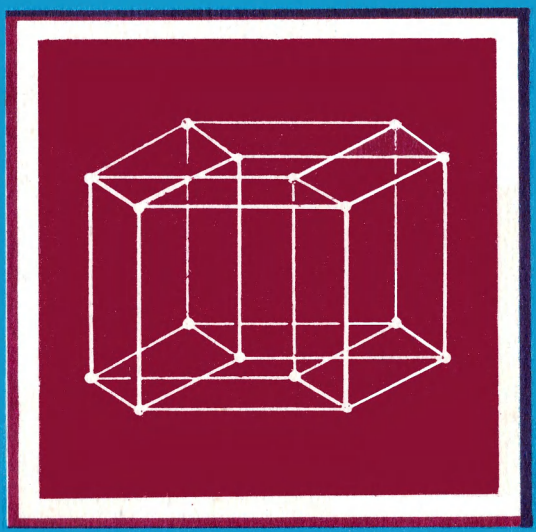
НОВОЕ
В ЖИЗНИ, НАУКЕ,
ТЕХНИКЕ

ЗНАНИЕ

В. К. Леонтьев

ТЕОРИЯ КОДИРОВАНИЯ



6/1977

СЕРИЯ
МАТЕМАТИКА,
КИБЕРНЕТИКА

НОВОЕ
В ЖИЗНИ, НАУКЕ,
ТЕХНИКЕ

Серия «Математика, кибернетика»
№ 6, 1977 г.
Издается ежемесячно с 1967 г.

В. К. Леонтьев,
кандидат физико-математических наук

ТЕОРИЯ КОДИРОВАНИЯ

ИЗДАТЕЛЬСТВО «ЗНАНИЕ»
Москва 1977

51
Л 47

СОДЕРЖАНИЕ

Леонтьев В. К.
Л 47 Теория кодирования. М., «Знание», 1977.

64 с. (Новое в жизни, науке, технике. Серия «Математика, кибернетика», 6. Издается ежемесячно с 1967 г.)

В брошюре рассмотрены основные задачи теории помехоустойчивого кодирования. Указана связь теории кодирования с другими разделами математики: алгеброй, комбинаторным анализом, дискретной геометрией.

Брошюра рассчитана на широкий круг читателей, интересующихся проблемами дискретной математики.

20 200 **51**

ВВЕДЕНИЕ

Хорошо известно, что любой процесс передачи информации в реальном канале связи сопровождается помехами. Вид этих помех может быть самым разнообразным: от замены одной или группы букв сообщения другими буквами до стирания или искажения символов. Как бороться с этими помехами? Увеличивать надежность канала связи техническими средствами? Но часто это бывает очень дорого, и, кроме того, есть некоторый предел надежности, которым обладает любое техническое устройство. Второй способ борьбы с помехами состоит в разумной организации передачи самих сообщений. Другими словами, этот способ состоит в специальном выборе системы кодирования передаваемых сообщений. Эта система кодирования должна обладать тем свойством, что даже при наличии искажений мы имели бы возможность правильно восстановить исходную информацию.

Оказалось, что задача помехоустойчивого кодирования допускает точную математическую формулировку в терминах дискретной геометрии, где основным объектом исследования выступает множество вершин единичного n-мерного куба.

Как устанавливается эта связь, как решаются задачи теории кодирования, какие в ней есть нерешенные задачи, обо всем этом рассказывается в брошюре. При этом упор делается на комбинаторное содержание результатов и методов, которыми они получаются. Интерпретации и технической реализации результатов места почти не уделяется, так как имеется достаточное число хорошо и доступно написанных книг, в которых можно найти как очень содержательные примеры применения идей теории информации, так и технические реализации систем кодирования. Некоторые из таких книг указаны в разделе «Литература».

§ 1. КОДИРОВАНИЕ И ДЕКОДИРОВАНИЕ

1. *Методы повышения надежности передачи.* Широко известными методами борьбы с помехами являются следующие:

1) передача в контексте;
2) дублирование сообщений;
3) передача с переспросом.

Разберем подробней каждый из этих способов.

1. Передача в контексте. С этим хорошо известным и общепринятым способом сталкивался каждый, кто, пытаясь передать по телефону с плохой слышимостью чью-либо фамилию, называл вместо букв, ее составляющих, какие-нибудь имена, первые буквы которых составляют данную фамилию. В данном случае правильному восстановлению искаженного сообщения помогает знание его смыслового содержания.

2. Дублирование сообщений. Этот способ тоже широко применяется в житейской практике, когда для того, чтобы быть правильно понятым, нужное сообщение повторяют несколько раз.

3. Передача с переспросом. В случае, когда получатель имеет связь с источником сообщений, для надежной расшифровки сообщений пользуются переспросом, т. е. просят повторить все переданное сообщение или часть его.

Общим во всех этих способах повышения надежности является введение избыточности, т. е. увеличение тем или иным способом объема передаваемого сообщения для возможности его правильной расшифровки при наличии искажений.

Следует отметить, что введение избыточности уменьшает скорость передачи информации, так как лишь только часть передаваемого сообщения представляет интерес для получателя, а избыточная его доля введена для предохранения от шума и не несет в себе полезной информации.

Естественно выбирать такие формы введения избыточности, которые позволяют при минимальном увеличении объема сообщения обеспечивать максимальную помехоустойчивость. Этой задачей и занимается теория кодирования. Для того, чтобы сделать эту задачу совершенно конкретной, необходимо фиксировать способ передачи информации от источника (слуховые сигналы, азбука Морзе, зрительные образы и т. д.) и выделить помехи, действующие во время передачи.

Мы рассмотрим самый простой и одновременно очень важный канал — так называемый двоичный-симметричный канал (*д. с. к.*) — и на примере этого канала разберем некоторые основные понятия и результаты современной математической теории кодирования.

II. *Двоичный канал.* Итак, предположим, что множество передаваемых сообщений представлено в виде двоичных последовательностей одинаковой длины, состоящих из символов 0,1. Нетрудно понять, что двоичными последовательностями достаточно большой длины может быть закодировано любое сообщение: слово, предложение, книга и т. д. Кодирование в двоичном алфавите удобно в техническом отношении, так как такое сообщение легко передавать по электрической линии связи, сопоставив элементарным сигналам — нулю и единице — электрические импульсы различной длительности. Широко известным практическим примером такого кодирования является код Морзе.

Предположим теперь, что при передаче элементарных сигналов могут происходить два вида ошибок: нуль переходит в единицу или единица переходит в нуль. При этом вероятность перехода единицы в нуль равна вероятности перехода нуля в единицу и равна p. Вероятность же безошибочной передачи символа равна $q=1-p$.

Далее предположим, что для передачи по *д. с. к.* используются последовательности из 0,1 длины n, т. е. элементарное сообщение представляет собой набор из 0,1 длины n. Каждое подмножество множества всех последовательностей длины n называется блоковым кодом, а сами эти последовательности — кодовыми словами.

Список возможных для передачи кодовых слов, т. е. сам код, известен и на приемном конце канала. При получении на этом конце некоторого слова (двоичной последовательности) получателю нужно принять решение относительно переданного слова. Это решение должно быть таким, чтобы обеспечить максимальную правильность или надежность приема. Измерение этой надежности может быть выполнено не единственным способом, и имеются различные критерии надежности, применяющиеся в тех или иных практических задачах.

Для более четкого представления об имеющейся ситуации рассмотрим следующий пример, заимствованный нами из книги У. Питерсона «Коды, исправляющие ошибки».

Пусть у нас имеются для передачи четыре сообщения: a, b, c, α. Предположим, что эти сообщения будут передаваться в виде двоичных последовательностей длины пять. Таким образом, первая задача состоит в сопоставлении каждому сообщению некоторого двоичного набора длины пять. Эта первая задача называется проблемой кодирования. Предположим, что мы решили ее следующим образом:

$$a = 11000$$
$$b = 00110$$
$$c = 10011$$
$$\alpha = 01101$$

Вторая задача состоит в следующем. Предположим, что на приемном конце мы получили некоторую последовательность $(\alpha_1\alpha_2\alpha_3\alpha_4\alpha_5)$, которая вследствие происшедших ошибок не совпадает ни с одной из кодовых последовательностей. Как нам решить, какое из сообщений a, b, c или α было передано? Эта вторая задача называется проблемой декодирования.

Сам процесс принятия решения на приемном конце обычно описывается с помощью таблицы декодирования. Эта таблица показывает, какое мы принимаем решение для каждой из возможных последовательностей длины пять, которые могут появиться на приемном конце. Пример такой таблицы изображен ниже.

