

ГЛАВА 1

ОСНОВНЫЕ ПОНЯТИЯ КЛАССИЧЕСКОЙ ТЕОРИИ ИНФОРМАЦИИ

§ 1.1. Энтропия случайной величины и сжатие данных

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

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

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

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

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

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

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

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

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

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

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

$$p_w = p_{x_1} \cdot \dots \cdot p_{x_n} = 2^{-n\left(-\frac{1}{n} \sum_{i=1}^n \log p_{x_i}\right)} \quad (1.4)$$

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

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

Непосредственно устанавливаются следующие свойства типичных слов:

- 1) существует не более $2^{n(H(X)+\delta)}$ типичных слов;
- 2) для достаточно больших n существует, по крайней мере, $(1-\varepsilon)2^{n(H(X)-\delta)}$ типичных слов;
- 3) множество нетипичных слов имеет вероятность $\leq \varepsilon$.

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

Задача 2. Докажите последнее утверждение.

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

§ 1.2. Пропускная способность канала с шумом

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

$$I(X; Y) = H(X) - H(X | Y), \quad (1.6)$$

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

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

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

$$I(X; Y) = H(X) + H(Y) - H(X, Y) = H(Y) - H(Y | X), \quad (1.7)$$

где в последней формуле уже $H(Y)$ может быть интерпретирована, как информационное содержание выхода, а $H(Y | X)$ как

его бесполезная составляющая, обусловленная *шумом*. Взаимная информация всегда неотрицательна: тот факт, что $H(X) \geq H(X | Y)$ легко вытекает из вогнутости функции $-x \log x$ (задача 3).

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

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

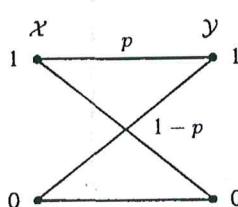


Рис. 1. Двоичный симметричный канал

каналом без памяти. Пропускная способность такого канала определяется как

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

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

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

$$h(p) = -p \log p - (1-p) \log(1-p), \quad (1.9)$$

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

$$C = 1 - h(p), \quad (1.10)$$

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

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

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

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

ЛЕММА.

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

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

$$H(X_1, \dots, X_n) = \sum_{i=1}^n H(X_i | X_1, \dots, X_{i-1}), \quad (1.12)$$

14 Гл. 1. ОСНОВНЫЕ ПОНЯТИЯ КЛАССИЧЕСКОЙ ТЕОРИИ ИНФОРМАЦИИ

которое легко доказать по индукции, используя формулу:

$$H(X, Y) = H(X) + H(Y | X). \quad (1.13)$$

Тогда взаимная информация равна

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

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

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

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

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

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

$$P_e(W, V) = \max_{1 \leq j \leq N} (1 - p(V^{(j)} | w^{(j)})), \quad (1.14)$$

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

$$\bar{P}_e(W, V) = \frac{1}{N} \sum_{i=1}^N (1 - p(V^{(j)} | w^{(j)})) \leq P_e(W, V), \quad (1.15)$$

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

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

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

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

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

где

$$\| a_{ji} \|_1 = \sup_{p_i} \frac{\sum_j |\sum_i a_{ji} p_i|}{\sum_i |p_i|}. \quad (1.17)$$

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

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

$$H(X | Y) \leq h(p_e) + p_e \log(|\mathcal{X}| - 1) \leq 1 + p_e \log |\mathcal{X}|. \quad (1.18)$$

Доказательство. Пусть E — индикатор ошибки оценивания,

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

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

$$H(E | X, Y) = H(E, X | Y) - H(X | Y) = 0, \quad (1.20)$$

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

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

16 Гл. 1. ОСНОВНЫЕ ПОНЯТИЯ КЛАССИЧЕСКОЙ ТЕОРИИ ИНФОРМАЦИИ

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

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

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

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

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

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

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

$$\begin{aligned} nC &= C_n \geq I(Z; Y^n) = H(Z) - H(Z | Y^n) \geq \\ &\geq \log N - 1 - P\{\widehat{Z}(Y^n) \neq Z\} \log N = \log N - 1 - \overline{P}_e(W, V). \end{aligned}$$

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

$$\begin{aligned} nC &\geq nR - 1 - p_e(n, 2^{nR})nR, \\ \frac{C}{R} &\geq (1 - p_e(n, 2^{nR})) - \frac{1}{nR}, \end{aligned}$$

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

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

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

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

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

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

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

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

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

ГЛАВА 2

СОСТОЯНИЯ И НАБЛЮДАЕМЫЕ

§ 2.1. Соглашения и обозначения

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

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

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

§ 2.2. Квантовые состояния

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

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

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

где s_i — собственные числа, $|e_i\rangle$ — собственные векторы оператора S , показывает, что всякое состояние является смесью не более чем d чистых состояний, где $d = \dim \mathcal{H}$. В квантовом статистическом ансамбле есть два вида стохастичности: во-первых, устрашимая в принципе стохастичность, обусловленная флуктуациями классических параметров процедуры приготовления, и во-вторых, неуничтожимая никакими усилиями квантовая стохастичность, присутствующая в любом чистом состоянии.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

нения спинов равновероятны), описываемое оператором плотности $S = I/2$.

