени, чем если бы состояние оставалось незакодированным.

3.6. Другие коды

Схемы Шора и Стина для отказоустойчивого измерения синдрома выше были описаны лишь для 7-кубитового кода, но их можно адаптировать и для более сложных кодов, способных исправлять множество ошибок [16, 33]. Измерение синдрома для более общих кодов лучше всего описывается с использованием формализма стабилизаторов. Согласно этому формализму, каждый генератор можно представить в виде произведения действующих на отдельный кубит операторов, где однокубитовые операторы выбраны из определяемого уравнением (5) множества $\{I, X, Y, Z\}$. Каждый генератор при возведении в квадрат дает единицу и имеет одинаковые количества собственных векторов с собственными значениями $+1$ и -1, так что определение его собственного значения вдвое сокращает размерность пространства. Если в блоке содержится n кубитов и имеется $n-k$ генераторов, то кодовое подпространство имеет размерность 2^k, то есть имеется k закодированных кубитов.

Например, 7-кубитовый код Стина представляет собой пространство, в котором все шесть генераторов стабилизатора

$$\begin{aligned} M_1 &= (IIIZZZZ), \\ M_2 &= (IZZIIZZ), \\ M_3 &= (ZIZIZIZ), \\ M_4 &= (IIIXXXX), \\ M_5 &= (IXXIIXX), \\ M_6 &= (XIXIXIX) \end{aligned} \tag{18}$$

имеют собственное значение $+1$. Сравнение с (1) показывает, что пространство с $M_1 = M_2 = M_3 = 1$ натянуто на кодовые слова, удовлетворяющие проверке четности Хэмминга, а поскольку адамаровский поворот базиса меняет местами Z и X, то пространство с $M_4 = M_5 = M_6 = 1$ натянуто на кодовые слова, удовлетворяющие проверке четности Хэмминга в повернутом базисе. Действительно, определяющим свойством кода Стина является то, что проверка четности Хэмминга удовлетворяется в обоих случаях.

Генераторы стабилизатора выбираются таким образом, чтобы оператор каждой подлежащей исправлению ошибки (также представляемый в виде произведения однокубитовых операторов $\{I, X, Y, Z\}$), а также произведение любых двух различных операторов ошибок антикоммутировали по крайней мере с одним генератором. Таким образом, каждая ошибка изменяет собственное значение одного из генераторов, а две независимые ошибки

всегда изменяют собственные значения своим, отличным от других, способом. Это означает, что измерение собственных значений всех генераторов стабилизатора дает полный синдром ошибки.[1]

Измерение генератора стабилизатора M не сложно. Сначала выполняется подходящее унитарное преобразование каждого кубита, так чтобы в новом базисе оператор M представлял собой произведение действующих на отдельные кубиты I и Z. (Для каждого кубита, находящегося под действием X в M, мы совершаем поворот

$$R = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{19}$$

а также поворот

$$R' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \tag{20}$$

для каждого кубита, находящегося под действием Y.) В этом базисе собственное значение M представляет собой четность битов, на которые действуют Z. Ее можно измерить (примерно так же, как это делалось при обсуждении 7-кубитового кода), применяя к служебным кубитам XOR-вентили, использующие в качестве источников кубиты в блоке, на которые в новом базисе действуют операторы Z из M. После возвращения в исходный базис эта процедура повторяется для следующего генератора стабилизатора, и так далее до тех пор, пока не будет получен полный синдром.

Мы можем сделать эту процедуру отказоустойчивой, заготовив для измерения каждого бита синдрома служебное состояние Шора, количество кубитов в котором равно *весу* соответствующего генератора стабилизатора (количеству не равных единичному однокубитовых операторов). Каждый служебный бит является целью лишь одного XOR-вентиля, так что многократные фазовые ошибки не возвращаются в данные. Рассмотренные выше процедуры проверки состояния Шора и измерения синдрома также можно соответствующим образом обобщить.

Для сложных кодов, которые либо кодируют множество кубитов, либо исправляют большое количество ошибок, обобщенный метод Шора требует значительно больше служебных кубитов и квантовых вентилей, чем

[1] На самом деле произведение операторов двух независимых ошибок может принадлежать стабилизатору. Тогда эти две ошибки будут иметь один и тот же синдром. Однако это не имеет значения, поскольку они также могут быть исправлены одним и тем же действием. Квантовые коды, сопоставляющие один синдром нескольким различным операторам ошибок, называются вырожденными.

это необходимо для выявления синдрома ошибки. В этом случае гораздо лучше воспользоваться обобщенным методом Стина. Идея Стина состоит в том, что в случае 7-кубитового кода для измерения всех M_1, M_2 и M_3 можно использовать одно 7-кубитовое служебное состояние. Действительно, приготовим исходное служебное состояние в виде равновзвешенной суперпозиции всех строк, удовлетворяющих проверке четности Хэмминга (то есть всех слов классического кода Хэмминга) применим к нему подходящие XOR-вентили, использующие кубиты блока в качестве источников, измерим все кубиты служебного состояния и, наконец, применим проверку четности Хэмминга к результату измерения. Три полученных бита четности представляют собой собственные значения операторов M_1, M_2 и M_3. Служебное состояние здесь приготовлено таким образом, чтобы, помимо этих собственных значений, никакая другая информация не могла быть извлечена из результата измерения; следовательно, эта процедура не разрушает когерентность квантовых кодовых слов.

Очевидно, эту процедуру можно адаптировать для одновременного измерения собственных значений любой совокупности операторов, имеющих вид произведений I и Z, действующих на отдельные кубиты. С этой целью по заданному списку k таких n-кубитовых операторов M_i $(1 \leqslant i \leqslant k)$ построим $k \times n$-матрицу H_Z, ij-элемент которой равен нулю (единице), если в j-й позиции M_i находится оператор I (Z). Затем приготовим n-кубитовое служебное состояние, представляющее собой равновзвешенную суперпозицию всех строк длины n, прошедших проверку четности матрицей H_Z. Применяя к этому состоянию соответствующие XOR-вентили и измеряя результат (а также применяя H_Z к результату измерения), спроецируем n-кубитовый блок на общее собственное состояние рассматриваемого семейства операторов M_i. Выполняя аналогичную процедуру в базисе, повернутом преобразованием Адамара, можно одновременно измерить собственные значения любой совокупности операторов, имеющих вид произведений I и X.

Среди генераторов стабилизатора могут быть и операторы, имеющие вид $M = \bar{Z}\bar{X}$, где $\bar{Z}$ — произведение операторов Z, действующих на одну совокупность кубитов, а $\bar{X}$ — произведение операторов X, действующих на другую совокупность кубитов. Поскольку квадрат каждого генератора M должен быть равен единице, число кубитов, на которые действует оператор Y, равный произведению Z и X, должно быть четным. Следовательно, $\bar{Z}$ и $\bar{X}$ коммутируют, и их можно измерить при помощи описанного выше метода. Однако такое измерение даст слишком много информации; мы хотим измерить произведение $\bar{Z}\bar{X}$, а не каждый из сомножителей $\bar{Z}$ и $\bar{X}$ в отдельности. Для осуществления желаемого измерения необходима дальнейшая модификация служебного состояния. Оно не должно выби-

раться одновременно удовлетворяющим проверке четности матрицами H_Z и H_X. Скорее, его следует приготовить так, чтобы биты четности H_Z и H_X были коррелированы, то есть служебное состояние должно представляться суммой таких строк, у которых оба бита четности либо тривиальны, либо нетривиальны. После измерения служебного состояния собственное значение оператора M получается суммированием четностей «$\bar{Z}$- и $\bar{X}$-измерения». Индивидуальные четности $\bar{Z}$- и $\bar{X}$- измерений являются абсолютно случайными и фактически ничего не сообщают о значениях операторов $\bar{Z}$ и $\bar{X}$.

Теперь мы можем описать метод Стина в его общей форме, в которой он применим к любому стабилизирующему коду. Если k логических кубитов закодированы в блоке из n кубитов, то существует $n-k$ независимых генераторов стабилизатора. Со списком этих генераторов ассоциируется матрица

$$\bar{H} = (H_Z \,|\, H_X), \tag{21}$$

имеющая $n-k$ строк и $2n$ столбцов. Позиции единиц в H_Z обозначают кубиты, на которые в перечисленных генераторах действует Z, а положения единиц в H_X обозначают кубиты, находящиеся под действием X; если единица возникает в одном и том же положении в H_Z и H_X, то на этот кубит действует произведение $Y = ZX$. Далее, готовится служебное $2n$-кубитовое *обобщенное состояние Стина* — равновзвешенная суперпозиция всех строк, удовлетворяющих проверке четности $\bar{H}$. Затем выполняется изображенная на рисунке 10 квантовая схема, измеряются служебные кубиты и к результату измерения применяется $\bar{H}$. Полученные биты четности являются собственными значениями генераторов стабилизатора, что дает полный синдром ошибки.

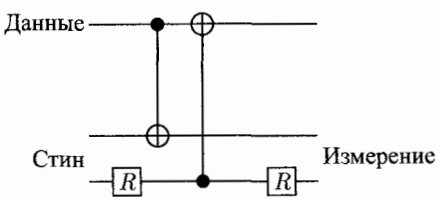


Рис. 10. Схема Стина для измерения синдрома. $2n$-кубитовое состояние Стина используется для определения синдрома для n-кубитового блока данных. Каждый XOR-вентиль на схеме представляет n выполняемых параллельно XOR-вентилей

Служебное состояние здесь приготовлено таким образом, что из результата измерения не может быть извлечена никакая другая информация,

кроме синдрома, и, следовательно, эта процедура не нарушает когерентность квантовых кодовых слов. В этой процедуре на каждый кубит в кодовом блоке действуют только два квантовых вентиля — необходимый минимум для определения как ошибки инвертирования бита, так и ошибки обращения фазы, поражающих любой кубит.

В заключение отметим, что Китаевым [34] была описана другая стратегия выполнения отказоустойчивого исправления ошибок. Он предложил семейство корректирующих ошибки квантовых кодов, способных исправить множество ошибок в кодовом блоке и требующих лишь четыре XOR-вентиля для вычисления каждого бита синдрома. В данном случае, даже если для вычисления каждого бита синдрома мы будем использовать лишь один служебный кубит (а не обобщенное служебное состояние, аналогичное состоянию Стина или Шора), из служебного кубита обратно в данные может вернуться лишь ограниченное количество ошибок. Тогда код можно выбрать таким образом, чтобы типичное количество ошибок, вернувшихся в данные в процессе вычисления синдрома, было заведомо меньше максимального числа ошибок, которое код еще может выдержать.

4. Отказоустойчивые квантовые вентили

Мы убедились, что кодирование может защитить квантовую информацию. Но мы хотим больше, нежели просто *хранить* квантовую информацию с высокой точностью воспроизведения; мы хотим оперировать квантовым компьютером, который *обрабатывает* эту информацию. Конечно, мы могли бы декодировать информацию, применить вентиль, а затем снова закодировать ее, но эта процедура на время подвергла бы опасности квантовую информацию. Если мы хотим, чтобы наш квантовый компьютер работал надежно, вместо этого,мы должны научиться применять квантовые вентили непосредственно к закодированным данным. В свою очередь, эти вентили должны удовлетворять принципам отказоустойчивости, если мы хотим избежать катастрофического распространения ошибок.

4.1. 7-кубитовый код

Действительно, существует ряд вентилей, которые можно легко применять с 7-кубитовым кодом Стина. Все три однокубитовых вентиля могут применяться *побитово*; то есть применение этих вентилей к каждому из семи кубитов в блоке обеспечивает применение того же вентиля к закодированному кубиту. Мы уже видели, что именно так действует поворот Адамара R [см. уравнение (11)]. То же самое можно сказать о вентиле NOT

(так как каждое нечетное кодовое слово Хэмминга является дополнением четного),[1] и вентиле сдвига фазы

$$P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{22}$$

(Нечетные кодовые слова Хэмминга имеют вес $\equiv 3 \pmod 4$, а четные — вес $\equiv 0 \pmod 4$, поэтому для выполнения P мы фактически применяем P^{-1} побитово.) XOR-вентиль также можно выполнять побитово, то есть путем применения к каждому кубиту блока целей своего XOR-вентиля, использующего соответствующий кубит из блока источников, как это показано на рисунке 11. Эта процедура работает, потому что четные кодовые слова образуют субкод, тогда как нечетные кодовые слова — его нетривиальный смежный класс.

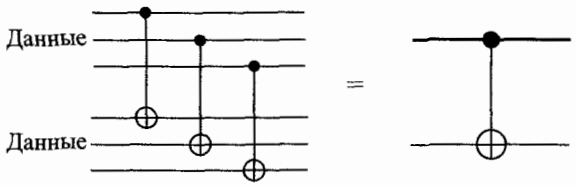


Рис. 11. Схематическое изображение трансверсального XOR-вентиля. Применяя XOR-вентиль к каждому кубиту из блока цели и соответствующему кубиту блока источника, мы выполняем XOR-вентиль, действующий на закодированные кубиты. Реализация вентиля отказоустойчива, поскольку на каждый кубит в обоих кодовых блоках действует свой вентиль

Таким образом, существуют простые отказоустойчивые процедуры выполнения вентилей NOT, R, P и XOR. Но к сожалению, сами по себе они не образуют универсального набора. Для того чтобы иметь возможность выполнять произвольные унитарные преобразования закодированной квантовой информации (с произвольной точностью), этот набор необходимо соответствующим образом дополнить. Следуя Шору [15], добавим к нему 3-кубитовый вентиль Тоффоли, который выполняется с помощью изображенной на рисунке 13 процедуры.[2]

Конструкция Шора отказоустойчивого вентиля Тоффоли предусматривает два этапа. На первом этапе три закодированных служебных блока го-

[1] Собственно, мы можем применить операцию NOT, действующую на закодированный кубит, при помощи лишь трех NOT, применяемых к выбранным кубитам в блоке.