11000	00110	10011	01101
11001	00111	10010	01100
11010	00100	10001	01111
11100	00010	10111	01001
10000	01110	11011	00101
01000	10110	00011	11101
11110	00000	01011	10101
01010	10100	11111	00001

Таким образом, если мы на выходе получили последовательность $\bar{\alpha}=10001$, находящуюся в третьем столбце таблицы декодирования, то мы считаем, что было передано сообщение $c=10011$. Аналогично принимается решение для любой из $2^5=32$ последовательностей, которые могут появиться на выходе приемного устройства. Следует отметить, что рассмотренная нами таблица не является единственно возможной таблицей декодирования, а представляет просто одну из возможных реализаций таких таблиц. В частности, перемена местами любых двух после-

довательностей, не лежащих в самой верхней строчке, опять дает нам некоторую новую таблицу декодирования. Аналогично обстоит дело и с кодированием. Выбор нами для сообщений a, b, c, и α написанных выше кодовых слов также не является неизбежным, а представляет лишь один из $C_{32}^4 = 35960$ возможных вариантов. Чем же по существу отличаются различные способы выбора кода и как нужно при заданном коде строить таблицу декодирования? Это и есть основные вопросы, которыми занимается теория кодирования.

Обратимся опять к таблице декодирования. Заметим, что эта таблица позволяет «исправить» любую одиночную ошибку в каждом из кодовых слов. Действительно, возьмем кодовое слово сообщения a, т. е. набор 11000. Если в нем изменить один из пяти символов на другой, то получатся следующие пять наборов.

11001
11010
11100
10000
01000

Все эти пять наборов расположены в первом столбце таблицы декодирования. А это означает, что при получении на приемном конце любого из этих наборов мы считаем, что было передано сообщение a, в данном случае это соответствует истине. Аналогично обстоит дело и с остальными тремя сообщениями, в чем можно убедиться точно таким же образом. Однако все двойные ошибки этот код исправить уже не может. Действительно, если при передаче сообщения a произошли ошибки в первых двух разрядах кодового слова, то на выходе мы получим нулевой набор (0, 0, 0,0, 0). Этот набор находится во втором столбце таблицы декодирования, и поэтому мы вынуждены принять ошибочное решение о передаче сообщения b.

В то же время некоторые комбинации двойных ошибок декодируются правильно. Если, например, передавалось слово 11000 и в третьем и четвертом символе этого слова произошли ошибки, то на приемном конце мы получим слово 11110, которое находится в первом столбце таблицы декодирования и, значит, будет восстановлено как слово 11000, т. е. правильно.

Следует отметить, что в принципе невозможно построить такую таблицу декодирования, которая давала бы возможность правильно восстановить кодовое слово при любой

комбинации ошибок. Действительно, если под воздействием помех кодовое слово сообщения a перешло в кодовое слово сообщения b, то на приемном конце мы естественно примем решение о передаче слова b, т. е. допустим ошибку. Таким образом, идеального решения задачи не существует вовсе, и мы должны стремиться выбрать код и таблицу декодирования таким образом, чтобы исправлять «максимальное число» ошибок, т. е. сделать вероятность ошибки декодирования как можно более меньшей.

III. *Геометрическая модель.* Для более точной и наглядной постановки этой задачи мы воспользуемся известной геометрической моделью единичного n-мерного куба.

Основной объект — это множество двоичных последовательностей длины n, интерпретируемое как множество вершин единичного n-мерного куба E^n.

Множество вершин куба E^n — это множество всех возможных сообщений. Нетрудно видеть, что куб E^n имеет ровно 2^n вершин. Любой конкретный код — это некоторое подмножество вершин E^n. Вершины этого куба мы будем обозначать так: $\bar{\alpha}=(\alpha_1\alpha_2 \ldots \alpha_n)$, где α_i равно нулю или единице. Вершины куба E^n мы иногда будем называть точками E^n.

Под воздействием помех точки кода $V \subseteq E^n$ будут переходить в некоторые другие точки E^n. Чтобы охарактеризовать положение этих «ошибочных» точек относительно первоначальных, вводится понятие расстояния между вершинами куба E^n.

Расстоянием, по Хэммингу, между вершинами $\bar{\alpha}=(\alpha_1\alpha_2 \ldots \alpha_n)$ и $\bar{\beta}=(\beta_1\beta_2 \ldots \beta_n)$ называется число разрядов, в которых эти вершины различаются. Математически это принято записывать так:

$$\rho(\bar{\alpha}, \bar{\beta}) = \sum_{i=1}^{n} |\alpha_i - \beta_i| \cdot$$

Например, расстояние между вершинами $\bar{\alpha}=(1, 1, 0)$ и $\beta=(0,1, 0)$ равно единице, так как эти вершины различаются лишь в первом разряде. Ясно, что расстояние между любыми двумя различными вершинами куба E^n не меньше единицы и не больше n.

Шаром $S_{\bar{\alpha}}(t)$ радиуса t с центром в точке $\bar{\alpha} \in E^n$ мы будем называть множество точек из E^n, находящихся от $\bar{\alpha}$ на расстоянии не больше чем t. Так, в случае $\bar{\alpha}=(0, 0, 0)$

и $t=1$ шар $S\bar{\alpha}$ (1) состоит из следующих точек: $\bar{\alpha}=(0, 0, 0)$, $\bar{\beta}=(1, 0, 0)$, $\bar{\gamma}=(0, 1, 0)$, $\bar{\delta}=(0, 0, 1)$.

Нетрудно видеть, что это определение является дискретным аналогом обычного определения шара. Число точек в каждом шаре радиуса t в E^n равно следующей сумме биноминальных коэффициентов:

$$S_n^t=1+C_n'+\ldots+C_n^t$$

(здесь через S_n^t обозначено число точек в шаре радиуса t из E^n).

Сферой $S_{\bar{\alpha}}^0(t)$ радиуса t с центром в точке $\bar{\alpha}\in E^n$ называется множество точек из E^n, находящихся от $\bar{\alpha}$ на расстоянии t. Так, в случае $\bar{\alpha}=(1, 0, 0)$ и $t=2$ сфера $S_{\bar{\alpha}}^0$ (2) состоит из следующих точек: $\bar{\beta}=(1, 1, 1)$, $\bar{\gamma}=(0, 1, 0)$, $\bar{\delta}=(0, 0, 1)$. Число точек на сфере радиуса t в E^n равно C_n^t.

Оказывается, что эти привычные геометрические понятия являются довольно удобным языком для описания многих фактов, относящихся к теории кодирования.

Пусть $V=\{\bar{\alpha}_1\bar{\alpha}_2\ldots\bar{\alpha}_s\}$ — некоторый код. Целесообразность его использования для передачи по каналу с шумом определяется возможностью этого кода обнаруживать и исправлять ошибки.

Говорят, что код V обнаруживает t ошибок, если изменение в любом кодовом слове $\bar{\alpha}_i$ любого числа $r\leqslant t$ символов приводит к слову $\bar{\beta}_i$, которое уже не входит в множество V. Смысл этого определения состоит в том, что оно выделяет класс кодов, которые можно использовать для передачи по каналу с шумом при наличии обратной связи. Другими словами, получив слово, которое не входит в список передаваемых сообщений, мы просим повторить передачу и делаем это до тех пор, пока не восстановим исходное сообщение.

Нетрудно понять, что код, рассмотренный нами выше, обнаруживает любую комбинацию из одной и двух ошибок, так как изменение любого из кодовых слов сообщений a, b, c или d в одном или двух местах приводит к слову, отличному от исходных.

Способность кода V обнаруживать любую комбинацию из t или меньшего числа ошибок определяется следующим условием: код V обнаруживает t-ошибок тогда и только тогда, когда расстояние Хэмминга между любыми двумя точками из V не меньше, чем $t+1$. Действительно, возьмем про-

извольную точку $\bar{\alpha} \in V$ и рассмотрим шар $S_{\bar{\alpha}}$ (t). Ясно, что под воздействием любого числа $r \leqslant t$-ошибок точка $\bar{\alpha}$ не выйдет за пределы этого шара. С другой стороны, по условию в шаре $S_{\bar{\alpha}}$ (t) нет других кодовых точек. Это и доказывает наше утверждение.

Если же мы хотим с помощью кода V правильно декодировать любое кодовое слово, в котором произошло t или меньше ошибок, то надо потребовать выполнение следующего условия: расстояние Хэмминга между любыми двумя точками $\bar{\alpha}_i$ и $\bar{\alpha}_j$ из кода V должно быть не меньше, чем $2t+1$.

Действительно, если это условие выполнено, то таблицу декодирования для кода V можно построить следующим образом: под каждым кодовым словом надо написать все точки из шара радиуса t с центром в этом кодовом слове.

Так как расстояние между любыми двумя кодовыми словами не меньше, чем $2t+1$, то каждая из выписанных точек встретится в таблице только один раз. Это и обеспечит возможность правильного декодирования.

Геометрически описанная ситуация выглядит следующим образом: каждое кодовое слово мы окружаем «защитным» шаром радиуса t. По условию эти шары не пересекаются. Любое из слов, полученных на выходе, принадлежит одному из «защитных» шаров. Каждое из таких слов мы декодируем в центр того «защитного» шара, которому оно принадлежит.