Другим важным примером двухуровневой системы является поляризация поперечного фотона.

§ 2.3. Квантовые наблюдаемые

Во всяком физическом эксперименте присутствуют две основные стадии: приготовление состояния и измерение. Даже если приготавливается чистое квантовое состояние, где нет классической стохастичности, результат измерения в данном ансамбле все равно может быть случаен. Итак, мы измеряем случайную величину, распределение $\mu_S^M(x)$ которой зависит от приготовления ансамбля S и от измерительного прибора M . Естественно ожидать, что смешивание ансамблей приводит к такому же смешиванию распределений, т. е. если $S = \sum_j p_j S_j$, то $\mu_S^M(x) = \sum_j p_j \mu_{S_j}^M(x)$.

Другими словами, вероятности исходов измерения должны быть аффинными функциями состояния. Этого на первый взгляд слабого ограничения оказывается достаточно для вывода обобщенной статистической формулы Борна.

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

$$M_x \geq 0, \quad \sum_{x \in \mathcal{X}} M_x = I, \quad (2.4)$$

и

$$\mu_S(x) = \text{Tr } S M_x. \quad (2.5)$$

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

Набросок доказательства. Эрмитов оператор $A = A^*$ в \mathcal{H} имеет (неединственное) представление $A = t_1 S_1 - t_2 S_2$, где $t_i \geq 0$, а $S_i \in S(\mathcal{H})$. В самом деле, существование представления $A = A_+ - A_-$ ($A_\pm \geq 0$) вытекает из спектрального разложения оператора A , а далее

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

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

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

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

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

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

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

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

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

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

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

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

$$S = \sum s_\omega |\omega\rangle\langle\omega|, \quad s_\omega \geq 0, \quad \sum s_\omega = 1,$$

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

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

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

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

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

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

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

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

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

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

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

- 1) система полна, т. е. $\{|\psi_j\rangle\}^\perp = \{0\}$;
- 2) матрица $P = [\langle\psi_j | \psi_k\rangle]$ идемпотентна, т. е. $P = P^2$.

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

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

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

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

Математический смысл неортогональных разложений единицы проясняет теорема Наймарка.

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

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

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

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

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

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

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

в которой определим псевдоскалярное произведение формулой

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

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

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

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

Этот оператор изометричен так как

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

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

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

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

§ 2.4. Составные квантовые системы

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

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

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

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

Тема 2.
Произведение

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

Таким образом, реализуя $\mathcal{H}_{1,2}$ как пространства ℓ^2 числовых последовательностей $\{c_j^1\}, \{c_k^2\}$, получим реализацию \mathcal{H} в виде пространства матриц $[c_{jk}]$ с величиной $\sum_{j,k} |c_{jk}|^2$ в качестве нормы.

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

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

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

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

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

и продолжая по линейности.

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

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

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

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

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

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

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

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

$$\tilde{\mathcal{H}} = \mathcal{H} \oplus \dots \oplus \mathcal{H} = \mathcal{H} \otimes \mathcal{H}_0,$$

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

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

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

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

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

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

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

§ 2.5. Парадокс ЭПР. Неравенство Белла

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

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

где векторы

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

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

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

а в квантовых вычислениях предпочитают обозначение

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

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

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

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

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

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

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

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

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

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

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

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

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

$$|\mathbf{E} X_1 Y_1 + \mathbf{E} X_1 Y_2 + \mathbf{E} X_2 Y_1 - \mathbf{E} X_2 Y_2| \leq 2, \quad (2.7)$$

где \mathbf{E} — соответствующее математическое ожидание.

Доказательство получается усреднением элементарного неравенства

$$-2 \leq X_1 Y_1 + X_1 Y_2 + X_2 Y_1 - X_2 Y_2 \leq 2.$$

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

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

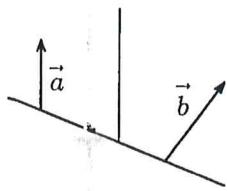


Рис. 2.
Направления спинов

неравенство

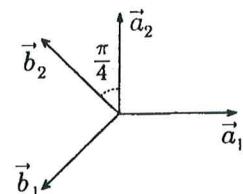


Рис. 3. Выбор
векторов a_j и b_k

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

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