[2] Нилл и др. [19, 20] описывают альтернативный способ комплектования универсального набора вентилей.

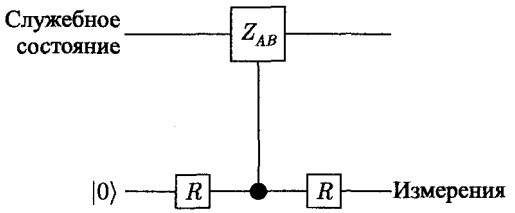


Рис. 12. Схема измерения, используемая на этапе приготовления служебного состояния для предложенной Шором реализации вентиля Тоффоли

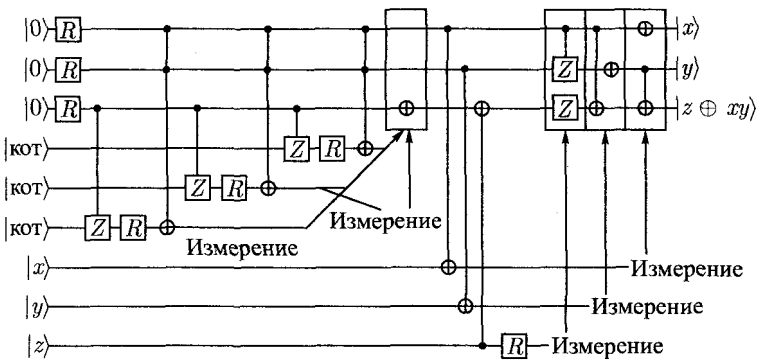


Рис. 13. Отказоустойчивый вентиль Тоффоли. Каждая линия представляет блок из семи кубитов, а вентили выполняются трансверсально. Для каждого измерения стрелка указывает на совокупность вентилей, которые применяются, если результатом измерения является 1, и не действуют, если результатом является 0

товятся в состоянии вида

$$|A\rangle_{\text{anc}} \equiv \frac{1}{2} \sum_{a=0,1} \sum_{b=0,1} |a, b, ab\rangle_{\text{anc}}. \tag{23}$$

На втором этапе, чтобы завершить выполнение вентиля, служебное состояние взаимодействует с тремя блоками данных. Прежде всего опишем способ приготовления служебного состояния. Для начала каждый из трех служебных блоков кодируется в состоянии $|0\rangle_{\text{code}}$. Для приготовления закодированного состояния ко всем трем блокам применяются побитовые вентили

Адамара

$$\frac{1}{\sqrt{8}} \sum_{a=0,1} \sum_{b=0,1} \sum_{c=0,1} |a,b,c\rangle_{\text{anc}}. \tag{24}$$

Отметим, что это состояние можно представить в виде

$$\frac{1}{\sqrt{2}}(|A\rangle_{\text{anc}} + |B\rangle_{\text{anc}}), \quad |B\rangle_{\text{anc}} \equiv \text{NOT}_3|A\rangle_{\text{anc}}, \tag{25}$$

где NOT_3 обозначает NOT-вентиль, действующий на третий закодированный кубит. В завершение приготовления служебного состояния эти три блока измеряются в базисе $\{|A\rangle_{\text{anc}}, |B\rangle_{\text{anc}}\}$; если получается результат $|A\rangle_{\text{anc}}$, приготовление завершено, если результат измерения $|B\rangle_{\text{anc}}$, для завершения процедуры применяется вентиль NOT_3.

Теперь объясним, как выполняется измерение в базисе $\{|A\rangle_{\text{anc}}, |B\rangle_{\text{anc}}\}$. Схематически оно осуществляется с помощью показанной на рисунке 12 схемы, где вентиль Z_{AB} (обусловленный управляющим битом) обращает относительную фазу $|A\rangle_{\text{anc}}$ и $|B\rangle_{\text{anc}}$. Из уравнений (23) и (25) можно увидеть, что Z_{AB} применяет фазовый множитель $(-1)^{ab+c}$, выражающийся через значения трех служебных блоков a, b и c. Если контрольный бит обозначается x, тогда вентилями, которые нам необходимо применить, являются $(-1)^{xab}$ и $(-1)^{xc}$ — произведение трехбитового и двухбитового фазовых вентилей.

Но трехбитовый фазовый вентиль так же сложен, как и вентиль Тоффоли, так что, похоже, мы попали в тупик. Однако из него можно выйти, выбрав управляющий блок таким образом, чтобы им был не закодированный кубит, а (проверенное) 7-битовое «кот-состояние»

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}(|0000000\rangle + |1111111\rangle). \tag{26}$$

Мы уже знаем, каким образом построить трансверсально действующие отказоустойчивые *двух-* и *однокубитовые* фазовые вентили. Их можно повысить до необходимых нам трех- и двухкубитовых вентилей простым управлением всех побитовых вентилей конструкции соответствующими битами кот-состояния. Наконец, мы применяем побитовый поворот Адамара к кот-состоянию и измеряем его четность, завершая таким образом выполнение схемы измерения на рисунке 12. (Мы получим схему на рисунке 13, заметив, что если кот-состояние находится в базисе полученном повернутом Адамара, то трехбитовый фазовый вентиль можно представить как вентиль Тоффоли с кот-состоянием в качестве цели; следовательно, при таком исполнении схемы измерения осуществляется один побитовый вентиль Тоф-

фоли). Конечно, для обеспечения точности необходимо повторение измерения.

Все это время три блока данных терпеливо ожидали готовности служебного кубита. Применением трех XOR-вентилей и поворота Адамара, состояние данных и служебного кубита преобразуются как

$$\begin{aligned}&\sum_{a=0,1}\sum_{b=0,1}|a,b,ab\rangle_{\text{anc}}|x,y,z\rangle_{\text{data}}\to\\&\to\sum_{a=0,1}\sum_{b=0,1}\sum_{w=0,1}(-1)^{wz}|a,b,ab\oplus z\rangle_{\text{anc}}|x\oplus a,y\oplus b,w\rangle_{\text{data}}.\end{aligned}\tag{27}$$

Теперь выполняется измерение каждого *блока данных*. Если результат измерения 0, то никакие действия не предпринимаются, но если результат измерения 1, то для завершения выполнения вентиля Тоффоли к служебному кубиту применяется определенный набор вентилей, как это показано на рисунке 13. Отметим, что эта процедура разрушает исходные блоки данных, а новыми блоками данных становятся те, что в начале являлись служебными. Важной особенностью этой конструкции является то, что все этапы тщательно построены в соответствии с принципами отказоустойчивости и минимизации распространения ошибок. Таким образом, для того чтобы в любом одном из блоков данных одновременно появились две ошибки, в ходе выполнения процедуры должны возникнуть две независимые ошибки.

Немного досадно, что отказоустойчивые вентили образуют дискретный набор, но это также неизбежная черта любой отказоустойчивой схемы. Нет смысла создавать континуум отказоустойчивых вентилей, поскольку как тогда мы сможем избежать ошибки, применяя неправильный вентиль, отличающийся от предполагаемого на малую величину? В любом случае, так как наши отказоустойчивые вентили образуют универсальный набор, их достаточно для приближения любого желаемого унитарного преобразования с любой желаемой точностью.

4.2. Другие коды

Шор [15] описал, как обобщить этот отказоустойчивый набор вентилей на более сложные коды, способные исправлять большее количество ошибок, а Готтесман [17,35] описал еще более общую процедуру, которую можно применить к любым квантовым стабилизирующим кодам.

Конструкция Готтесмана начинается со следующего наблюдения: для любого стабилизирующего кода существует отказоустойчивая реализация

однокубитовых вентилей X и Z, действующих на каждый закодированный кубит. Вспомним, что мы уже видели в разделе 3.6, что для стабилизирующего кода с размером блока n любой «оператор ошибки» M (представленный в виде тензорного произведения n матриц из набора $\{I, X, Y, Z\}$) можно записать в виде $\bar{Z}\bar{X}$ и, следовательно, однозначно представить в виде двоичной строки длины $2n$. Если имеется k закодированных в блоке логических кубитов, то стабилизатор кода генерируется $n-k$ такими операторами. Коммутирующие со всеми элементами стабилизатора операторы ошибок сами образуют группу. Генераторы этой группы представляются двоичными строками длины $2n$, которые должны удовлетворять $n-k$ независимым двоичным условиям; следовательно, имеется $n+k$ независимых генераторов. Из них $n-k$ являются генераторами стабилизатора, но есть $2k$ дополнительных независимых операторов ошибок, не принадлежащих стабилизатору, но коммутирующих с ним. Эти $2k$ операторов сохраняют кодовое подпространство, но действуют нетривиально на кодовые слова и, следовательно, их можно интерпретировать как операции, действующие на закодированные логические кубиты.

Собственно, эти $2k$ операторов можно выбрать в качестве однокубитовых операций $\hat{Z}_i$ и $\hat{X}_i$, где $i = 1, 2, 3, \ldots k$ нумерует закодированные кубиты. Сначала отметим, что $n-k$ генераторов стабилизатора можно расширить до максимального набора n коммутирующих операторов; k дополнительных операторов можно идентифицировать как операторы $\hat{Z}_i$. Мы можем выбрать состояния вычислительного базиса в кодовом подпространстве таким образом, чтобы они одновременно были собственными состояниями всех операторов $\hat{Z}_i$, с собственным значением $+1$, соответствующим значению 0, и собственным значением -1 соответствующим значению 1. Тогда $\hat{Z}_i$ обращает фазу i-го кубита. Остальные k генераторов $\hat{X}_i$, которые коммутируют со стабилизатором, но не коммутируют со всеми $\hat{Z}_i$, можно выбрать так, чтобы они удовлетворяли соотношениям

$$[\hat{Z}_i, \hat{Z}_j] = 0 = [\hat{X}_i, \hat{X}_j], \quad [\hat{Z}_i, \hat{X}_j] = 0, \ (i \neq j), \qquad \hat{Z}_i\hat{X}_i + \hat{X}_i\hat{Z}_i = 0. \quad (28)$$

Так как $\hat{X}_i$ антикоммутирует с $\hat{Z}_i$, он обращает собственное значение $\hat{Z}_i$ и, следовательно, значение i-го кубита. Все эти операции выполняются с применением максимум одного однокубитового вентиля к каждому кубиту в блоке; следовательно, эти операции без сомнения отказоустойчивые.

Мы также видели в разделе 3.6, что возможно выполнение отказоустойчивого *измерения любого* оператора $\bar{Z}\bar{X}$, а значит, в частности, возможно отказоустойчивое измерение любого оператора ошибки $\hat{X}_i$, $\hat{Y}_i$, $\hat{Z}_i$. Готтесман [17] показал, что если возможно выполнение вентиля Тоффоли

(который является универсальным для *классических* вычислений, сохраняющих набор состояний вычислительного базиса), то для универсальных квантовых вычислений достаточно однокубитовых вентилей X и Z вместе с возможностью измерения X, Y и Z для любого кубита. Следовательно, если мы покажем, что, действуя на любые три кубита, можно построить отказоустойчивый вентиль Тоффоли, то этим завершим доказательство возможности универсальных отказоустойчивых квантовых вычислений для любого стабилизирующего кода.

Конструкция отказоустойчивого вентиля Тоффоли достаточно сложна, поэтому доказательство лучше построить следующим образом: Готтесман показал, что в любом стабилизирующем коде отказоустойчивый XOR-вентиль может применяться к любой паре кубитов (независимо от того, находятся или нет эти два кубита в одном и том же кодовом блоке). Используя XOR-вентиль, а также однокубитовые вентили и измерения, которые, как мы только что видели, можно выполнять отказоустойчиво, можно построить все вентили, необходимые в конструкции Шора вентиля Тоффоли. Таким образом, с помощью любого стабилизирующего кода можно реализовать отказоустойчивую схему вентиля Тоффоли, что достаточно для выполнения универсальных отказоустойчивых квантовых вычислений.

Несмотря на то, что отказоустойчивые квантовые вычисления, в принципе, можно реализовать с помощью любого стабилизирующего кода, некоторые из них лучше подходят для этой цели. Например, существует 5-кубитовый код, способный исправить одну ошибку [36,37]; Готтесман продемонстрировал универсальный набор отказоустойчивых вентилей для этого кода. Но их реализация достаточно сложна. 7-кубитовый код Стина требует блок большего размера, но он гораздо удобнее для вычислений.