Таким образом мы выяснили, что важнейшей характеристикой кода V, определяющей его корректирующие возможности, является минимальное расстояние между кодовыми точками. Эта характеристика обозначается через $d(V)$ и называется кодовым расстоянием. Отметим, что в рассмотренном выше примере код V имеет кодовое расстояние 3, т. е. он способен исправлять любую единичную ошибку и обнаруживать любую комбинацию из двух ошибок. Таблица декодирования в рассмотренном нами примере как раз составлена таким образом, что под каждым кодовым словом находятся пять точек из шара радиуса единица с центром в этом кодовом слове.

IV. *Задачи теории кодирования.* Теперь мы уже в состоянии точно сформулировать одну из важнейших задач теории кодирования. Она состоит в следующем. Требуется построить код, исправляющий t ошибок и имеющий максимально возможное число точек. В геометрической поста-

новке эта же задача звучит следующим образом: среди вершин единичного n-мерного куба E^n требуется выделить максимальное число таким способом, чтобы расстояние между любыми двумя выделенными вершинами было не меньше, чем $2t+1$. Это максимальное число обозначается обычно через $A(n, 2t+1)$.

Другая, связанная с предыдущей, задача состоит в расположении s точек в вершинах E^n так, чтобы наименьшее из попарных расстояний между ними было возможно большим. Это расстояние обозначается через $d(s, n)$.

Выражения «построить», «выделить», «расположить» нуждаются в уточнении, так как нам вовсе не безразлично, в каком виде будет задан искомый код. Самый простой способ — перечисление всех кодовых точек — является неэффективным, требует большой памяти. Поэтому нужен такой способ задания кода, который позволяет просто восстановить каждую точку кода по ее номеру. Другими словами, код должен иметь «простую реализацию». С некоторыми просто реализуемыми кодами мы познакомимся ниже. С другой стороны, мы хотим иметь возможность просто восстанавливать исходное сообщение на выходе.

Таким образом, «хороший код» должен удовлетворять следующим трем естественным требованиям:

1) исправлять много ошибок, т. е. иметь большое кодовое расстояние;

2) иметь несложную реализацию;

3) обладать простым алгоритмом исправления ошибок на приемном конце.

Следует отметить, что эти требования в значительной степени являются противоречивыми, так как код V, исправляющий много ошибок, вовсе не обязан иметь простую реализацию и тем более простой алгоритм декодирования. Поэтому на практике применяются коды, которые обладают в «достаточной» мере всеми тремя перечисленными выше качествами.

Для количественной оценки свойств кода требуется иметь оценки отклонения его параметров от параметров «идеального» кода. Для получения таких оценок нужно иметь хотя бы приближенные значения важнейших параметров «идеального» кода, т. е. значения $A(n, 2t+1)$ и $d(s, n)$. Следует отметить, что функции $A(n, 2t+1)$ и $d(s, n)$ не являются единственными параметрами, характеризующими качество кода. Не менее важными являются также такие параметры кода, как вероятность правильного

декодирования и вероятность обнаружения ошибки, с которыми мы познакомимся ниже. Имеется также еще ряд других важных критериев, применяющихся для оценки качества кодов, но на них мы не будем останавливаться.

§ 2. ОЦЕНКИ КОРРЕКТИРУЮЩИХ ВОЗМОЖНОСТЕЙ КОДОВ

1. *Верхние оценки корректирующих способностей кодов.* Одним из первых результатов в теории корректирующих кодов является следующее неравенство Хэмминга:

$$A(n,\ 2t+1) \leqslant \frac{2^n}{1+C_n^1+\ldots+C_n^t}. \qquad (1)$$

Это неравенство вытекает из следующих соображений.

Если в множестве $V \in E^n$ расстояние между любыми двумя точками не меньше, чем $2t+1$, то шары радиуса t с центрами в точках множества V не имеют пересечений. Поэтому общее число точек в этих шарах равно:

$$|V|\,(1+C_n^1+\ldots+C_n^t),$$

где $|V|$ — число точек в коде V, а $1+C_n^1+\ldots+C_n^t$ — число точек в шаре радиуса t.

Так как число точек, попавших в шары, очевидно, не превосходит общего числа точек в E^n, то

$$|V|\,(1+C_n^1+\ldots+C_n^t) \leqslant 2^n,$$

откуда и следует неравенство Хэмминга.

Прежде чем перейти к установлению еще некоторых неравенств для функции $A(n,\ 2t+1)$, нам понадобится еще один взгляд на множество E^n, позволяющий глубже проникнуть в структуру этого множества.

Введем в E^n операцию сложения по модулю два (mod 2). Для этого сначала введем эту же операцию в множество из двух элементов $\{0,1\}$. Положим:

$$0 \oplus 0 = 1$$
$$0 \oplus 1 = 1$$
$$1 \oplus 0 = 1$$
$$1 \oplus 1 = 0$$

Если наряду с этой операцией рассмотреть обычную операцию умножения, то множество $\{0,1\}$ с этими двумя операциями будет представлять собой так называе-

мое поле Галуа, которое обозначается через $GF(2)$. Суммой двух точек $\bar{\alpha}=(\alpha_1\alpha_2 \ldots \alpha_n)$, $\bar{\beta}=(\beta_1\beta_2 \ldots \beta_n) \in E^n$ называется такая точка $\bar{\gamma}=(\gamma_1\ \gamma_2\ \ldots\ \gamma_n)$, координаты которой определяются равенствами:

$$\gamma_i=\alpha_i \oplus \beta_i. \quad i=1, 2, \ldots, n$$

Эта сумма тоже обозначается знаком $\oplus$.

Пример 1. Пусть $\bar{\alpha}=(1, 0, 0, 1, 1)$, $\bar{\beta}=(0, 1, 0, 1, 1)$. Тогда

$$\bar{\alpha} \oplus \bar{\beta}=(1, 1, 0, 0, 0)$$

Введенная операция сложения по mod 2 обладает многими свойствами, присущими обычной операции сложения. Например:

1) $\bar{\alpha} \oplus \bar{\beta}=\bar{\beta} \oplus \bar{\alpha}$ — перестановочность слагаемых;

2) $\bar{\alpha} \oplus (\bar{\beta} \oplus \bar{\gamma})=(\bar{\alpha} \oplus \bar{\beta}) \oplus \bar{\gamma}$ — ассоциативность.

Но у неё есть и некоторые необычные свойства. Например, из равенства

$$\bar{\alpha} \oplus \bar{\beta}=\bar{\gamma}$$

следуют два таких равенства:

$$\bar{\alpha}=\bar{\beta} \oplus \bar{\gamma};$$
$$\bar{\beta}=\bar{\alpha} \oplus \bar{\gamma}.$$

Множество E^n с этой операцией является линейным векторным пространством над полем $GF(2)$. В дальнейшем элементы E^n мы будем называть также и векторами. Под нормой вектора $\bar{\alpha}=(\alpha_1\alpha_2 \ldots \alpha_n)$ из E^n понимается число единичных координат этого вектора. Норма вектора $\bar{\alpha}$ обозначается через $\|\bar{\alpha}\|$. С помощью нормы и операции сложения по mod 2 легко получить следующую формулу для расстояния Хэмминга:

$$\rho(\bar{\alpha}, \bar{\beta}) = \|\bar{\alpha} \oplus \bar{\beta}\|.$$

В теории кодирования очень часто норму вектора называют также весом этого вектора. Все множество точек E^n по весам распадается на $(n+1)$ — слой: E_0^n, E_1^n $\ldots E_n^n$, где E_i^n — все точки из E^n веса i. Слой E_i^n содержит ровно C_n^i — точек.

Сдвигом множества $V \in E^n$ на точку $\bar{\alpha} \in E^n$ называется множество $V \oplus \bar{\alpha} = \{\bar{\alpha}_i \oplus \bar{\alpha}, \bar{\alpha}_i \in V\}$.

Пример 2. Пусть $V=\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ и $\bar{\alpha}=(1, 1, 1)$. Тогда сдвигом множества V на точку $\bar{\alpha}$ будет множество $V\oplus\bar{\alpha}=\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$.

Преобразование сдвига является аналогом хорошо известного в элементарной геометрии преобразования параллельного переноса. Это преобразование обладает тем полезным свойством, что не изменяет расстояний между точками множества V. В частности,

$$d(V\oplus\bar{\alpha})=d(V),$$

т. е. кодовые расстояния множества V и $V\oplus\bar{\alpha}$ одинаковы, что позволяет из одних кодов получать другие с теми же характеристиками.

Рассмотрим теперь некоторые простые применения этого преобразования.