5. Порог безошибочности квантовых вычислений

Существуют квантовые коды, способные исправить t ошибок, где t может быть сколь угодно большим. Если мы используем такой код и придерживаемся принципов отказоустойчивости, то непоправимая ошибка возникнет только при появлении минимум $t+1$ независимых ошибок в одном блоке до завершения процесса исправления. Поэтому если вероятность возникновения ошибки на квантовый вентиль или вероятность возникновения ошибки хранения на единицу времени имеет порядок ϵ, то вероятность появления ошибки в расчете на один вентиль, действующий на закодированные данные, будет иметь порядок ϵ^{t+1}, что гораздо меньше, чем ϵ, если эта величина достаточно мала. Действительно, может показаться, что при выборе кода со сколь угодно большим t мы можем сделать вероятность воз-

никновения ошибки в расчете на один вентиль сколь угодно малой, но, оказывается, это не так, по крайней мере, не для большинства кодов. Проблема в том, что по мере увеличения t сложность кода резко возрастает, и соответственно возрастает сложность процедуры исправления. В конечном счете, мы достигаем момента, когда выполнение исправления занимает так много времени, что становится возможным накопление $t+1$ ошибок в блоке до полного завершения этапа исправления, и, таким образом, способность кода исправлять ошибки становится сомнительной.

Предположим, что количество вычислительных этапов, необходимых для выполнения измерения синдрома, растет вместе с t, как степень t^b. Тогда вероятность того, что до завершения измерения накопится $t+1$ ошибок, будет вести себя как

$$\text{Block Error Probability} \sim (t^b \epsilon)^{t+1}, \tag{29}$$

где ϵ — вероятность появления ошибки на одном этапе. Тогда мы можем выбрать t, чтобы минимизировать вероятность ошибки ($t \sim e^{-1}\epsilon^{-1/b}$, при большом t), и получить

$$\text{Minimum Block Error Probability} \sim \exp\left(e^{-1} b \epsilon^{-1/b}\right). \tag{30}$$

Таким образом, если мы рассчитываем безукоризненно выполнить все T циклов исправления ошибок, то наши вентили должны иметь точность

$$\epsilon \sim (\log T)^{-b}. \tag{31}$$

Аналогично, для выполнения квантового вычисления с участием T квантовых вентилей необходимы элементарные вентили заданной точности.

В первоначально описанной Шором процедуре [15] степень, характеризующая сложность измерения синдрома, равна $b = 4$; с помощью более оптимизированной процедуры можно достичь несколько меньших значений b. Размер блока используемого кода растет вместе с t как t^2 (для рассмотренных Шором кодов), поэтому при выборе кода для оптимизации вероятности появления ошибки, размер блока имеет порядок $(\log T)^2$. Конечно, описываемый уравнением (31) скейлинг гораздо предпочтительнее, чем точность $\epsilon \sim T^{-1}$, которая потребовалась бы, если бы кодирование не использовалось вообще. Но при любой заданной точности существует предел продолжительности вычисления, при которой еще можно не опасаться появления ошибок.

Это ограничение можно преодолеть, используя код специального типа — *каскадный* код [18–24]. Чтобы понять идею каскадного кода, представим, что мы используем исправляющий ошибки квантовый код Стина,

кодирующий единственный кубит в 7-кубитовом блоке. Но если мы внимательнее рассмотрим один из семи кубитов в блоке, то обнаружим, что на самом деле это не один кубит, а другой блок из семи закодированных с помощью того же кода Стина кубитов. А когда мы с еще бóльшим разрешением исследуем один из семи кубитов в уже этом блоке, мы обнаружим, что он также является блоком из семи кубитов, и так далее (см. рис. 14). Если в этой иерархии каскадного соединения всего имеется L уровней, то отдельный кубит фактически кодируется в блоке размера 7^L. Каскадное соединение оказывается полезным, поскольку, действуя по принципу «разделяй и властвуй», оно позволяет более эффективно избавляться от ошибок. При таком методе сложность этой процедуры уже не так стремительно растет с повышением способности квантового кода исправлять ошибки.

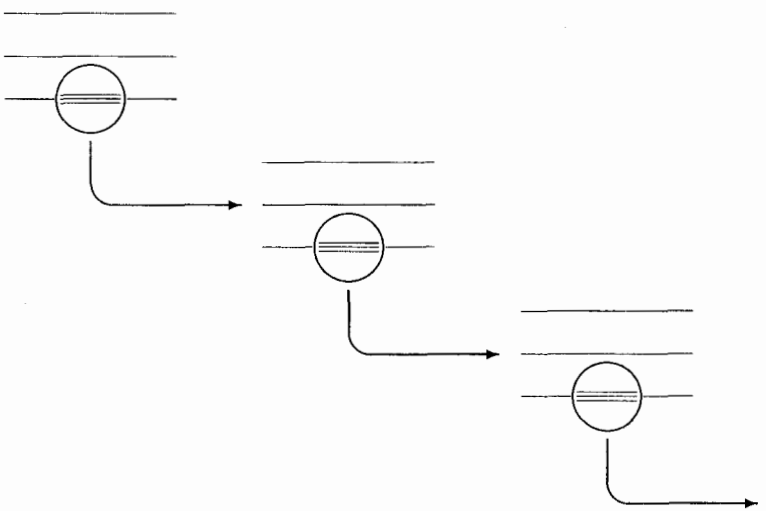


Рис. 14. Каскадное кодирование. Каждый кубит в блоке, при ближайшем рассмотрении, сам по себе является закодированным субблоком

Мы убедились, что 7-кубитовый код Стина может исправить одну ошибку. Если вероятность возникновения ошибки на кубит равна ϵ, ошибки некоррелированы, а исправление отказоустойчиво, тогда вероятность отказа при исправлении имеет порядок ϵ^2. Если два кода соединяются в каскад, образуя блок размера 7^2, то ошибка в нем возникает только при повреждении двух из его субблоков размера семь, что происходит с вероятностью порядка $(\epsilon^2)^2$. При добавлении еще одного уровня каскадирования

ошибка в блоке размера 7^3 возникает только в случае повреждения двух из его субблоков размера 7^2, а вероятность такого события имеет порядок $\left((\epsilon^2)^2\right)^2$, и так далее. При наличии L уровней каскадного соединения кодов вероятность появления ошибки имеет порядок ϵ^{2^L}, тогда как размер блока равен 7^L. Если частота появления ошибок в основных вентилях достаточно мала, то отнесенную к одному вентилю вероятность появления ошибки можно уменьшить с помощью каскадного соединения кодов. Если это так, то добавление следующего уровня каскадирования приведет к дальнейшему понижению вероятности ошибки, и так далее. В этом состоит природа порога безошибочности квантовых вычислений: если кодирование значительно снижает вероятность появления ошибки, то, добавляя достаточное количество уровней каскадирования, частоту появления ошибок можно сделать сколь угодно малой. Но если частоты появления ошибок изначально слишком высоки, тогда вместо улучшения кодирование только усугубит существующее положение вещей.

Чтобы проанализировать эту ситуацию, необходимо принять некоторую конкретную модель ошибок. Я выберу наиболее простую из возможных квазиреалистическую модель: некоррелированые стохастические ошибки.[1] На каждом такте вычислений каждый кубит в устройстве запутывается с окружением. Пусть запутывание описывается уравнением (4) с тем лишь отличием, что теперь четыре состояния окружения предполагаются взаимно ортогональными, а «ошибочные состояния» — имеющими одинаковые нормы. Таким образом, три типа ошибок (инвертирование бита, обращение фазы, обе одновременно) предполагаются равновероятными. Полная вероятность ошибки на каждом такте вычислений обозначается как ϵ_{store}. Помимо этих ошибок запоминающего устройства, повреждающих хранящиеся кубиты, существуют ошибки, вносимые самими квантовыми вентилями. Для каждого типа вентиля вероятность появления ошибки при каждом его выполнении обозначается ϵ_{gate} (при независимых значениях, приписываемых вентилям каждого типа). Если вентиль действует более чем на один кубит (XOR или Тоффоли), то могут возникнуть коррелированые ошибки. Сделаем пессимистическое допущение, что ошибка многокубитового вентиля всегда повреждает все кубиты, на которые он действует; например, неправильный XOR-вентиль вносит ошибки как в кубит источника, так и в кубит цели. Это предположение (среди прочих) сделано лишь для упрощения анализа. При более реалистичных допущениях мы обнаружили бы, что вполне можно пережить и несколько более высокие частоты появления ошибок.

[1] Более детальную характеристику модели ошибок я приведу в разделе 6.

Эффективность каскадного кодирования можно проанализировать, построив систему *потоковых уравнений*, описывающих эволюцию модели ошибки при переходе от одного уровня каскадирования к другому. Пусть, например, мы хотим выполнить XOR-вентиль с последующим этапом исправления ошибок в кубитах, закодированных с помощью каскадного кода Стина с L уровнями каскадирования (размер блока 7^L). Эти процедуры можно описать с помощью таких же операций, но действующих на субблоки размера 7^{L-1}. Таким образом, вероятность появления ошибки $\epsilon^{(L)}$ для вентиля, действующего на блок размера L, может быть выражена через вероятность появления ошибки $\epsilon^{(L-1)}$ для вентиля, действующего на блок размера $L-1$. Это соотношение представляет собой одно из потоковых уравнений. В принципе, решая систему потоковых уравнений, можно получить выражение для вероятности появления ошибки «на уровне L» через параметры модели ошибок и исследовать ее поведение с ростом L. Если вероятность появления ошибки в блоке стремится к нулю с ростом его размера L, то это означает, что вероятности элементарных ошибок лежат «ниже порога». Так как вероятности элементарных ошибок могут зависеть от множества параметров, то на самом деле порог представляет собой некоторую гиперповерхность в многомерном параметрическом пространстве рассматриваемой модели.

Метод Стина измерения синдрома очень хорошо подходит для каскадного кодирования. Все вентили корректирующей схемы (рис. 9) можно выполнять *трансверсально*; выполнение вентилей на элементарных кубитах кодового блока означает применение точно таких же вентилей к информации, закодированной в каждом блоке размера семь, в каждом суперблоке размера 7×7, и так далее. Аналогично, когда при завершении вычисления синдрома мы измеряем элементарные кубиты в служебном блоке, тогда (после применения к кубитам классической коррекции ошибок Хэмминга) мы также измеряем кубиты, закодированные в каждом блоке размера семь, и (после применения коррекции ошибок Хэмминга к блокам) в каждом суперблоке размера 7×7, и так далее. Таким образом, необходимая для извлечения синдрома обработка квантовых данных может быть выполнена *одновременно на всех уровнях каскадного кода*.[1] После чего результаты некоторой классической обработки указывают, какие однокубитовые вентили необходимо применить ко всем элементарным кубитам, чтобы завершить этап исправления на всех уровнях одновременно.

[1] *Разрушающее* измерение закодированного служебного блока можно выполнить одновременно на всех уровнях. Процедура *неразрушающего* измерения блока (проецирование блока на $|0\rangle_{\text{code}}$ или $|1\rangle_{\text{code}}$) гораздо более трудоемка; в каждый момент времени она должна выполняться на одном уровне.

Таким образом, нетрудно понять, как (по крайней мере, в принципе) можно оценить порог безошибочности [38]. На каждом уровне каскадного кода блок из семи кубитов отказывает, если ошибки возникают как минимум в двух его субблоках. Если p_L — вероятность появления ошибки в блоке на уровне L, то вероятность ошибки на уровне $L+1$ равна

$$p_{L+1} \sim \binom{7}{2} p_L^2 + \ldots = 21 p_L^2 + \ldots \tag{32}$$

(с точностью до слагаемых более высокого порядка по p_L), что будет меньше p_L при $p_L < 1/21$. Следовательно, если вероятность появления ошибки в каждом элементарном кубите $p_0 < 1/21$, то на первом уровне кодирования она будет меньше этой величины, еще меньше — на втором уровне и так далее — значение порога безошибочности p_0 равно 1/21.

Предположим, мы выполняем исправление ошибок каждый раз при осуществлении XOR-вентиля или однокубитового вентиля. Грубо говоря, p_0 — это вероятность ошибки на кубит информации в момент начала цикла исправления ошибок. Для оценки порога безошибочности мы следуем схеме на рис. 9 и увеличиваем вследствие появления ошибок значение p_0 (включая возможные ошибки хранения), которые возникли при только что выполненных квантовых вентилях и еще не были устранены предыдущим циклом исправления ошибок. Мы получаем выражение для p_0 через вероятности ошибки вентиля и ошибки хранения; это выражение приравнивается к $1/21$ для определения порога.

Поступая таким образом и предполагая, что ошибки запоминающего устройства пренебрежимо малы, а все однокубитовые и XOR-вентили имеют одинаковую вероятность появления ошибки ϵ_{gate}, мы [38] грубо оцениваем порог частоты ошибок вентиля как

$$\epsilon_{\text{gate},0} \sim 6 \cdot 10^{-4}. \tag{33}$$

Аналогично, если ошибки вентиля пренебрежимо малы, предполагаемый порог частоты ошибок запоминающего устройства равен

$$\epsilon_{\text{store},0} \sim 6 \cdot 10^{-4}. \tag{34}$$

Пороги для ошибок вентиля и запоминающего устройства по существу одинаковы, так как метод Стина хорошо приспособлен для работы с ошибками хранения. Кубиты редко лежат без дела; практически на каждом этапе на каждый из них действует вентиль. Следовательно, требование к точности запоминающего устройства значительно менее жесткое по сравнению с предыдущими оценками порога, основанными на методе исправления Шора [23, 35, 39].

Однако, как показывает более тщательный анализ, по ряду причин фактическое значение порога, которое можно вывести из схемы на рисунке 9, несколько ниже оценок (33) и (34). Наиболее серьезное требование состоит в том, что для осуществления исправления необходим запас закодированных на уровне L хорошо проверенных состояний $|0\rangle_{\text{code}}$. Отдельное (и весьма сложное) вычисление требуется для определения порога надежного кодирования. Кроме того, необходимо проанализировать реализацию Шора вентиля Тоффоли, чтобы гарантировать возможность его надежного применения к соединенным в каскад блокам.[1] Наконец, для получения точного результата необходимо ограничить вклады высшего порядка в вероятность отказа, которые были опущены в уравнении (32). Полный анализ, учитывающий все эти факторы, еще не выполнен, но вполне разумно предположить, что конечные величины порогов запоминающего устройства и вентиля будут превышать 10^{-4}. Конечно, при более совершенной кодирующей схеме и/или протоколе исправления ошибок, возможно, удастся достичь гораздо более высокого значения порога безошибочности.

Также следует задать вопрос, какой размер блока необходим для обеспечения некоторой заданной точности. Грубо говоря, если пороговая частота ошибок вентиля равна ϵ_0, а фактическая частота ошибок элементарного вентиля равна $\epsilon < \epsilon_0$, то L-кратное каскадирование кода понижает частоту появления ошибок до

$$\epsilon^{(L)} \sim \epsilon_0 \left(\frac{\epsilon}{\epsilon_0}\right)^{2^L}. \tag{35}$$

Таким образом, чтобы быть достаточно уверенными в том, что мы можем безошибочно завершить вычисление при T вентилях, выбранный нами размер блока 7^L должен иметь порядок

$$\begin{matrix}\text{block}\\ \text{size}\end{matrix} \sim \left[\frac{\log \epsilon_0 T}{\log \epsilon_0/\epsilon}\right]^{\log_2 7}. \tag{36}$$

Если каскадный код имеет размер блока n и может исправить $t+1$ ошибок, степень $\log_2 7 \sim 2.8$ в уравнении (36) заменяется на $\log n/\log(t+1)$; эта степень приближается к 2 для рассмотренных Шором семейств кодов, но для «хороших» кодов, в принципе, может приблизиться к 1.

Когда частоты появления ошибок находятся ниже порога безошибочности, также возможно сколь угодно продолжительное хранение *неизвест-*

[1] Элементарные вентили Тоффоли не должны быть такими же точными, как одно- и двухкорпусные вентили — вполне приемлема частота появления ошибок вентиля Тоффоли порядка 10^{-3}, если достаточно низки частоты других ошибок. Этот вывод приятен, поскольку вентили Тоффоли сложнее в применении и на практике, вероятно, имеют меньшую точность.

ного квантового состояния. Однако, как мы уже отмечали в разделе 3.5, если вероятность возникновения ошибки запоминающего устройства на один такт вычисления равна ϵ, то первоначальное кодирование состояния можно выполнить с точностью воспроизведения, не превышающей $F = 1 - O(\epsilon)$. При каскадном кодировании мы можем сколь угодно долго хранить неизвестную квантовую информацию с достаточно высокой точностью воспроизведения, но достичь сколь угодно высокой точности воспроизведения не можем.

Каскадирование — важная теоретическая конструкция, так как она позволяет утверждать, что возможны сколь угодно продолжительные вычисления. Но пока частоты ошибок достаточно далеки от пороговых значений, каскадное кодирование, возможно, будет не лучшим способом выполнения определенного вычисления заданной длины. Действительно, выбранный из первоначально описанного Шором семейства код может оказаться более эффективным, чем каскадный 7-битовый код. Более того, каскадный 7-битовый код и коды Шора кодируют лишь один кубит квантовой информации в достаточно большом кодовом блоке. Но в разделе 4 мы видели, что отказоустойчивые квантовые вычисления можно выполнять, используя любые стабилизирующие коды, в том числе и те, что путем кодирования множества кубитов в одном блоке делают более эффективным использование объема памяти. Если надежность наших аппаратных средств близка к порогу безошибочности, то эти коды будут работать неэффективно. Но с усовершенствованием «железа» можно использовать более эффективные коды и таким образом повышать надежность нашего квантового компьютера при меньших затратах объема памяти.

6. Модели ошибок

Отказоустойчивая схема должна быть приспособлена для защиты от тех типов ошибок, которые могут с большей вероятностью нанести ущерб конкретному устройству. И любое утверждение о величине допустимых частот возникновения ошибок (как и оценка порога безошибочности, набросок которой мы только что сделали) бессмысленно, пока не будет тщательным образом определена модель ошибок. Подытожим некоторые важные предположения о характере ошибок, лежащие в основе нашей оценки порога безошибочности.

- **Случайные ошибки.** Мы предположили, что ошибки не имеют систематической составляющей.[1] Ошибки, имеющие случайные фазы,

[1] Нилл и др. [19] показали существование порога безошибочности для гораздо более общих моделей ошибок.

накапливаются по сценарию случайного блуждания, так что с количеством применяемых вентилей приблизительно линейно растет *вероятность* ошибки. Но если ошибки имеют систематические фазы, тогда линейно с числом применяемых вентилей может расти *амплитуда* вероятности ошибки. Следовательно, чтобы наш квантовый компьютер работал хорошо, частота систематических ошибок должна удовлетворять более жестким требованиям, нежели частота случайных ошибок. Иначе говоря, если мы допускаем, что систематические фазы тайно договариваются всегда складываться конструктивно, и если для случайных ошибок порог безошибочности равен ϵ_0, то для (максимально законспирированных) систематических ошибок он будет иметь порядок ϵ_0^2. Несмотря на то, что систематические ошибки могут стать проблемой для квантовых инженеров будущего, они не должны представлять собой непреодолимое препятствие. Моя позиция состоит в следующем: (1) даже если наши аппаратные средства предрасположенные к появлению ошибок с систематическими фазами, эти ошибки будут стремиться к взаимному уничтожению в ходе достаточно продолжительного вычисления [40–42], и (2) так как систематические ошибки можно, в принципе, понять и устранить, с фундаментальной точки зрения более важно иметь представление об ограничениях, накладываемых на работу машины случайными ошибками.

- **Некоррелированные ошибки.** Мы предположили, что ошибки не коррелируют как в пространстве, так и во времени. Таким образом, когда мы говорим, что вероятность возникновения ошибки на кубит равна (например) $\epsilon \sim 10^{-5}$, то фактически имеем в виду, что для двух заданных кубитов вероятность одновременного повреждения ошибками обоих равна $\epsilon^2 \sim 10^{-10}$. Это очень сильное предположение. Действительно важное условие состоит в том, чтобы коррелированные ошибки, повреждающие множество кубитов в одном и том же кодовом блоке, были в высшей степени маловероятны, так как кодирующие схемы будут отказывать при появлении нескольких ошибок в одном блоке. Квантовые инженеры будущего столкнутся с проблемой конструирования таких устройств, в которых кубиты одного блока были бы тщательно изолированы друг от друга.

- **Максимальный параллелизм.** Мы предположили, что в течение одного такта может параллельно выполняться несколько квантовых вентилей. Это допущение позволяет выполнять исправление ошибок во всех кодовых блоках одновременно, и поэтому важно для контроля ошибок хранения кубитов. (В противном случае, добавление к коду

следующего уровня каскадирования вело бы к росту вероятности отказа, поскольку каждому не занятому в процессе отдельному кубиту пришлось бы дольше ожидать своей очереди исправления ошибок.) Если мы пренебрегаем ошибками запоминающего устройства, то для анализа порога безошибочности параллельность выполнения операций не обязательна, но для ускорения вычислений она, безусловно, желательна.

- **Независимая от числа кубитов частота появления ошибок.** Мы предположили, что частоты ошибок не зависят от количества хранящихся в нашем устройстве кубитов. Неявно это допущение касается природы аппаратных средств. Например, это предположение было бы необоснованным, если бы все кубиты хранились в единственной ионной ловушке и делили бы один фононный канал передачи информации [43].
- **Вентили могут действовать на любую пару кубитов.** Мы предположили, что наша машина снабжена набором базовых вентилей, которые можно применить к любой паре хранящихся кубитов (или тройке кубитов, в случае вентиля Тоффоли), независимо от их близости друг к другу. На практике возможны издержки как по времени выполнения, так и по частоте появления ошибок, связанные с перемещением кубитов для того, чтобы вентиль мог эффективно действовать на конкретную пару. Выбор архитектуры, минимизирующей эти издержки, мы оставим конструкторам машин. При наличии вентилей, действующих только на соседние кубиты, порог по-прежнему будет существовать [21], но он станет значительно ниже.
- **Новые служебные кубиты.** Мы предположили, что наш компьютер имеет доступ к достаточному запасу новых служебных кубитов. Служебные кубиты используются как для выполнения вентилей (Тоффоли), так и для осуществления исправления ошибок. Вместе с накоплением эффектов случайных ошибок генерируется энтропия, а процесс исправления ошибок выбрасывает ее из вычислительного устройства в служебный регистр. В принципе, пока поставляются свежие служебные кубиты, вычисление может продолжаться сколь угодно долго, но на практике мы захотим очищать служебные кубиты и использовать их повторно. Стирание служебного кубита неизбежно вызовет рассеяние мощности и выделение тепла; поэтому потребуется охлаждение устройства.
- **Отсутствие ошибок утечки.** Мы пренебрегли вероятностью утечки. В нашей модели квантового компьютера каждый из кубитов живет

в двумерном гильбертовом пространстве, и мы предполагаем, что при появлении ошибки этот кубит либо запутывается с окружающей средой, либо поворачивается в двумерном пространстве в непредсказуемом направлении. Но существует и другой возможный тип ошибки, при которой кубит просачивается из двумерного в более широкое пространство [44]. Чтобы контролировать ошибки утечки, мы можем повторно запрашивать каждый кубит (например, используя показанную на рисунке 15 схему определения места утечки), не пытаясь точно диагностировать, что произошло с просочившимся кубитом [23]. При возникновении утечки кубит разрушается, и его необходимо отбросить;[1] мы заменяем его свежим кубитом в стандартном состоянии, скажем, в состоянии $|0\rangle$. Затем мы можем выполнить стандартное измерение синдрома, которое спроецирует этот кубит на такое состояние, что ошибку можно будет исправить с помощью простого унитарного преобразования.[2] При использовании каскадного кодирования детектирование утечки должно выполняться только на самых нижних уровнях кодирования. Схема определения достаточно проста, поэтому наличие ошибок утечки лишь незначительно повлияет на порог безошибочности.

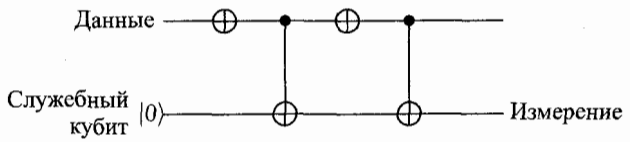


Рис. 15. Квантовая схема детектирования утечки. Допустим, что при утечке данных XOR-вентиль действует тривиально, тогда при возникновении утечки результат измерения равен 0, в противном случае — 1

Предположения нашей модели ошибок в достаточной степени реалистичны, чтобы обеспечить разумную оценку того, насколько хорошо может работать квантовый компьютер при наличии шума. Предположим, например, мы хотим, чтобы наш квантовый компьютер решил сложную задачу факторизации с использованием алгоритма Шора; каким техническим требованиям должна отвечать машина? Располагая самым известным класси-

[1] Конечно, мы можем впоследствии использовать его повторно.

[2] Действительно, так как еще до измерения синдрома мы знаем, что поврежденный кубит находится в определенной позиции в кодовом блоке, для диагностики и коррекции ошибки в известном положении мы можем применить ускоренную версию исправления ошибок [45].

ческим алгоритмом факторизации и самой быстродействующей существующей машиной, факторизацию 130-разрядного (432-битового) числа можно выполнить за несколько месяцев [46]. Чтобы решить эту задачу с помощью алгоритма Шора, мы должны быть в состоянии хранить около $5 \cdot 432 = 2160$ кубитов и выполнить около $38 \cdot (432)^3 \sim 3 \cdot 10^9$ вентилей Тоффоли [47]. Чтобы иметь достаточно шансов выполнить такое вычисление с приемлемой точностью, необходимо, чтобы вероятность ошибки на вентиль Тоффоли была меньше 10^{-9}, а вероятность ошибки запоминающего устройства на время выполнения вентиля — меньше 10^{-12}.

Согласно потоковым уравнениям каскадирования для 7-кубитового кода,[1] для закодированных данных эти частоты возникновения ошибок могут быть достигнуты, если частоты ошибок на уровне отдельных кубитов равны $\epsilon_{\text{store}} \sim \epsilon_{\text{gate}} \sim 10^{-6}$ и если используется три уровня каскадирования, так что размер блока, кодирующего каждый кубит, равен $7^3 = 343$. С учетом дополнительных служебных кубитов, необходимых для выполнения вентилей и (параллельного) исправления ошибок, требуемое машине общее число кубитов будет иметь порядок 10^6.