Пусть $V\subset E^n$ и $d(V)=d$. Возьмем некоторое множество V' из E^n, обладающее следующим свойством: расстояние между любыми двумя точками из V' не превосходит $d-1$. Пусть $V=\{\bar{\alpha}_1\bar{\alpha}_2 \dots \bar{\alpha}_m\}$ и $V'=\{\bar{\beta}_1\bar{\beta}_2 \dots \bar{\beta}_k\}$. Рассмотрим сдвиги множества V на вектора множества V', т. е. следующие множества:

$$V\oplus\bar{\beta}_1,\ V\oplus\bar{\beta}_2 \dots V\oplus\bar{\beta}_k.$$

Заметим, что никакие два из этих множеств не имеют общих точек. Действительно, если $z\in V\oplus\bar{\beta}_i$ и $z\in V\oplus\bar{\beta}_j$, то

$$z=\bar{\alpha}_s\oplus\bar{\beta}_i=\bar{\alpha}_r\oplus\bar{\beta}_j.$$

Отсюда следует, что

$$\bar{\alpha}_s\oplus\bar{\alpha}_r=\bar{\beta}_i\oplus\bar{\beta}_j$$

или

$$\|\bar{\alpha}_s\oplus\bar{\alpha}_r\| = \|\bar{\beta}_i\oplus\bar{\beta}_j\|.$$

Но

$$\|\bar{\alpha}_s\oplus\bar{\alpha}_r\| = \rho(\bar{\alpha}_s, \bar{\alpha}_r)\geqslant d(V)=d,$$

а с другой стороны,

$$\|\bar{\beta}_i\oplus\bar{\beta}_j\| = \rho(\bar{\beta}_i\bar{\beta}_j)\leqslant d-1.$$

Противоречие.

Так как множества $V\oplus\bar{\beta}_i$ не пересекаются и каждое из них содержит ровно $|V|$ точек, то

$$|V|\cdot|V'|\leqslant 2^n. \qquad (2)$$

Из этого неравенства легко получить упомянутое выше неравенство Хэмминга, а также неравенство Джоши.

Возьмем в качестве V' шар радиуса t в E^n. Ясно, что расстояние между любыми двумя точками в этом шаре не превосходит $2t$, а число точек в этом шаре равно $S_n^t = 1 + C_n' + \ldots + C_n^t$, т. е. $|V'| = S_n^t$. Из приведенного выше неравенства получаем:

$$|V| \leqslant 2^n |S_n^t,$$

т. е. оценку Хэмминга.

Если же мы возьмем в качестве V' множество двоичных последовательностей длины n, у которых на первых $(d-1)$ местах стоят всевозможные наборы длины $(d-1)$, а на остальных местах — нули, то расстояние между любыми двумя из этих наборов не превосходит $(d-1)$, а число их есть 2^{d-1}. Опять применяя полученное выше неравенство, получаем

$$|V| \cdot 2^{d-1} \leqslant 2^n,$$

т. е.

$$A(n, d) \leqslant 2^{n-d+1}.$$

Это неравенство принадлежит Джоши.

Обе оценки, Хэмминга и Джоши, являются достижимыми для некоторых t и d. Другими словами, существуют последовательности кодов $\{V_n\}$ и $\{M_n\}$, таких, что $|V_n| = 2^n |S_n^t$ и $|M_n| = 2^{n-d+1}$. В случае оценки Хэмминга это происходит при $t=1$, а в случае оценки Джоши — при $d=2$. Эти случаи мы подробно рассмотрим несколько позже.

Имеется еще ряд верхних оценок для функции $A(n, d)$, но они выглядят несколько сложней, чем предыдущие, и мы на них не будем останавливаться.

Другой важной в теории кодирования функцией является максимальное число исправляемых ошибок при заданном числе сообщений. Более точно эта функция определяется так: $d(s, n)$ — максимальное кодовое расстояние в классе всех подмножеств E^n, содержащих ровно s точек. Функции $A(n, d)$ и $d(s, n)$ связаны очевидным соотношением: $A(n, d(s, n)) = s$. Величина $d(s, n)$ представляет для теории кодирования столь же существенный интерес, как и $A(n, d)$, и для нее также получен ряд верхних и нижних оценок.

Следующая верхняя оценка функции $A(n, d)$ основана на вычислении суммы всех попарных расстояний между точками фиксированного множества $V \in E^n$.

Итак, пусть $V = \{\bar{x}_1 \bar{x}_2 \ldots \bar{x}_s\}$ и $\bar{x}_i = \{\alpha_1^i\ \alpha_2^i \ldots \alpha_n^i\}$. Имеем:

$$\sum_{i,\ j} \rho(\bar{x}_i, \bar{x}_j) = \sum_{i,\ j} \sum_{t=1}^{n} |\alpha_t^i - \alpha_t^j| = \sum_{i,\ j} \sum_{t=1}^{n} (\alpha_t^i - \alpha_t^j)^2 =$$

$$= \sum_{i,\ j} \sum_{t=1}^{n} \alpha_t^i + \sum_{i,\ j} \sum_{t=1}^{n} \alpha_t^j - 2 \sum_{i,\ j} \sum_{t=1}^{n} \alpha_t^i \alpha_t^j = s \sum_{t=1}^{n} \sum_{i=1}^{s} \alpha_t^i +$$

$$+ s \sum_{t=1}^{n} \sum_{j=1}^{s} \alpha_t^j - 2 \sum_{i,\ j} \sum_{t=1}^{n} \alpha_t^i \alpha_t^i.$$

Здесь мы воспользовались двумя очевидными фактами:
1) $(\alpha_t^i)^2 = \alpha_t^i$, так как α_t^i есть нуль или единица;
2) $|\alpha_t^i - \alpha_t^j| = (\alpha_t^i - \alpha_t^j)^2$ по этой же причине.

Если выписать все точки множества V в матрицу G

$$G = \begin{Vmatrix} \alpha_1^1 \alpha_2^1 \ldots \alpha_n^1 \\ \alpha_1^2 \alpha_2^2 \ldots \alpha_n^2 \\ \ldots\ldots\ldots \\ \alpha_1^s \alpha_2^s \ldots \alpha_n^s \end{Vmatrix},$$

то $\sum_{i=1}^{s} \alpha_t^i$ будет равна числу k_t единиц в t-м столбце этой матрицы. Отсюда

$$\sum_{i,\ j} \rho(\bar{x}_i, \bar{x}_j) = 2s \sum_{t=1}^{n} k_t - 2 \sum_{t=1}^{n} \sum_{i=1}^{s} \alpha_t^i \sum_{j=1}^{s} \alpha_t^j = 2s \sum_{t=1}^{n} k_t - 2 \sum_{t=1}^{n} k_t^2 =$$

$$= 2 \sum_{t=1}^{n} k_t (s - k_t).$$

Так как в сумме слева каждое расстояние учитывается два раза: один раз как $\rho(\bar{x}_i \bar{x}_j)$, другой как $\rho(\bar{x}_j, \bar{x}_i)$, то

$$\sum_{i<j} \rho(\bar{x}_i, \bar{x}_j) = \sum_{t=1}^{n} k_t (s - k_t).$$

Так как $k_t (s - k_t) \leqslant s^2/4$, то

$$\sum_{i<j} \rho(\bar{x}_i, \bar{x}_j) \leqslant \frac{ns^2}{4}. \qquad (3)$$

Это неравенство можно использовать для получения верхней оценки функции $A(n, d)$.

Пусть код $V=\{x_1 x_2 \ldots x_s\}$ имеет минимальное расстояние d. Тогда

$$\rho(\bar{x}_i \bar{x}_j) \geqslant d. \quad i, j=1,2,\ldots, s.$$

Используя это неравенство совместно с предыдущим, получаем

$$s \leqslant \frac{2d}{2d-n}$$

или неравенство

$$A(n, d) \leqslant \frac{2d}{2d-n}, \tag{4}$$

справедливое при $2d>n$.

Неравенство (4) носит название границы Плоткина.

II. Обсуждение границ. Границы Хэмминга и Джоши справедливы при всех значениях параметров n и d, но граница Хэмминга становится точнее границы Джоши при $d \geqslant 3$. При $d=2$ точнее граница Джоши. Неравенство Плоткина справедливо лишь при больших значениях d, удовлетворяющих неравенству $2d>n$. В этом случае оно дает более точные границы для функции $A(n, d)$, чем неравенства Хэмминга и Джоши.

Существуют и более точные верхние границы для функции $A(n, d)$ и связанной с ней функцией $d(s, n)$, но они выглядят несколько сложней, чем предыдущие, и мы на них не будем останавливаться.

Первый приведенный нами способ доказательства неравенства Хэмминга носит название метода «плотной упаковки». Это название связано с тем, что правая часть этого неравенства дает верхнюю оценку числа непересекающихся шаров радиуса t, которые можно разместить в вершинах E^n при «самой плотной» упаковке.

Коды V, для которых неравенство Хэмминга превращается в равенство, называются совершенными или плотно упакованными. Эти коды соответствуют разбиению вершин множества E^n на непересекающиеся шары радиуса t. Совершенные коды являются оптимальными, так как они имеют максимальное число точек среди всех кодов, исправляющих t ошибок. Совершенные коды обладают многими интересными свойствами, связанными с их геометрической структурой. Более подробно об этом будет сказано в разделе «Важнейшие классы кодов».