При достаточно высокой частоте появления ошибок запоминающего устройства каскадирование может стать наиболее эффективной кодирующей процедурой. Но если доминируют ошибки вентилей (и если частота их появления не слишком близка к пороговой), тогда лучше будут работать другие коды. Например, Стин [48] обнаружил, что с использованием кода с размером блока 55, способного исправить пять ошибок, эта задача факторизации может быть решена квантовым компьютером при наличии $4 \cdot 10^5$ кубитов и частоте появления ошибок вентиля порядка 10^{-5}. При более низких частотах появления ошибок можно применять коды, более эффективно использующие объем памяти путем кодирования множества кубитов в одном блоке [17].

Несомненно, квантовый компьютер, содержащий около миллиона кубитов и имеющий отнесенную к одному вентилю частоту появления ошибки около одной на миллион, был бы очень мощным и ценным устройством (при достаточной скорости обработки). Конечно, с точки зрения текущего состояния технологии [49–52], эти числа выглядят удручающими. Но даже машина, удовлетворяющая гораздо менее жестким техническим требованиям, может быть очень полезной [53]. Во-первых, помимо факторизации квантовые компьютеры способны на многое другое, и некоторые из этих задач (в частности, квантовое моделирование [54]) можно выполнить с по-

[1] Этот анализ [23] был выполнен скорее для измерения синдрома методом Шора, нежели методом Стина, на который мы ссылались в разделе 5 при обсуждении каскадного кодирования.

мощью менее надежного или меньшего по размеру устройства. Более того, наша оценка порога безошибочности, возможно, по ряду причин слишком консервативна. В частности, она была получена в предположении, что фазовые и амплитудные ошибки в кубитах равновероятны. Располагая более реалистичной моделью, лучше представляющей вероятности ошибок в данном устройстве, к ней можно было бы лучше приспособить схему исправления ошибок и, следовательно, пережить более высокую частоту появления ошибок. Но даже при сформулированных допущениях приведенный анализ отказоустойчивой схемы не вполне точен; при более тонком анализе можно ожидать обнаружение несколько более высокого порога безошибочности, возможно, значительно более высокого. Также значительных усовершенствований можно достичь путем модификации схемы отказоустойчивости, или с помощью новых, более эффективных, способов применения универсального набора отказоустойчивых вентилей или средств выполнения измерений синдрома ошибки. С учетом различных усовершенствований уже не кажется удивительным то, что квантовый компьютер сможет эффективно работать при вероятности появления ошибок в расчете на один вентиль, например, порядка 10^{-3}.[1]

Конечно, частота появления ошибок, скажем, 10^{-5} весьма претенциозна, но, наверное, эта величина не лежит за рамками возможно достижимого в будущем. В любом случае, сейчас мы имеем четкое представление о степени надежности работы полезного квантового компьютера, что само по себе представляет невероятный прогресс по сравнению с тем, что мы имели всего лишь два года назад.

7. Топологические квантовые вычисления

7.1. Эффект Аронова–Бома и правила суперотбора

Теперь, когда мы не сомневаемся в возможности исправления квантовых ошибок, важно взглянуть на эту проблему шире, а именно, попытаться выйти за рамки анализа абстрактных схем и исследовать потенциальные физические условия, в которых можно надежно хранить и обрабатывать квантовую информацию. В частности, мы могли бы рассчитывать на создание *внутренне* отказоустойчивых квантовых вентилей, не требующих активного вмешательства оператора вычислительной машины для защиты ее от шума. Важный шаг в направлении к этой цели был недавно сделан Алексеем Китаевым [25], настоящий раздел основан на его идеях.

[1] Действительно, Залкой были предложены более оптимистичные по сравнению с моими оценки порога безошибочности [24].

Топологические идеи естественным образом возникают при обсуждении коррекции квантовых ошибок и отказоустойчивых вычислений. Топология интересуется «глобальными» свойствами объекта, которые остаются неизменными при его локальной деформации. Основная идея коррекции квантовых ошибок включает хранение и обработку квантовой информации в «глобальной» форме, устойчивой к локальным возмущениям. Конструкция отказоустойчивого вентиля должна позволять ему действовать на эту глобальную информацию так, чтобы направленное им на закодированные данные действие оставалось неизменным даже при некоторой деформации данного вентиля, то есть даже при его неидеальном выполнении.

Пытаясь найти физическую реализацию отказоустойчивых квантовых вычислений, зададимся вопросом: существуют ли системы, в которых физические взаимодействия имеют топологическую природу? Несомненно, топология лежит в основе *эффекта Аронова–Бома*. Если электрон обходит вокруг идеально заэкранированного магнитного соленоида, его волновая функция приобретает фазу $e^{ie\Phi}$, где e — заряд электрона, а Φ — запертый внутри соленоида магнитный поток. Эта фаза Аронова–Бома является топологическим свойством пройденного электроном пути — она зависит лишь от количества оборотов вокруг соленоида и остается неизменной при непрерывной деформации траектории обхода (см. рис. 16). Это заставляет нас подумать о такой реализации квантовых вычислений, в которой закодированная информация могла бы измеряться и обрабатываться с помощью взаимодействий Аронова–Бома — устойчивых к локальным возмущениям взаимодействий топологической природы.

Полезно выразить эти доводы еще раз на языке правил суперотбора. Правило суперотбора, как я использую здесь этот термин, возникает (в теории поля или спиновой системы, определенной в бесконечном пространственном объеме), если гильбертово пространство распадается на взаимно ортогональные секторы, каждый из которых сохраняется под действием любой локальной операции. Пожалуй, самым известным примером является правило суперотбора заряда в квантовой электродинамике. Электрическое поле, создаваемое зарядом, имеет бесконечный радиус действия. Следовательно, никакая локальная операция не в состоянии создать или уничтожить заряд. Действительно, для этого необходимо создать или уничтожить простирающиеся в бесконечность силовые линии электрического поля, но ни одна локальная процедура не может справиться с этой задачей.

Взаимодействие Аронова–Бома также имеет бесконечный радиус действия; по мере обхода вокруг соленоида электрон приобретает фазу Аронова–Бома независимо от их взаимной удаленности. Поэтому можно сказать, что ни одна локальная операция не может уничтожить заряд, при-

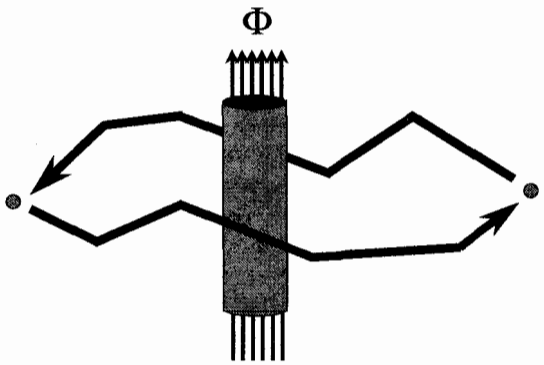


Рис. 16. Топологические взаимодействия. Фаза Аароновa–Бома, приобретаемая электроном при обходе трубки потока, остается неизменной при непрерывной деформации его пути

нимающий участие в явлении Аароновa–Бома. Если мы рассмотрим два несущих такие заряды объекта, находящихся на значительном расстоянии друг от друга и хорошо изолированных от других заряженных объектов, то любой процесс, изменяющий заряд на одном из этих двух объектов, должен был бы действовать когерентно во всей содержащей их области. Таким образом, в присутствии локальных возмущений заряды достаточно устойчивы; мы можем ударить по частице молотком или повредить ее любым другим способом, но переносимый ею заряд мы этим не изменим.

Следуя Китаеву [25], мы можем представить себе *топологический квантовый компьютер* — устройство, в котором квантовая информация кодируется в квантовых числах, переносимых квазичастицами, лежащими на двумерной плоскости и влияющими друг на друга посредством дальнодействующего взаимодействия Аароновa–Бома. При нулевой температуре случайная перестановка квантовых чисел между квазичастицами (ошибка) возникает только из-за явления квантового туннелирования, влекущего за собой виртуальную перестановку заряженных объектов. Амплитуда такого процесса имеет порядок e^{-mL}, где m — масса самого легкого заряженного объекта (в естественных единицах), а L — расстояние между двумя квазичастицами. Если квазичастицы удерживаются на достаточно большом расстоянии друг от друга, вероятность появления ошибки, поражающей закодированную информацию, будет чрезвычайно низкой. При конечной температуре T появляется дополнительный источник ошибок, обусловленный

неизбежной при $T > 0$ генерацией плазмы заряженных частиц с плотностью, пропорциональной фактору Больцмана $e^{-\Delta/T}$, где Δ — массовая щель (не обязательно равная «массе кривизны» m). Иногда одна из частиц плазмы может проскользнуть незамеченной между двумя нашими частицами — носителями данных, что приведет к перестановке зарядов и, следовательно, к ошибке. Итак, для достижения приемлемо низкой частоты появления ошибок необходимо поддерживать уровень температуры значительно ниже щели Δ (в противном случае нам пришлось бы тщательно контролировать термически возбужденную плазму).

7.2. Дробный квантовый эффект Холла (и не только)

Для того чтобы наше устройство было способно выполнять интересные вычисления, применяемые им явления Аронова–Бома должны быть *неабелевыми*. Только в этом случае можно построить сложные унитарные преобразования, выполняя множество следующих друг за другом перестановок частиц. Такие неабелевы эффекты Аронова–Бома могут возникнуть в системах с неабелевыми калибровочными полями. Природа оказалась весьма благосклонной — она снабдила нас некоторыми фундаментальными неабелевыми калибровочными полями, но, к сожалению, не слишком большим количеством, и, похоже, ни одно из них не подходит для практических квантовых вычислений. В таком случае нам остается надеяться на то, что подходящие для реализации идеи Китаева неабелевы эффекты Аронова–Бома могут возникать как сложные кооперативные явления в (двумерных электронных или спиновых) системах, в которых существуют лишь короткодействующие фундаментальные взаимодействия.

Тот факт, что дальнодействующие явления Аронова–Бома могут возникать в таких системах (как это показали наблюдения дробного квантового эффекта Холла), является одним из наиболее замечательных открытий последних десятилетий. Электронная система в режиме квантового эффекта Холла настолько фрустрирована, что ее основное состояние представляет собой в высшей степени запутанное состояние с простирающимися на большие расстояния сильными квантовыми корреляциями. Следовательно, когда одна квазичастица обходит вокруг другой, даже если они расположены на большом расстоянии друг от друга, многоэлектронная волновая функция приобретает нетривиальную фазу Берри (такую как $e^{2\pi i/3}$). Эта фаза Берри во всех ее наблюдаемых проявлениях неотличима от фазы Аронова–Бома, происходящей от фундаментального калибровочного поля, а ее экспериментальные последствия впечатляющи [55].

Наблюдаемые в режиме квантового эффекта Холла фазы Берри — абелевы (хотя имеются некоторые серьезные признаки того, что при соответ-

ствующих условиях могут возникать неабелевы фазы Берри [56,57]) и, следовательно, не особенно интересны с точки зрения квантовых вычислений. Однако Китаев [25] описал семейство простых спиновых систем с локальными взаимодействиями, в которых возможно существование квазичастиц с неабелевыми фазами Берри. (Гамильтониан системы настолько фрустрирует спины, что основное состояние представляет собой чрезвычайно запутанное состояние с бесконечным радиусом квантовых корреляций.) Эти модели настолько просты (хотя, к сожалению, они требуют четырехчастичных взаимодействий), что можно даже представить такой материал, который достаточно хорошо описывается одной из моделей Китаева. Основные топологические свойства модели относительно безразличны к точным микроскопическим деталям, поэтому технологическая проблема «подгонки» параметров материала, возможно, не будет слишком сложной. Более того, если бы можно было управлять процессом переноса отдельных квазичастиц (возможно, при помощи подходящего магнитного пинцета), тогда эта система могла бы функционировать как отказоустойчивый квантовый компьютер.

Модель Китаева представляет собой систему спинов, располагающихся на ребрах квадратной решетки. Гамильтониан выражается в виде суммы двух взаимно коммутирующих четырехчастичных операторов, один из которых соответствует узлам решетки, а второй — плакетам (см. рис. 17). Поскольку они взаимно коммутируют, гамильтониан несложно диагонализовать путем диагонализации каждого его слагаемого по отдельности. Операторы на узлах напоминают локальные калибровочные симметрии (действующие независимо на каждом узле), а состояние, минимизирующее эти слагаемые, инвариантно относительно локальной симметрии, подобно физическим состояниям, подчиняющимся закону Гаусса в калибровочной теории. Операторы на плакетах подобны операторам «магнитного потока» в калибровочной теории, а эти слагаемые минимизируются, когда магнитный поток всюду обращается в нуль. Возбуждениями в такой системе являются состояния, в которых закон Гаусса нарушен на изолированных узлах (эти точки представляют собой «электрически заряженные» квазичастицы), и состояния, в которых магнитные потоки не равны нулю на изолированных плакетах (квазичастицы магнитного потока). Квантовое запутывание основного состояния таково, что нетривиальная фаза Берри, связанная с обходом заряда вокруг потока, идентична фазе Аароновa–Бома в аналогичной калибровочной теории.

Эти явления Аароновa–Бома стабильны даже при деформации гамильтониана данной модели. Действительно, если деформация достаточно мала, мы можем изучать ее влияние, используя теорию возмущений. Но пока воз-

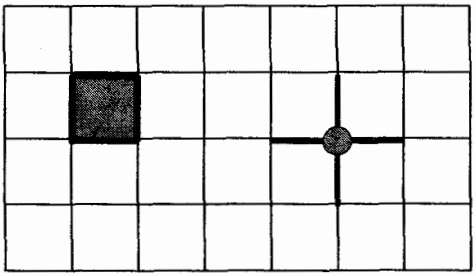


Рис. 17. Спиновая модель Китаева. Спины располагаются на ребрах решетки. Взаимодействуют четверки спинов, встречающихся на одном узле или окружающих один плакет

мущения пространственно локальны, топологические эффекты устойчивы, так как результаты теории возмущений определяются лишь суммой локализованных воздействий. Любая деформация модели, разрушающая дальнодействующие топологические взаимодействия, должна иметь непертурбативный характер.

Можно предвидеть два типа непертурбативных эффектов [58]. Основным состоянием теории мог бы оказаться «конденсат потока» с бесконечным количеством магнитных возбуждений. В таком случае возникло бы дальнодействующее взаимодействие притяжения между заряженными частицами и их античастицами. Разделение зарядов стало бы невозможным, и не возникло бы никаких дальнодействующих эффектов. В калибровочной теории это явление назвали бы *электрическим конфайнментом*. И наоборот, в основном состоянии мог бы образоваться конденсат электрически заряженных квазичастиц. Тогда возник бы конфайнмент магнитных возбуждений, и вновь дальнодействующие эффекты Аронова – Бома были бы разрушены. В калибровочной теории это назвали бы явлением Хиггса (или магнитным конфайнментом).

Таким образом, деформируя гамильтониан Китаева, мы можем ожидать, что в конечном счете столкнемся с границей раздела фаз, за пределами которой возникает электрический конфайнмент или явление Хиггса. Размер заключенной в эти границы области определяет, как именно должен быть изготовлен материал, чтобы он вел себя по предписанию Китаева. Чрезвычайно важный вопрос для разработчика материала состоит в следующем: смогут ли искусно подобранные *двухчастичные* взаимодействия так фрустрировать спиновую систему, чтобы в ней образовалось сильно запутанное

основное состояние, а между квазичастицами возбуждений возникли неабелевы взаимодействия Ааронова–Бома?

Дробный квантовый эффект Холла и модели Китаева преподнесли впечатляющий урок. Мы обнаружили калибровочные эффекты, возникающие как коллективные явления в системах с одними лишь короткодействующими взаимодействиями. Возможно, стоит подумать о том, что известные в природе калибровочные симметрии могут иметь подобное происхождение.

7.3. Топологические взаимодействия

Как уже отмечалось, в спиновых моделях Китаева существует два типа зарядов, переносимых локализованными квазичастицами, которые можно назвать «электрическими» и «магнитными» зарядами. В модели простейшего типа «магнитным потокам», переносимым частицами, можно сопоставить элементы некоторой конечной группы G, а «электрическим зарядам» — неприводимые представления[1] группы G. Если частица с зарядом, соответствующим неприводимому представлению $D^{(\nu)}$, квантовые числа которой закодированы во внутренней волновой функции $|\psi^{(\nu)}\rangle$, обходит вокруг потока, обозначенного групповым элементом $u \in G$, то ее волновая функция преобразуется по правилу

$$|\psi^{(\nu)}\rangle \to D^{(\nu)}(u)|\psi^{(\nu)}\rangle. \tag{37}$$

Используя это взаимодействие, мы можем *измерить* магнитный поток путем рассеяния на нем подходящей заряженной частицы [59]. Например, мы могли бы сконструировать изображенный на рисунке 18 флюксинтерферометр Маха–Цендера, чувствительный к относительной фазе, приобретаемой заряженными частицами, проходящими по траекториям справа или слева от потока. Если мы подходящим образом сбалансируем интерферометр, то сможем различить, скажем, два значения потока u_1, $u_2 \in G$; поток u_1 будет обнаруживаться появлением частицы в одном плече интерферометра, а поток u_2 — появлением частицы в другом плече. Конечно, сконструированный нами интерферометр не будет безупречным, но, тем не менее, измерение потока может быть отказоустойчивым: при наличии большого количества заряженных налетающих частиц и при многократном повторении измерения мы можем определить поток с достаточно высокой статистической достоверностью.

[1] Также возможно существование «дионов», переносящих заряды обоих типов; классификация переносимого дионом заряда достаточно тонкая, однако нам не потребуется детальное обсуждение свойств этих квазичастиц.

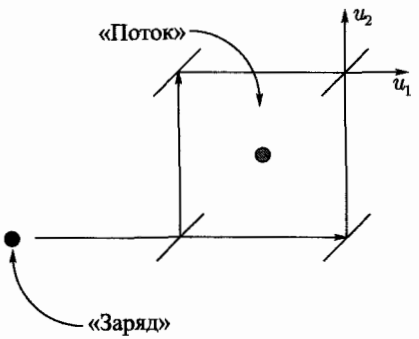


Рис. 18. Схематическое изображение интерферометра Маха–Цендера для измерения потока. Измеряемый поток помещается внутрь. Если поток имеет значение u_1, то пробный заряд появляется в одном плече, а при значении u_2 — в другом

Если два потока u_1 и u_2 принадлежат одному и тому же классу сопряженных элементов группы G, то существует связывающая их симметрия, а вся локальная физика не зависит от значения потока (см. ниже). Следовательно, когерентность суперпозиции потоков

$$a|u_1\rangle + b|u_2\rangle \tag{38}$$

не будет разрушаться локальными взаимодействиями с окружающей средой. Тем не менее, флюкс-интерферометр (действующий многократно) спроецирует флаксон на одно из двух собственных состояний $|u_1\rangle$ (с вероятностью $|a|^2$) или $|u_2\rangle$ (с вероятностью $|b|^2$).

Теперь представим, что два флаксона были тщательно откалиброваны, так что известно, что один из них несет поток u_1, а второй — поток u_2. Осторожно перемещая первый флаксон относительно второго, «переставим» их, как это показано на рисунке 19, после чего вновь выполним измерение потоков. После такого обмена, обход заряженной частицы вокруг правого флаксона топологически эквивалентен следующему обходу, совершаемому до перестановки: сначала вокруг правого флаксона, затем вокруг левого и, наконец, вокруг правого флаксона в противоположном направлении. Мы приходим к выводу, что обмен изменяет квантовые числа флаксонов по правилу

$$|u_1\rangle|u_2\rangle \to |u_2\rangle|u_2^{-1}u_1u_2\rangle, \tag{39}$$

что представляет собой нетривиальное взаимодействие, если два потока не коммутируют [60]. Таким образом, даже в отсутствие любых электрических

зарядов между некоммутирующими потоками существуют собственные интересные взаимодействия Аароно́ва–Бома. Так как обход одного потока вокруг другого может привести к *сопряжению* его значения, два флаксона, несущие на себе сопряженные потоки, должны рассматриваться как *неразличимые* частицы [61]. Перестановка двух таких объектов может изменить их внутренние квантовые числа; эти неразличимые в двумерии частицы, которые подчиняются экзотической неабелевой разновидности статистики, мы будем называть *неабелевыми анионами* [62].

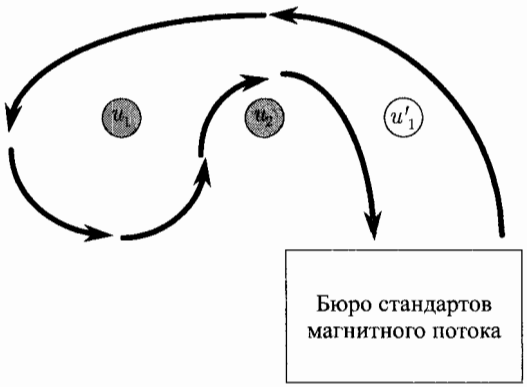


Рис. 19. Обменное взаимодействие потоков. Обозначенный u_1 поток из своего исходного положения (заштрихованное) перемещается в новое (не заштрихованное), а затем заново измеряется. Изображенная траектория заряженной частицы, что обходит исходное положение потока, топологически эквивалентна траектории, охватывающей новое положение; следовательно, значение потока u_1 меняется на $u'_1 = u_2^{-1} u_1 u_2$

Мы будем использовать обменное взаимодействие (39) в качестве фундаментальной логической операции в нашем квантовом компьютере Аароно́ва–Бома. Однако в действительности может оказаться удобным кодирование кубитов в парах флаксонов с тривиальным полным потоком [25]: мы будем рассматривать пары флаксон-антифлаксон вида $|u, u^{-1}\rangle$, но такие, в которых флаксон и антифлаксон удерживаются на достаточном расстоянии друг от друга, чтобы случайный обмен квантовыми числами между ними был маловероятен. Для выполнения логической схемы мы можем протащить одну пару сквозь другую, как это показано на рисунке 20. Так как полный поток, проходящий через середину внешней пары, тривиален, ее состояние не изменяется, но внутренние потоки сопрягаются внешним

потоком:

$$|u_1, u_1^{-1}\rangle|u_2, u_2^{-1}\rangle \to |u_2, u_2^{-1}\rangle|u_2^{-1}u_1u_2, u_2^{-1}u_1^{-1}u_2\rangle; \tag{40}$$

эта операция, очевидно, изоморфна результату перестановки отдельных потоков, описываемому уравнением (39). Использование пар вместо отдельных флаксонов имеет два преимущества. Во-первых, так как каждая пара имеет тривиальный полный поток, они не взаимодействуют, пока одна из них не протаскивается сквозь другую; следовательно, мы можем свободно перемещать пары по устройству, не вызывая никаких нежелательных взаимодействий с удаленными парами. Во-вторых, что более важно, пара может переносить заряд, даже если каждый ее элемент в этом отношении является нейтральным [63, 64]. Заряд пары можно измерить, и эта операция измерения заряда станет основным элементом универсального набора квантовых вентилей. Операцию (40) можно рассматривать как *классический* логический вентиль; она преобразует одно собственное состояние потока в другое. Чтобы выполнять интересные квантовые вычисления, мы должны уметь готовить когерентные суперпозиции собственных состояний потока. Именно это мы можем осуществлять, измеряя заряд пары.

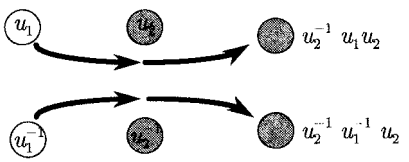


Рис. 20. Взаимодействие «протаскивания». Одна пара потоков протаскивается сквозь другую. Внешний поток не изменяется, но внутренний поток сопрягается внешним

Предположим, что u_0 и $u_1 \in G$ связаны соотношением $u_1 = v^{-1}u_0v$ для некоторой величины $v \in G$. Тогда, если мы рассматриваем собственные состояния потоков $|u_0, u_0^{-1}\rangle$ и $|u_1, u_1^{-1}\rangle$ в качестве состояний вычислительного базиса, эффект протаскивания одной из двух пар сквозь пару $|v, v^{-1}\rangle$ можно интерпретировать как NOT- или X-вентиль:

$$|u_0, u_0^{-1}\rangle \longleftrightarrow |u_1, u_1^{-1}\rangle \tag{41}$$

(см. рис. 21). Теперь предположим, что мы хотим приготовить одно из состояний

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|u_0, u_0^{-1}\rangle \pm |u_1, u_1^{-1}\rangle). \tag{42}$$

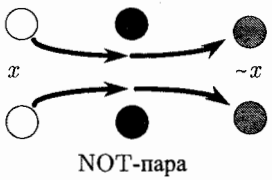


Рис. 21. NOT-вентиль. Протаскивание вычислительной пары потоков сквозь пару NOT обращает значение закодированного бита

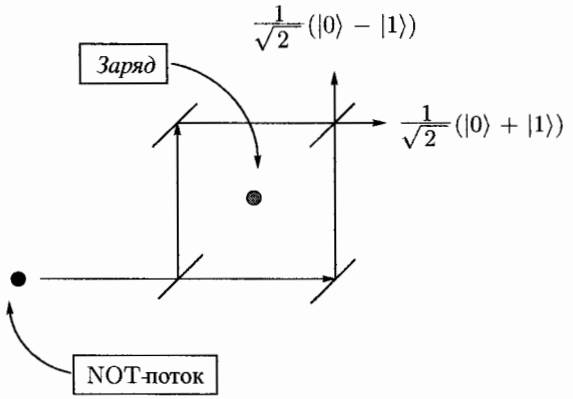


Рис. 22. Схематическое изображение интерферометра Маха – Цендера для измерения заряда. Пара потоков, заряд которой должен быть измерен, помещается внутрь. Если пробный поток NOT появляется в одном плече, то было приготовлено состояние заряда $|+\rangle$; если он появляется в другом плече, то было приготовлено состояние $|-\rangle$

Мы можем спроецировать когерентную суперпозицию $|u_0, u_0^{-1}\rangle$ и $|u_1, u_1^{-1}\rangle$ на базис $\{|\pm\rangle\}$, путем рассеяния флаксона $|v\rangle$ на паре или, дру-

гими словами, оперируя *зарядовым интерферометром*, как это показано на рисунке 22. Когда флаксон $|v\rangle$ движется вокруг пары, он приобретает тривиальную фазу Аронова–Бома, если пара находится в состоянии $|++\rangle$, и нетривиальную фазу -1, если пара находится в состоянии $|--\rangle$. Если интерферометр должным образом сбалансирован, то налетающая частица $|v\rangle$ будет детектироваться в одном плече интерферометра, когда состоянием пары является $|+\rangle$, и в другом, когда состоянием пары является $|-\rangle$. Это пример измерения заряда. Хотя интерферометр не будет безупречен, измерение заряда (как и измерение потока) может быть отказоустойчивым, если оно повторяется достаточное количество раз.