Неравенство Джоши обращается в равенство при $d=2$. Это означает, что максимальное число точек в E^n с кодовым расстоянием два равно 2^{n-1}, т. е. ровно половине всех точек E^n. Простейшим примером множества, состоящего из 2^{n-1} точек и имеющего кодовое расстояние два, является так называемый «счетчик четности», т. е. множество вершин E^n, имеющих четное число единичных координат. Действительно, так как каждая точка «счетчика четности» имеет четный вес, то и расстояние между любыми двумя точками «счетчика четности» есть четное число, т. е. по меньшей мере равно двум.

«Счетчик четности» является оптимальным кодом с обнаружением одной ошибки. Действительно, любая единичная ошибка в кодовом слове изменяет «четность» числа единиц, что и позволяет обнаруживать ошибку. Оптимальность этого кода состоит в том, что он имеет максимальное число точек среди всех кодов с обнаружением одной ошибки.

Граница Плоткина также является достижимой. Коды, на которых достигается эта граница, имеют специальную алгебраическую структуру. Эти коды обладают большой корректирующей способностью и, как показывает неравенство Плоткина, имеют небольшой объем (число кодовых точек). Более подробно об этих кодах будет сказано в разделе, посвященном специальным классам кодов.

III. *Граница Варшамова — Гилберта.* Существует простой метод построения кодов, исправляющих ошибок, который фактически основан лишь на определении этих кодов. Этот метод состоит в следующем.

1) выбираем произвольную точку $\bar{x}_1 \in E^n$ и называем ее первой кодовой точкой;

2) «окружаем» точку $\bar{x}_1$ шаром радиуса $2t$ и в качестве точки $\bar{x}_2$ выбираем любую из точек E^n, не попавших в этот шар;

3) «окружаем» точку $\bar{x}_2$ шаром радиуса $2t$ и в качестве точки $\bar{x}_3$ выбираем любую из точек E^n, не попавших ни в первый, ни во второй шар.

Продолжаем эту процедуру до тех пор, пока каждая из точек E^n не будет принадлежать одному из построенных шаров. Центры этих шаров и будут образовывать код $V = \{\bar{x}_1, \bar{x}_2 \ldots \bar{x}_s\}$. По построению центры этих шаров удалены друг от друга на расстояние не менее чем $2t+1$, т. е. код V действительно исправляет t ошибок.

Этот способ построения, несмотря на свою банальность, позволяет дать хорошую оценку снизу для числа точек кода V. Действительно, если построение завершено, то s шаров радиуса $2t$ «покрывают» весь куб E^n. Так как каждый из шаров радиуса $2t$ содержит $S_n^{2t}=1+C_n^1+\ldots+C_n^{2t}$ точек, то для «покрытия» всех 2^n вершин E^n нужно по меньшей мере $2^n/S_n^{2t}$ шаров. Отсюда следует оценка для s:

$$s \geqslant \frac{2^n}{1+C_n^1+\ldots+C_n^{2t}}. \qquad (5)$$

Это неравенство и носит название границы Варшамова—Гилберта.

При малых значениях t граница Варшамова—Гилберта сильно отличается от известных верхних границ функции $A(n, 2t+1)$. Однако при t, пропорциональных n, и при больших n эта граница находится вблизи известных верхних границ, и вполне может случиться так, что при отмеченных выше условиях неравенство Варшамова—Гилберта дает асимптотически правильный результат.

Следует отметить, что теорема Варшамова—Гилберта, несмотря на конструктивный характер доказательства, скорее должна быть отнесена к теоремам существования. Дело в том, что о построенном коде мы фактически ничего не знаем. Единственная информация о нем состоит в том, что этот код исправляет t ошибок. Никаких же сведений о «внутренних» свойствах этого кода мы не имеем. Тем не менее при каждых конкретных значениях n и t описанная выше процедура может быть реализована, и мы получим код V с исправлением t ошибок и с числом точек не менее той величины, которую нам гарантирует граница Варшамова—Гилберта.

Описанная выше ситуация не является уникальной, а представляет лишь частный пример процедуры построения «неконструктивного» множества. Имеется много других задач, в которых может быть использован подобный способ рассуждений; полученные при помощи этого способа множества являются в том же смысле неконструктивными, как и коды Варшамова — Гилберта.

§ 3. ВАЖНЕЙШИЕ КЛАССЫ КОДОВ

1. *Групповые коды.* Обратимся к конструктивным методам построения кодов, исправляющих ошибки. Здесь нам

существенную помощь окажет «алгебраический» взгляд на множество E^n как на n-мерное векторное пространство с операцией сложения векторов по mod 2, которая была определена нами раньше.

Теперь мы в качестве кодов будем рассматривать не произвольные подмножества E^n, а лишь те подмножества $V=\{\bar{x}_1\ \bar{x}_2\ \ldots\ \bar{x}_s\}\subseteq E^n$, которые вместе с любыми двумя своими точками $\bar{x}_i$ и $\bar{x}_j$ содержат и их сумму, т. е. точку $\bar{x}_i\oplus\bar{x}_j$. Такие подмножества называют групповыми кодами.

Пример 3. Пусть $n=3$. Рассмотрим код $V=\{\bar{x}_1=(1, 0, 0),\ \bar{x}_2=(0, 1, 0),\ \bar{x}_3=(1, 1, 0),\ \bar{x}_4=(0, 0, 0)\}$. Легко проверить справедливость следующих равенств:

$$\bar{x}_1\oplus\bar{x}_2=\bar{x}_3$$
$$\bar{x}_1\oplus\bar{x}_3=\bar{x}_2$$
$$\bar{x}_1\oplus\bar{x}_4=\bar{x}_1$$
$$\bar{x}_2\oplus\bar{x}_3=\bar{x}_1.$$

Отсюда и следует, что сумма по mod 2 любых двух точек множества V опять является точкой этого же множества, т. е. V-групповой код.

Групповые коды благодаря своей алгебраической структуре занимают центральное место в теории корректирующих кодов. Первое преимущество этих кодов состоит в том, что для задания их вовсе не обязательно перечислять все элементы. Достаточно запомнить лишь небольшую долю всех элементов, по которым затем можно однозначно восстановить весь код. Другими словами, весь групповой код можно задавать множеством образующих элементов, которые составляют базис подпространства. При этом за базис кода V можно взять любое множество из k линейно-независимых векторов подпространства V (k — размерность пространства V). Другими словами, за множество образующих кода можно взять любое множество из k векторов этого кода, сумма любых r, из которых не равна нулю ($r=1, 2, \ldots, k$.) Например, в рассмотренном выше коде V за множество образующих можно принять векторы $\bar{x}_1$ и $\bar{x}_2$. Ясно, что:

$$\bar{x}_3=\bar{x}_1\oplus\bar{x}_2;$$
$$\bar{x}_4=\bar{x}_1\oplus\bar{x}_1,$$

т. е. любой вектор этого кода может быть получен как сумма некоторого числа базисных векторов. Если множество образующих векторов кода V выписать в матрицу G, то мы

получим порождающую матрицу кода V. Любой из z^k векторов кода V может быть представлен в виде суммы r строк матрицы G ($r=1, 2, \ldots, k$). Таким образом, для задания всех 2^k векторов кода V достаточно иметь всего k строк его порождающей матрицы. При этом за строки матрицы G могут быть взяты любые k линейно-независимых векторов кода V. Код V, задаваемый порождающей матрицей размеров $k \times n$, называется групповым (n, k)-кодом.

Для кода, описанного в предыдущем примере, порождающей будет следующая матрица:

$$G = \left\| \begin{matrix} 100 \\ 010 \end{matrix} \right\|$$

Другим, но связанным с предыдущим способом описания групповых кодов является описание с помощью проверочной матрицы.

В пространстве E^n обычным образом определяется скалярное произведение векторов $\bar{x}=(x_1 x_2 \ldots x_n)$, $\bar{y}=(y_1 y_2 \ldots \ldots y_n)$, т. е.

$$(\bar{x}, \bar{y}) = \sum_{i=1}^{n}{}' x_i y_i.$$

Знак Σ' означает, что сложение ведется по mod 2. Так, введенное скалярное произведение принимает лишь два значения: нуль или единица. При этом, как обычно, вектора $\bar{x}$ и $\bar{y}$ называются ортогональными, если $(\bar{x}, \bar{y})=0$.

Вместе с групповым (n, k)-кодом V удобно рассматривать и двойственный к нему код V^*, элементами которого являются те и только те векторы из E^n, которые ортогональны всем векторам кода V. Можно показать, что множество V^* является групповым $(n, n-k)$ кодом. В чисто алгебраическом смысле множество V^* является ортогональным подпространством для пространства V. Порождающая матрица H двойственного кода V^* называется проверочной матрицей кода V. Таким образом, вектор $\bar{v}$ из E^n принадлежит коду V тогда и только тогда, когда выполнено соотношение

$$\bar{v} \cdot H^{\mathrm{T}} = 0.$$

Здесь H^{T} — матрица, транспорированная к матрице H.