7.4. Универсальные топологические вычисления

Работая с парами флаксонов в качестве состояний вычислительного базиса, мы увидели, как выполняется операция перестановки (или «протаскивания») в (40), как измеряется поток (с использованием предварительно откалиброванных зарядов), и как измеряется заряд (с использованием предварительно откалиброванных потоков). Предположим теперь, что мы можем создать большой запас пар потоков. С помощью локальных процессов можно создать пары, несущие нулевой заряд и тривиальный поток; с точностью до нормировки, состояние такой пары имеет вид

$$|\text{charge zero}\rangle = \sum_u |u, u^{-1}\rangle, \tag{43}$$

где суммирование ведется по всему классу сопряженных элементов группы G. Поскольку при сопряжении любым элементом группы G это состояние остается неизменным, оно имеет тривиальные взаимодействия Аронова–Бома с любым потоком и, следовательно, не имеет детектируемого заряда. После рождения такой пары можно выполнить измерение потока, проецирующее ее состояние на одно из собственных состояний пары потоков $|u, u^{-1}\rangle$. Выполняя множество таких измерений для большого количества пар, мы собираем большой резервуар калиброванных пар потоков, которые при необходимости можно извлекать в процессе вычисления.

Но является ли наш квантовый компьютер универсальным — можем ли мы получить достаточное приближение любого желаемого унитарного преобразования? Чтобы исследовать этот вопрос, вспомним упомянутый в разделе 4.2 результат: для универсальных квантовых вычислений достаточно универсальных *классических* вычислений в совокупности с возможностью *выполнять* однокубитовые вентили X и Z и *измерять* X, Y и Z [17]. Действительно, существуют такие группы G, что операции (40) оказывается

достаточно для универсальных классических вычислений. Мы обнаружили [65], что если $G = A_5$ — группа четных перестановок пяти объектов, то из уравнения (40) можно построить вентиль Тоффоли. Например, в качестве состояний вычислительного базиса можно выбрать

$$u_0 = (125), \quad u_1 = (234); \tag{44}$$

то есть потоки, соответствующие циклам длины три (3-циклам) с одним общим объектом. Тогда вентиль Тоффоли можно построить в общей сложности из 16 элементарных операций «протаскивания»; кроме того, для ускорения выполнения этой операции используется шесть служебных пар. Ни в одной из меньших, чем A_5, групп вентиль Тоффоли не был обнаружен.[1] Так как A_5 также является наименьшей из конечных неразрешимых групп, напрашивается вывод, что неразрешимость — необходимое условие для порожденных сопряжением классических вычислений.[2]

Как уже говорилось, X-вентиль можно осуществить, протаскивая вычислительную пару вихрей сквозь пару с потоком v, таким, что $u_1 = v^{-1}u_0v$; здесь мы выбираем $v = (14)(35)$. Оказывается, Z-вентиль можно построить при помощи шести этапов «протаскивания» и четырех служебных пар. Измерение Z аналогично измерению потока, и мы уже видели, что измерение X можно выполнить путем измерения заряда пары, а именно, используя в зарядовом интерферометре в качестве налетающей частицы v. Остается лишь подтвердить, что мы можем измерить Y. Хотя по данной схеме измерение Y нельзя выполнить точно, оказывается, что вентиль «контролируемое Y» можно построить с помощью 31 этапа «протаскивания» и семи служебных пар. Обращаясь к другому изобретенному Китаевым трюку [67], для выполнения Y-измерения с любой желаемой точностью мы можем использовать вентиль «контролируемое Y» повторно.[3] Следовательно, мы построили набор универсальных вентилей, используя лишь взаимодействия Аароно́ва–Бома потоков и зарядов; мы располагаем отказоустойчивым универсальным квантовым компьютером.

[1] Ранее Китаев сообщал, что универсальные классические вычисления возможны для $G = S_5$.

[2] Конечная группа *неразрешима*, если она имеет нетривиальную подгруппу, коммутант которой совпадает с ней самой. (Коммутантом группы называется множество всех возможных произведений коммутаторов $aba^{-1}b^{-1}$ элементов рассматриваемой группы. Имеет место следующий критерий: если группа G некоммутативна и не имеет нетривиальных нормальных подгрупп, то она неразрешима. — *Прим. ред.*) Баррингтон [66] нашел признак разделения групп на разрешимые и неразрешимые по вычислительной сложности группового умножения.

[3] Фактически, измерение Y (который имеет собственные значения $\pm i$) с использованием вентиля контролируемое Y не работает, поскольку метод Китаева не различает связанные комплексным сопряжением собственные значения. То, что мы действительно построили — это вентиль контролируемое ωY, где $\omega = e^{2\pi i/3}$.

К сожалению, спиновая модель, на которой основана данная конструкция, не так проста. Так как группа A_5 имеет порядок 60, реализующая данную схему спиновая модель Китаева имеет 60-компонентный спин, расположенный на каждом (!) ребре решетки. Остается надеяться, что будет обнаружена более простая реализация вычислений Аронова–Бома.

7.5. Является ли природа отказоустойчивой?

Открытие коррекции квантовых ошибок и отказоустойчивости настолько изменило наше представление о квантовой информации, что впору задаться вопросом об их потенциальном значении для фундаментальной физики. Действительно, основные вопросы, имеющие отношение к потере квантовой информации, озадачивали физиков на продолжении двадцати лет.

В 1975 году, Стивен Хокинг [68] доказал, что квантовая информация неизбежно теряется при образовании черной дыры и ее последующем полном испарении. Суть доказательства предельно проста: из-за сильно искаженной причинной структуры пространства-времени черных дыр испускаемое излучение находится фактически на *том же* временном срезе, что и исчезающее за горизонтом событий коллапсирующее вещество. Если квантовая информация, первоначально закодированная в коллапсирующем веществе, должна в конечном счете возродиться в информации, закодированной в микросостоянии излучения, то она должна находиться в двух местах одновременно. Другими словами, квантовая информация должна быть *клонирована*, что, как известно, при обычных допущениях квантовой теории невозможно [69, 70]. Хокинг делает вывод, что не все физические процессы могут управляться унитарной временной эволюцией; законы квантовой теории нуждаются в пересмотре.

Эти аргументы убедительны, но многие физики им не доверяют. Возможно, одна из причин для скептицизма состоит том, что для природы выглядит странным допускать потерю даже маленького бита информации [71]. Если процессы с участием черных дыр могут разрушать информацию, то можно ожидать, что потеря информации неизбежна на масштабе порядка планковской длины $\left(G\hbar/c^3\right)^{1/2} \sim 10^{-33}$ см, на котором виртуальные черные дыры непрерывно возникают как квантовые флуктуации. Становится сложно понять, почему квантовая информация может быть так легко разрушена на планковском масштабе, но так хорошо сохраняется на больших расстояниях, которые мы можем исследовать экспериментально, — в конце концов, нарушения законов квантовой механики ни разу не наблюдались.

Наше сегодняшнее понимание отказоустойчивых квантовых вычислений дает полезный и потенциально продуктивный способ разобраться

в данной проблеме. В спиновых моделях Китаева мы можем представить, что разрушающие квантовую информацию локальные процессы достаточно просты. Тем не менее, если бы нам потребовалось проследить эволюцию системы с более грубым разрешением, отслеживая лишь информацию, закодированную в зарядах пространственно разделенных квазичастиц, мы с поразительной точностью наблюдали бы унитарную эволюцию; мы не обнаружили бы ни одного признака шума, происходящего ниже этого уровня.[1]

Таким образом, весьма приятно полагать, что Природа внесла в свою структуру отказоустойчивость, скрыв от нас квантовый шум на масштабе Планка. Открытие того, что квантовые системы можно стабилизировать с помощью соответствующих методов кодирования, заставляет нас задаться вопросом: является ли природа отказоустойчивой? Если да, то квантовая механика может доминировать (с превосходной точностью) на промежуточных масштабах длин, но спотыкаться как на планковском масштабе (где высока частота появления «ошибок»), так и на макроскопическом масштабе (где стремительна декогерентизация).

Настоящая работа была частично поддержана Управлением перспективного планирования оборонных научно-исследовательских работ (DARPA), грант DAAH04-96-1-0386, управляемым Военным исследовательским центром, и Министерством энергетики, грант DE-FG03-92-ER40701. Я благодарен полезным беседам и переписке с Доритом Аароновым, Дэвидом Бэкманом, Джоном Кортизом, Эриком Деннисом, Дэвидом Дивинченцо, Джара Эвслином, Крисом Фуксом, Шамом Какаде, Алексеем Китаевым, Мэнни Ниллом, Раймондом Лафламмом, Эндрю Ландалом, Сетом Ллойдом, Майклом Нильсеном, Уолтом Огберном, Питером Шором, Эндрю Стином и Кристофом Залкой. Я отдельно благодарю Дэниэла Готтесмана за множество плодотворных дискуссий об отказоустойчивых квантовых вычислениях.

Литература

[1] R. P. Feynman, Simulating physics with computers, *Int. J. Theor. Phys.*, **21**, 467–482 (1982); перевод: Р. Фейнман, Моделирование физики на компьютерах. *Квантовый компьютер и квантовые вычисления*. — Ижевск, РХД (1999).

[1] Аналогичный язык можно было бы использовать для характеристики осуществления каскадного кода: ошибки редки, когда мы исследуем квантовую информацию при слабом разрешении, но появляются намного чаще, если мы рассматриваем кодовый блок на более глубоких уровнях каскадирования.

[2] D. Deutsch, Quantum theory, the Church-Turing principle and the universal quantum computer, *Proc. Roy. Soc. Lond.*, A **400**, 97–117 (1985); перевод: Д. Дойч, Квантовая теория, принцип Черча–Тьюринга и универсальный квантовый компьютер. *Квантовый компьютер и квантовые вычисления.* — Ижевск, РХД (1999).

[3] P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings of the 35th Annual Symposium on Fundamentals of Computer Science* (Los Alamitos, CA, IEEE Press, 1994), pp. 124–134; Расширенная версия: SIAM J. Comp. **26**, 1484–1509 (1997), online preprint quant-ph/9508027; перевод: П. Шор, Полиномиальные по времени алгоритмы разложения числа на простые множители и нахождения дискретного логарифма для квантовых компьютеров. *Квантовый компьютер и квантовые вычисления.* — Ижевск, РХД (1999).

[4] R. Landauer, Is quantum mechanics useful? *Phil. Tran. R. Soc. Lond.*, **353**, 367–376 (1995).

[5] R. Landauer, The physical nature of information, *Phys. Lett.*, A **217**, 188–193 (1996).

[6] R. Landauer, Is quantum mechanically coherent computation useful? In *Proc. Drexel-4 Symposium on Quantum Nonintegrability-Quantum-Classical Correspondence*, Philadelphia, PA, 8 September 1994, ed. D. H. Feng and B.-L. Hu (Boston, International Press, 1997).

[7] W.G. Unruh, Maintaining coherence in quantum computers, *Phys. Rev.*, A **51**, 992–997 (1995).

[8] S. Haroch and J.-M. Raimond, Quantum computing: dream or nightmare? *Phys. Today*, **49** (8), 51–52 (1996).

[9] W.H. Zurek, Decoherence and the transition from quantum to classical, *Phys. Today*, **44**, 36–44 (1991).

[10] P. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev.*, A **52**, R2493–R2496 (1995).

[11] A.M. Steane, Error correcting codes in quantum theory, *Phys. Rev. Lett.*, **77**, 793–797 (1996).

[12] A.M. Steane, Multiparticle interference and quantum error correction, *Proc. Roy. Soc. Lond.*, A **452**, 2551–2577 (1996).

[13] J. von Neumann, Probabilistic logics and synthesis of reliable organisms from unreliable components, in *Automata Studies*, ed. C. E. Shannon and J. McCarthy (Princeton, Princeton University Press, 1956).

[14] P. Gács, Reliable computation with cellular automata, *J. Comp. Sys. Sci*, **32**, 15–78 (1986).

[15] P. Shor, Fault-tolerant quantum computation, in *Proceedings of the Symposium on the Foundations of Computer Science*, Los Alamitos, CA, IEEE Computer Society Press, 1996, pp. 56–65 (online preprint http://lanl.arxiv.org/abs/quant-ph/9605011, 1996).