Существует простой алгоритм для нахождения матрицы G по матрице H, и наоборот.

Пример 4. Рассмотрим групповой (5,3) код V с порождающей матрицей

$$G = \left\| \begin{matrix} 10010 \\ 01010 \\ 00101 \end{matrix} \right\|$$

Этот код содержит $2^3 = 8$ векторов, каждый из которых есть сумма некоторых строк матрицы G. Проверочной матрицей для кода V служит следующая матрица H:

$$H = \left\| \begin{matrix} 11010 \\ 10101 \end{matrix} \right\|$$

Вектор $\bar{v}=(v_1\ v_2 v_3 v_4 v_5)$ принадлежит коду V тогда и только тогда, когда выполнены следующие два соотношения:

$$v_1 \oplus v_2 \oplus v_4 = 0;$$
$$v_1 \oplus v_3 \oplus v_5 = 0.$$

Нетрудно видеть, что все векторы кода V удовлетворяют этим соотношениям. Чтобы убедиться в этом, достаточно в выписанные соотношения подставить строки матрицы G.

В общем виде соотношение

$$\bar{v} \cdot H^{\mathrm{T}} = 0$$

соответствует $(n-k)$-линейным уравнениям вида:

$$\sum_{j=1}^{n} v_j h_{ij} = 0, \quad i = 1, 2, \ldots n-k,$$

где $H=(h_{ij})$.

Эти уравнения называются проверками на четность.

Можно показать, что путем сложения строк и перестановки столбцов порождающую матрицу любого (n, k)-кода V можно представить в следующей приведенно-ступенчатой форме

$$G' = \left\| \begin{matrix} 10 \ldots 0 & P_{11} \ldots P_{1,\,n-k} \\ 01 \ldots 0 & P_{21} \ldots P_{2,\,n-k} \\ \ldots\ldots\ldots & \ldots\ldots\ldots \\ 00 \ldots 1 & P_{k1} \ldots P_{k\ n-k} \end{matrix} \right\|$$

При этом матрицам G и G' соответствует один и тот же код V. Так как каждый вектор кода V есть сумма некоторого числа строк матрицы G' и, что то же самое, сумма всех ее

строк, взятых с коэффициентами 0 или 1, то каждый кодовый вектор $\overline{v} \in V$ можно представить в следующем виде:

$$\overline{v}=(a_1\, a_2 \ldots a_k\; c_1 c_2 \ldots c_{n-k}).$$

Здесь a_i принимают произвольно и независимо значения 0 или 1, а

$$c_j = \sum_{i=1}^{k}{}' a_i p_{ij}.$$

Таким образом, первые k символов кодового вектора могут быть выбраны произвольно, а следующие n—k-символов являются линейными комбинациями с коэффициентами 0 и 1 первых символов.

Эти первые k символов называются информационными, а последующие $(n-k)$-избыточными, или проверочными. Код V, порождающая матрица G которого задана в приведенно-ступенчатой форме, называется систематическим. В случае систематического кода избыточные символы имеют простой и наглядный смысл — они просто являются линейными комбинациями информационных символов.

Пример 5. Рассмотрим код V с порождающей матрицей

$$G=\left\|\begin{matrix}10010\\11001\\11100\end{matrix}\right\|$$

Проделаем с матрицей G следующие операции: сначала первую строчку прибавим ко второй и третьей, а потом в полученной матрице прибавим вторую строку к третьей. Тогда последовательно получим:

$$G=\left\|\begin{matrix}10011\\11001\\11100\end{matrix}\right\| \rightarrow \left\|\begin{matrix}10011\\01010\\01111\end{matrix}\right\| \rightarrow G'=\left\|\begin{matrix}10011\\01010\\00101\end{matrix}\right\|$$

Каждый вектор кода V имеет три информационных символа и два проверочных и может быть представлен в следующем виде:

$$\overline{v}=(a_1 a_2 a_3 c_1 c_2),$$

где

$$c_1=a_1 \oplus a_2;$$
$$c_2=a_1 \oplus a_3.$$

Придавая символам a_1, a_2, a_3 произвольные значения 0 или 1, мы, используя эти формулы, получим все кодовые векторы.

Групповые коды, кроме простого описания, имеют еще ряд полезных свойств, к обсуждению которых мы приступаем.

В первую очередь отметим следующее обстоятельство. Так как $\rho(\bar{x}, \bar{y}) = \|\bar{x} \oplus \bar{y}\|$ и для группового кода V из условия $\bar{x}, \bar{y}, \in V$ следует, что $\bar{x} \oplus \bar{y} \in V$, то кодовое расстояние в групповом коде равно просто минимальному весу кодового вектора, т. е.

$$d(V) = \min_{\bar{x} \in V} \|\bar{x}\|. \tag{6}$$

При этом минимум берется по всем векторам кода V, отличным от нулевого. Таким образом, если в случае произвольного кода V для нахождения величины $d(V)$ мы должны сравнить $C^2_{|V|}$ попарных расстояний, то для группового кода это число сравнений равно просто $|V|$.

Однако не это обстоятельство является самым важным для оценки кодового расстояния. Оказывается, что корректирующую способность групповых кодов можно довольно просто выразить в терминах линейной зависимости столбцов проверочной матрицы, что в ряде случаев дает возможность эффективного построения хороших кодов.

Пусть V-групповой (n, k) код с проверочной матрицей H и $h_1 h_2 \ldots h_n$ — столбцы этой матрицы. Тогда если любая совокупность из $(w-1)$ или меньшего числа столбцов матрицы H линейно независима, то кодовое расстояние V не меньше чем w.

Действительно, пусть это условие выполнено. Пусть $\bar{v} = (v_1\ v_2 \ldots v_n) \in V$. Тогда

$$\bar{v} \cdot H^T = 0$$

или, записывая это условие через столбцы матрицы H, получаем

$$\sum_{i=1}^{n} v_i h_i = 0. \tag{7}$$

Эта сумма соответствует линейной комбинации столбцов $\{h_1 h_2 \ldots h_n\}$. При этом число столбцов, в действительности участвующих в этой комбинации, равно числу ненулевых коэффициентов v_i или, другими словами, равно

весу вектора $\bar{v}$. Если предыдущее условие выполнено и вектор $\bar{v}$ ненулевой, то равенство (7) возможно лишь в случае, если вес этого вектора не менее чем w. Учитывая соотношение (6), получаем, что $d(V) \geqslant w$.

Пример 6. В проверочной матрице H (5, 3)-кода V из примера 4 есть два одинаковых столбца: 2-й и 4-й. Поэтому $w-1=1$. Отсюда следует, что $d(V) \geqslant 2$. На самом деле $d(V)=2$, так как, например, вторая строка в матрице G имеет вес два.

Простейшим, но важным следствием приведенного выше результата является следующее утверждение: если все столбцы проверочной матрицы H кода V различны, то кодовое расстояние V не меньше, чем три. Действительно, если все столбцы матрицы H различны, то сумма никаких двух из них не равна нулю, т. е. любые два столбца H линейно независимы. Отсюда $d(V) \geqslant 3$.

Вообще, как показывает приведенное выше условие, для построения кода с большой корректирующей способностью достаточно построить матрицу H, у которой сумма любых двух, трех, четырех и т. д. столбцов отлична от нулевого вектора.

Пример 7. Рассмотрим матрицу H:

$$H = \begin{Vmatrix} 10011 \\ 01010 \\ 00101 \end{Vmatrix}$$

Нетрудно убедиться в том, что в матрице H сумма никаких двух столбцов не равна нулю. Поэтому (5,2)-код V с проверочной матрицей H имеет кодовое расстояние не меньше, чем три. Код V имеет порождающую матрицу G:

$$G = \begin{Vmatrix} 10101 \\ 01011 \end{Vmatrix}$$

и состоит из следующих векторов:

$$V = \begin{Vmatrix} 10101 \\ 01011 \\ 11110 \\ 00000 \end{Vmatrix}$$

Таким образом, $d(V)=3$.

Кроме кодового расстояния, важной метрической характеристикой множества V, полностью описывающей его

корректирующие способности, является так называемый спектр этого множества. Пусть $V=\{\bar{x}_1\bar{x}_2 \ldots \bar{x}_m\}$ и $r_s(\bar{x}_i)$ есть число точек кода V, находящихся на расстоянии s от точки $\bar{x}_i$. Тогда последовательность чисел $r(\bar{x}_i)=\{r_1(\bar{x}_i)r_2(\bar{x}_i) \ldots r_n(\bar{x}_i)\}$ называется спектром точки $\bar{x}_i$.