[16] A.M. Steane, Active stabilization, quantum computation and quantum state synthesis, *Phys. Rev. Lett.*, **78**, 2252–2255 (1997).

[17] D. Gottesman, A theory of fault-tolerant quantum computation, *Phys. Rev.* A **57**, 127–137 (1998) (online preprint http://lanl.arxiv.org/abs/quant-ph/9702029, 1997).

[18] E. Knill and R. Laflamme, Concatenated quantum codes, Techn. Report LAOR-96-2808. (online preprint http://lanl.arxiv.org/abs/quant-ph/9608012, 1996).

[19] E. Knill, R. Laflamme, and W. Zurek, Accuracy threshold for quantum computation, (online preprint http://lanl.arxiv.org/abs/quant-ph/9610011, 1996).

[20] E. Knill, R. Laflamme, and W. Zurek, Resilient quantum computation: error models and thresholds, *Proc. Roy. Soc. Lond.* A **454**, 365–384 (1998) (online preprint http://lanl.arxiv.org/abs/quant-ph/9702058, 1997).

[21] D. Aharonov and M. Ben-Or, Fault tolerant quantum computation with constant error. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing* ACM. New York, 1998, p.176. (online preprint http://lanl.arxiv.org/abs/quant-ph/9611025, 1996).

[22] А.Ю. Китаев, Квантовые вычисления: алгоритмы и исправления ошибок, *Успехи мат. наук*, **52**, стр. 53–112 (1997).

[23] J. Preskill, Reliable quantum computers, *Proc. Roy. Soc. London* A, **454**, pp. 385–410 (1998) (online preprint http://lanl.arxiv.org/abs/quant-ph/9705031, 1997); перевод: Дж. Прескилл, Надежные квантовые компьютеры (в этом издании).

[24] C. Zalka, Threshold estimate for fault tolerant quantum computing (online preprint http://lanl.arxiv.org/abs/quant-ph/9612028, 1996).

[25] A.Yu. Kitaev, Fault-tolerant quantum computation by anyons, *Annals of Phys*, **303**, 2–30 (2003), (online preprint http://lanl.arxiv.org/abs/quant-ph/9707021, 1997).

[26] F.J.MacWilliams, and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, (New York, North-Holland Publishing Company, 1977); перевод: Ф.Дж. Мак-Вильямс, Н.Дж. Слоэн, *Теория кодов, исправляющих ошибки*, Связь, М., 1979.

[27] E. Knill and R. Laflamme, A theory of quantum error-correcting codes, *Phys. Rev.* A **55**, 900–911 (1997).

[28] A.R. Calderbank and P.W. Shor, Good quantum error-correcting codes exist, *Phys. Rev.*, A **54**, 1098–1105 (1996).

[29] D. Gottesman, Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev.*, A **54**, 1862–1868 (1996).

[30] A.R. Calderbank, E.M. Rains, P.W. Shor, and N.J.A. Sloane, Quantum error correction and orthogonal geometry, *Phys. Rev. Lett.*, **78**, 405–408 (1997).

[31] A.R. Calderbank, E.M. Rains, P.W. Shor, and N.J.A. Sloane, Quantum error correction via codes over GF(4). IEEE Trans. Inf. Theory, 44(4), pp. 1369–1387 (online preprint http://lanl.arxiv.org/abs/quant-ph/9608006, 1996).

[32] J. Evslin, S. Kakade, and J. Preskill, unpublished (1996).

[33] D. DiVincenzo and P. Shor, Fault-tolerant error correction with efficient quantum codes. *Phys. Rev. Lett.*, **77**, 3260–3263 (1996).

[34] A.Yu. Kitaev, Quantum error correction with imperfect gates, in *Proceedings of the Third International Conference on Quantum Communication and Measurement*, Ed O. Hirota, A.S. Holevo, and C.M. Caves, pp 181–188 (New York, Plenum, 1997).

[35] D. Gottesman, Stabilizer codes and quantum error correction. Ph.D. thesis, California Institute of Technology (online preprint http://lanl.arxiv.org/abs/quant-ph/9705052, 1997).

[36] C. Bennett, D. DiVincenzo, J. Smolin, and W. Wootters, Mixed state entanglement and quantum error correction, *Phys. Rev.*, A **54**, 3824–3851 (1996).

[37] R. Laflamme, C. Miquel, J.P. Paz, and W. Zurek, Perfect quantum error correction code, *Phys. Rev. Lett.*, **77**, 198–201 (1996).

[38] D. Gottesman and J. Preskill, unpublished (1997).

[39] D. Gottesman, J. Evslin, S. Kakade, and J. Preskill, unpublished (1996).

[40] K. Obenland and A.M. Despain, Simulation of factoring on a quantum computer architecture, in *Proceedings of the 4th Workshop on Physics and Computation*, Boston, November 22–24, 1996, (Boston, New England Complex Systems Institute, 1996).

[41] K. Obenland, and A.M. Despain, Impact of errors on a quantum computer architecture, Technical Report, Information Science Institute, University of Southhern California, Oct 1, 1996; (online preprint http://www.isi.eu/acal/quantum/quantumoperrors.ps, 1996).

[42] C. Miquel, J.P. Paz, and W.H. Zurek, Quantum computation with phase drift errors, *Phys. Rev. Lett.*, **78**, 3971–3974 (1997); (online preprint http://lanl.arxiv.org/abs/quant-ph/9704003, 1997).

[43] J.I. Cirac and P. Zoller, Quantum computations with cold trapped ions, *Phys. Rev. Lett.*, **74**, 4091–4094 (1995).

[44] M.B. Plenio and P.L. Knight, Decoherence limits to quantum computation using trapped ions, *Proc. Roy. Soc. Lond.*, A **453**, 2017–2041 (1997).

[45] M. Grassl, Th. Beth, and T. Pellizzari, Codes for the quantum erasure channel, *Phys. Rev.*, A **56**, 33–38 (1997).

[46] A.K. Lenstra, J. Cowie, M. Elkenbracht–Huizing, W. Furmanski, P.L. Montgomery, D. Weber, J. Zayer, RSA factoring-by-web: the world-wide status (online document http://www.npac.syr.edu/factoring/status.html, 1996).

[47] D. Beckman, A. Chari, S. Devabhaktuni, and J. Preskill, Efficient networks for quantum factoring, *Phys. Rev.*, A **54**, 1034–1063 (1996).

[48] A.M. Steane, Space, time, parallelism and noise requirements for reliable quantum computing *Fortsch. Phys.*, **46**, 443–458 (1998); (online preprint http://lanl.arxiv.org/abs/quant-ph/9708021, 1997).

[49] C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano, and D.J. Wineland, Demonstration of a fundamental quantum logic gate, *Phys. Rev. Lett.*, **75**, 4714–4717 (1995).

[50] Q.A. Turchette, C.J. Hood, W. Lange, H. Mabuchi, and H.J. Kimble, Measurement of conditional phase shifts for quantum logic. *Phys. Rev. Lett.*, **75**, 4710–4713 (1995).

[51] D.G. Cory, A.F. Fahmy, and T.F. Havel, Nuclear magnetic resonance spectroscopy: an experimentally accessible paradigm for quantum computing. *In Proceedings of the 4th Workshop on Physics and Computation*, T. Toffoli, M. Biafore, and J. Leao (eds.), Boston, New England Complex Systems Institute, pp. 87–91 (1996).

[52] N. Gershenfeld and I. Chuang, Bulk spin resonance quantum computation. *Science*, **275**, 350–356 (1997).

[53] J. Preskill, Quantum computing: pro and con *Proc. Roy. Soc. Lond.* A **454**, 469–486 (1998) (online preprint http://lanl.arxiv.org/abs/quant-ph/9705032, 1997); перевод в *Квантовые вычисления: за и против.* — РХД, Ижевск (1999).

[54] S. Lloyd, Universal quantum simulators, *Science*, **273**, 1073–1078 (1996); correction in *Science* **279**, 1113–1117 (1998).

[55] R. Prange and S. Girvin (eds.), *The Quantum Hall Effect*, (New York, Springer-Verlag, 1987); перевод *Квантовый эффект Холла*, под ред. Р. Пренджа и С. Гирвина. — М.: Мир (1989).

[56] N. Read and E. Rezayi, Quasiholes and fermionic zero modes of paired fraction quantum Hall states: the mechanism for nonabelian statistics, *Phys. Rev.* B, **54**, 16864–16887 (1996); (online preprint http://lanl.arxiv.org/abs/cond-mat/9609079, 1996).

[57] C. Nayak and F. Wilczek, $2n$ quasihole states realize 2^{n-1}-dimensional spinor braiding statistics in paired quantum Hall states, *Nucl. Phys.* B, **479**, 529–553 (1996); (online preprint http://lanl.arxiv.org/abs/cond-mat/9605145, 1996).

[58] G. t'Hooft, On the phase transition toward permanent quark confinement, *Nucl. Phys.*, B **138**, 1–25 (1978).

[59] M. Alford, S. Coleman, and J. March-Russell, Disentangling nonabelian discrete quantum hair, *Nucl. Phys.*, B **351**, 735–748 (1991).

[60] F.A. Bais, Flux metamorphosis, *Nucl. Phys.*, B **170**, 32–43 (1980).

[61] H.-K. Lo and J. Preskill, Nonabelian vortices and nonabelian statistics, *Phys. Rev.*, D **48**, 4821–4834 (1993)

[62] G. Moore and N. Read, Nonabelions in the fractional quantum Hall effect, *Nucl. Phys.*, B **360**, 362–396 (1991).

[63] M.G. Alford, K. Benson, S. Coleman, J. March-Russell, and F. Wilczek, Interactions and excitations of nonabelian vortices, *Phys. Rev. Lett.*, **64**, 1632–1635 (1990).

[64] J. Preskill and L.M. Krauss, Local discrete symmetry and quantum mechanical hair, *Nucl. Phys.*, B **341**, 50–100 (1990).

[65] R.W. Ogburn and H. Preskill, Topological quantum computation, in *Lect. Notes in Comp. Sci.* C.P. Williams (ed.) **1509**, 341–356, Springer–Verlag (1999).

[66] D.A. Barrington, Bounded width polynomial size branching programs recognoze exactly those languages in NC^1, *J. Comp. Sys. Sci.*, **38**, 150–164 (1989).

[67] A.Yu. Kitaev, Quantum measurements and the abelian stabilizer problem (online preprint http://lanl.arxiv.org/abs/quant-ph/9511026, 1995).

[68] S.W. Hawking, Breakdown of predictability in gravitational collapse, *Phys. Rev.*, D **14**, 2460–2473 (1976).

[69] D. Dieks, Communication by electron-paramagnetic-resonance devices. *Phys. Lett.*, A **92**, 271–272 (1982).

[70] W.K. Wootters and W.H. Zurek, A single quantum cannot be cloned, *Nature* **299**, 802–803 (1982); *Nature* **304**, 188-189 (1983).

[71] T. Banks, L. Susskind, and M.E. Peskin, Difficulties for the evolution of pure states into mixed states, *Nucl. Phys.*, B **244**, 125–134 (1984).

Интересующие Вас книги нашего издательства можно заказать почтой или электронной почтой:

subscribe@rcd.ru

Внимание: дешевле и быстрее всего книги можно приобрести через наш Интернет-магазин:

http://shop.rcd.ru

Книги также можно приобрести:

1. Москва, ИМАШ, ул. Бардина, д. 4, корп. 3, к. 415,
   тел.: (499) 135–54–37, (495) 641–69–38
2. МГУ им. Ломоносова (ГЗ, 1 этаж)
3. Магазины:

   Москва: «Дом научно-технической книги» (Ленинский пр., 40)

   «Московский дом книги» (ул. Новый Арбат, 8)

   Книжный магазин «ФИЗМАТКНИГА» (г. Долгопрудный, Новый корпус МФТИ, 1 этаж, тел. 409–93–28)

   С.-Пб.: «С.-Пб. дом книги» (Невский пр., 28)

Джон Прескилл

Квантовая информация и квантовые вычисления

Том 2

Дизайнер В. А. Толстолуцкая
Технический редактор А. В. Широбоков
Компьютерная верстка А. В. Моторин
Корректор О. З. Логунова

Подписано в печать 23.11.2011. Формат $60 \times 84^{1}/_{16}$.
Печать офсетная. Усл. печ. л. 18,14. Уч. изд. л. 19,45.
Гарнитура Таймс. Бумага офсетная № 1. Заказ № 11-55.

АНО «Ижевский институт компьютерных исследований»
426034, г. Ижевск, ул. Университетская, 1.
http://shop.rcd.ru E-mail: mail@rcd.ru Тел./факс: (+73412) 500–295

Дж. Прескилл — известный физик-теоретик, профессор теоретической физики отделения физики, математики и астрономии Калифорнийского технологического института (Калтех). Область научных интересов — физика элементарных частиц и космология, топологические дефекты, непертурбативные методы квантовой теории поля, квантовые аспекты ранней Вселенной и черных дыр. В середине 90-х годов увлекся теорией квантовой информации, квантовых вычислений и кодирования. В настоящее время — один из ведущих специалистов в этой области. Руководитель Института квантовых вычислений, а также Центра физики информации при Калтехе.

ISBN 978-5-4344-0030-5