Ясно, что спектр точки $\bar{x}_i$ показывает, сколько точек кода V находится от $\bar{x}_i$ на расстоянии 1, сколько на расстоянии 2 и т. д.

Рассмотрим матрицу $A=(r_i(\bar{x}_j))$. Ясно, что строками этой матрицы являются спектры всех точек кода V. В случае, когда множество V является групповым (n, k)-кодом, спектр каждой точки один и тот же, т. е. не зависит от самой точки. В частности, если $\bar{x}=(0, 0, \ldots, 0)\in V$, то $r_s(\bar{x})=a_s$, где a_s — число точек кода V, имеющих вес, равный s. В этом случае все строки матрицы одинаковы, и поэтому она задается одной строкой $r=(a_0a_1 \ldots a_n)$. Последовательность $r=(a_0a_1 \ldots a_n)$ называется весовым спектром, или просто спектром группового кода V. Все метрические характеристики этого кода могут быть более или менее сложным способом выражены в терминах его спектра. В частности,

$$d(V)=\min_{\substack{a_i\neq 0\\ i\geqslant 1}} i.$$

Пример 8. Спектр кода V, описанного в предыдущем примере, имеет следующий вид: $r=(1, 0, 0, 2, 1, 0)$.

В дальнейшем мы рассмотрим еще ряд примеров на вычисление спектров конкретных групповых кодов. Вообще задача вычисления спектра группового (n, k)-кода по его порождающей или проверочной матрице является очень актуальной и довольно сложной. Существенным вспомогательным моментом, облегчающим решение этой задачи, является полученная американским математиком Дж. Мак-Вильямс формула, связывающая спектры кода и двойственная к нему.

Итак, пусть V-групповой (n, k)-код, имеющий спектр $r=(a_0a_1 \ldots a_n)$ и V^* — двойственный к нему $(n, n-k)$-код, имеющий спектр $r^*=(b_0b_1 \ldots b_n)$. Формула Мак-Вильямс устанавливает связь между r и r^* в следующем виде:

$$\sum_{i=0}^{n} b_i(1+z)^{n-i}(1-z)^i = 2^{n-k}\sum_{i=0}^{n} a_iz^i. \qquad (8)$$

С помощью этой формулы, зная спектр любого из кодов V или V^*, можно вычислить спектр другого. Действительно, пусть

$$(1+z)^{n-i}(1-z)^i = \sum_{s=0}^{n} \varphi_s(n, i)\, z^s,$$

где коэффициенты $\varphi_s(n, i)$ могут быть вычислены по следующим формулам:

$$\varphi_s(n, i) = \sum_{j=0}^{n} (-1)^j C_i^j C_{n-i}^{s-j}.$$

Тогда

$$\sum_{i=0}^{n} b_i (1+z)^{n-i}(1-z)^i = \sum_{i=0}^{n} b_i \sum_{s=0}^{n} \varphi_s(n, i)\, z^s =$$

$$= \sum_{s=0}^{n} z^s \sum_{i=0}^{n} b_i \varphi_s(n, i)$$

или, используя формулу Мак-Вильямс, получаем:

$$\sum_{s=0}^{n} z^s \sum_{i=0}^{n} b_i \varphi_s(n, i) = 2^{n-k} \sum_{s=0}^{n} a_s z^s.$$

Отсюда, приравнивая коэффициенты при одинаковых степенях z, выводим:

$$a_s = \frac{1}{2^{n-k}} \sum_{i=0}^{n} b_i \varphi_s(n, i),$$

что и позволяет вычислить спектр r по спектру r^*. Аналогичным образом можно установить формулу для вычисления спектра r^* по r.

Формула Мак-Вильямс оказывается полезной при значениях k, близких к n, так как в этом случае разность $n - k$ мала, и спектр двойственного кода можно вычислить непосредственно. Однако польза этой формулы идет гораздо дальше этого утилитарного применения.

Кроме формулы (8), полезным является и другое соотношение, также принадлежащее Мак-Вильямс:

$$\sum_{i=0}^{n} b_i \varphi_s(n, i)(1+z)^{n-i}(1-z)^i = 2^{n-k} \sum_{r=0}^{n} C(r, s)\, z^r. \quad (9)$$

Здесь $C(r, s)$ — число точек из E^n, имеющих вес r и находящихся на расстоянии s от некоторой точки кода V. Это тождество справедливо при всех $s \leqslant t$, где t — число ошибок, исправляемых кодом V.

Из этого тождества можно вывести другое полезное равенство, которое нам понадобится в дальнейшем. Пусть V — групповой (n, k)-код, исправляющий t ошибок, и $C(s)$ — число точек из E^n веса s, попавших хотя бы в один шар радиуса t с центром в кодовой точке. Тогда справедливо тождество:

$$\sum_{i=0}^{n} b_i \varphi_t(n-1, i-1)(1+z)^{n-i}(1-z)^i = \\ = 2^{n-k} \sum_{s=0}^{n} C(s) z^s . \qquad (10)$$

Здесь $\{b_0 b_1 \ldots b_n\}$ — спектр кода V^*, двойственного коду V. Числа $C(s)$ показывают, сколько точек веса s покрыто шарами радиуса t с центрами в кодовых точках.

II. *Коды Хэмминга.* Как мы уже отмечали, в случае $d(V)=2$ максимальным по мощности кодом с расстоянием 2 является, например, «счетчик четности», т. е. множество всех двоичных последовательностей с четным числом единичных координат. Это множество состоит из 2^{n-1} точек, и, как показывает неравенство Джоши, большего достичь нельзя.

Следующий случай: $d(V) = 3$. Любой код с $d(V) = 3$ способен исправлять любую единичную ошибку. Из неравенства Хэмминга следует, что число точек в таком коде не превышает следующей величины: $|V| \leqslant \frac{2^n}{n+1}$.

Выше мы уже отмечали, что для того чтобы групповой код имел кодовое расстояние 3, достаточно потребовать, чтобы в его проверочной матрице все столбцы были различны. Поэтому если мы возьмем в качестве матрицы H такую матрицу, у которой столбцами будут все $(2^m - 1)$-ненулевых двоичных наборов длины m (m — любое натуральное число), то код V с этой проверочной матрицей будет исправлять одну ошибку. Вычислим параметры этого кода.

Так как число столбцов в матрице H равно $n = 2^m - 1$, то длина кодовых последовательностей равна $2^m - 1$. Число строк в матрице H равно m, т. е. порождающая матрица двойственного кода V^* имеет m строк. Поэтому порождающая матрица самого кода V имеет $n - m$ строк,

т. е. число точек кода V равно 2^{n-m}. Так как $n = 2^m - 1$, то

$$|V| = 2^{2^m - m - 1} = \frac{2^{2^m - 1}}{2^m} = \frac{2^n}{n+1}.$$

Сравнивая это соотношение с границей Хэмминга, мы заключаем, что построенный нами код является совершенным, т. е. шары радиуса единица с центрами в кодовых точках покрывают без пересечений все вершины куба E^n.

Описанный нами код был впервые открыт американским математиком Р. В. Хэммингом и носит его имя. Коды Хэмминга обладают многими замечательными свойствами, важнейшим из которых является простота декодирования.

Рассмотрим способ декодирования, открытый самим Р. В. Хэммингом. Так как код V не зависит от порядка расположения столбцов в проверочной матрице H, то это расположение можно выбрать следующим образом: в качестве i-го столбца матрицы H берется двоичное представление числа i. В случае $m = 3$ это расположение выглядит следующим образом:

$$H = \begin{Vmatrix} 0001111 \\ 0110011 \\ 1010101 \end{Vmatrix}$$

Действительно, 1 = 001, 2 = 010, 3 = 011 и т. д.

Предположим, что по каналу передавался некоторый вектор $\bar{v}$ и в результате одиночной ошибки этот вектор перешел в вектор $\bar{u}$, отличающийся от $\bar{v}$ ровно в одном разряде. Этот вектор можно представить в следующем виде:

$$\bar{u} = \bar{v} \oplus \bar{e}.$$

Здесь $\bar{e} = (0,0, \ldots, 1, 0, \ldots, 0)$ — вектор, содержащий единицу на том месте, где произошла ошибка, и содержащий нули на остальных местах.

Вычислим так называемый синдром, т. е. произведение вектора $\bar{u}$ на H^T. Имеем:

$$\bar{u} \cdot H^T = (\bar{v} \oplus \bar{e}) \cdot H^T = \bar{v} \cdot H^T \oplus \bar{e} \cdot H^T = \bar{e} \cdot H^T,$$

так как $\bar{v} \in V$. Отсюда следует, что синдром вектора $\bar{u}$ в точности равен тому столбцу матрицы H, номер которого соответствует номеру ошибки в кодовом слове $\bar{v}$. Но мы расположили столбцы в матрице H таким способом, что

каждый столбец есть двоичное представление его номера. Поэтому вектор $\bar{e} \cdot H^T$ есть двоичное представление номера ошибки. Зная номер ошибки, мы изменяем значение соответствующего разряда и тем самым исправляем ошибку.

Пример 9. Пусть вектор $\bar{v}$ кода V, описанного выше, выбран следующим: $\bar{v} = (0, 1, 1, 1, 1, 0, 0)$. Предположим, что на выходе принят вектор $\bar{u}$ с ошибкой в 5-м символе, т. е. $\bar{u} = (0, 1, 1, 1, 0, 0)$. Имеем:

$$\bar{u} \cdot H^T = (0,1,1,1,0,0,0) \cdot \begin{Vmatrix} 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{Vmatrix} = (1,0,1)$$

Но $101 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 5$, т. е. ошибка произошла в 5-м разряде. Теперь кодовое слово можно правильно восстановить, изменив пятый символ в слове $\bar{u}$.

Возникает еще вопрос о явном описании самого кода V. По самому определению вектор $\bar{v}$ принадлежит коду V тогда и только тогда, когда выполнено соотношение:

$$\bar{v} \cdot H^T = 0.$$

При $\bar{v} = (v_1 v_2 v_3 v_4 v_5 v_6 v_7)$ и $m = 3$ это соотношение эквивалентно следующим трем уравнениям:

$$\begin{aligned} v_4 \oplus v_5 \oplus v_6 \oplus v_7 &= 0 \\ v_2 \oplus v_3 \oplus v_6 \oplus v_7 &= 0 \\ v_1 \oplus v_3 \oplus v_5 \oplus v_7 &= 0. \end{aligned}$$

Если выбрать в качестве проверочных 1-й, 2-й и 4-й символы, то каждое кодовое слово можно представить в следующем виде: $v = (p_1 p_2 v_3 p_4 v_5 v_6 v_7)$. Информационным символам $v_3 v_5 v_6$ и v_7 можно произвольно и независимо придавать любые значения, а проверочные символы p_1, p_2 и p_4 выбираются таким способом, чтобы были удовлетворены приведенные выше три уравнения. В частности, если $v_3 = 1$, $v_5 = 1$, $v_6 = 0$ и $v_7 = 0$, то $p_4 = 1$, $p_2 = 1$, $p_1 = 0$, и мы получаем вектор $\bar{v}$, использованный выше.

Коды Хэмминга являются в высшей степени конструктивными, так как для их задания требуется знать лишь

число m-проверочных символов кода. Вычислим теперь спектр $(2^m - 1,\ 2^m - m - 1)$-кода Хэмминга.

Так как код Хэмминга является совершенным, то любая точка куба E^n попадает в шар радиуса единица с центром в какой-нибудь кодовой точке. Поэтому введенная выше величина $C(s)$ — число точек E^n веса s, попавших хотя бы в один шар радиуса t (в нашем случае $t = 1$) с центром в кодовой точке, равна общему числу точек веса s, т. е. C_n^s. Отсюда с использованием формулы (10) получаем:

$$\sum_{i=0}^{n} b_i \varphi_1 (n-1, i-1)(1+z)^{n-i}(1-z)^i = \\ = 2^{n-k} \sum_{s=0}^{n} C_n^s z^s = 2^{n-k}(1+z)^n. \qquad (11)$$

Можно показать, что $\varphi_1 (n-1, -1) = 2^{n-k}$. Отсюда

$$2^{n-k}(1+z)^n + \sum_{i=1}^{n} b_i \varphi_1 (n-1, i-1)(1+z)^{n-i}(1-z)^i = \\ = 2^{n-k}(1+z)^n, \qquad (12)$$

т. е.

$$\sum_{i=1}^{n} b_i \varphi_1 (n-1, i-1)(1+z)^{n-i}(1-z)^i = 0.$$

Так как в левой части стоит тождественно равный нулю полином, то все его коэффициенты равны нулю, т. е.

$$b_i \varphi_1 (n-1,\ i-1) = 0. \qquad i = 1, 2, \ldots, n.$$

Но

$$\varphi_1 (n-1, i-1) = \sum_{j=0}^{n-1} (-1)^j C_{i-1}^j C_{n-i}^{i-1} = n - i - (i-1) = \\ = n - 2i + 1. \qquad (13)$$

Отсюда следует, что $\varphi_1 (n-1,\ i-1)$ равен нулю лишь при $i = \frac{n+1}{2}$. Поэтому

$$b_1 = b_2 \ldots = b_{\frac{n-1}{2}} = b_{\frac{n+3}{2}} \ldots = b_n = 0.$$

Эти равенства показывают, что в двойственном коде V^* все ненулевые векторы имеют одинаковый вес, равный $\frac{n+1}{2} = 2^{m-1}$.

Из формулы Мак-Вильямс теперь получаем:

$$\sum_{i=0}^{n} a_i z^i = \frac{1}{2^{n-k}} \sum_{i=0}^{n} b_i (1+z)^{n-i}(1-z)^i =$$

$$= \frac{1}{n+1}\left[(1+z)^n + n(1+z)^{\frac{n-1}{2}}(1-z)^{\frac{n+1}{2}}\right]. \qquad (14)$$

Из этой формулы можно получить явные выражения для чисел a_i, т. е. найти спектр кода Хэмминга. Ниже мы используем этот результат для нахождения средней вероятности необнаружения ошибки этого кода.

III. *Эквидистантные коды.* Изучая $(n, 3)$-код Хэмминга, мы нашли другой интересный код V^*, являющийся двойственным к коду Хэмминга. Этот код содержит ровно $2^{2^m-1-(2^m-m-1\ldots)} = 2^m$ точек, а кодовое расстояние есть $\frac{n+1}{2}$. Как показывает неравенство Плоткина, при $d = \frac{n+1}{2}$ никакой код с расстоянием $d = \frac{n+1}{2}$ не может содержать более $(n+1)$-й точки. Поэтому код V^* является максимальным по числу точек в классе всех кодов с расстоянием $\frac{n+1}{2}$. Так как каждая ненулевая точка кода V^* имеет вес $\frac{n+1}{2}$, то все попарные расстояния между точками кода V^* одинаковы и равны этому весу. Таким образом, код V^* является представителем класса так называемых эквидистантных кодов, т. е. кодов, у которых все попарные расстояния между точками одинаковы. Простейшим примером такого кода в E^n является следующее множество точек:

$$\bar{\alpha}_1 = (1\,0\,0\ldots 0)$$
$$\bar{\alpha}_2 = (0\,1\,0\ldots 0)$$
$$\ldots\ldots$$
$$\bar{\alpha}_n = (0\,0\,0\ldots 1)$$

Расстояние между любыми двумя точками из этого множества равно двум. Нетрудно видеть, что геометрически это множество представляет собой сферу радиуса единица с центром в нулевой точке. Ясно, что все векторы построенного кода получаются из первого последовательным «сдвигом» единицы из 1-го разряда во 2-й, из 2-го в 3-й и т. д. Используя аналогичный прием, можно постро-

ить другой эквидистантный код, взяв за начальную последовательность последовательность с 2-мя единицами:

$$\bar{\beta}_1 = (1\ 1\ 0\ 0\ \ldots\ 0)$$
$$\bar{\beta}_2 = (0\ 0\ 1\ 1\ \ldots\ 0)$$
$$\ldots\ldots\ldots\ldots$$
$$\beta^n/_2 = (0\ 0\ 0\ 0\ \ldots\ 1\ 1)$$

Так мы получим эквидистантный код с расстоянием 4, состоящий из $m = \frac{n}{2}$ точек в случае четного n и $m = \left[\frac{n}{2}\right]$ в случае нечетного.

Мы доказали, что код V^*, двойственный к коду Хэмминга, является эквидистантным и содержит $(n+1)$-точку. Далее, использовав неравенство Плоткина, мы установили, что этот код имеет максимальное число точек среди всех кодов с расстоянием $\frac{n+1}{2}$. Однако этот результат на самом деле не связан с расстоянием в коде V^*, а является лишь следствием геометрической природы этого кода — его эквидистантностью. На самом деле имеет место такой факт: если в n-мерном пространстве имеется некоторое число точек, попарные расстояния между которыми одинаковы, то число этих точек не превосходит $(n+1)$. На плоскости таким эквидистантным множеством является правильный треугольник, в 3-мерном пространстве — тетраэдр и т. д.

Таким образом, число точек любого эквидистантного кода в E^n не превосходит $n+1$ для $n = 2, 3$. Докажем этот факт для произвольного n. Итак, пусть $V \subseteq E^n$, $V = \{\bar{\alpha}_1\ \bar{\alpha}_2 \ldots \bar{\alpha}_s\}$ и $\rho(\bar{\alpha}_i, \bar{\alpha}_j) = d$.
$i, j = 1, 2, \ldots, s$.

Будем считать, что нулевой вектор принадлежит коду V, иначе можно перейти к коду $V \oplus \bar{\alpha}_1$, который также является эквидистантным и уже содержит нулевой вектор. Пусть этим вектором будет $\bar{\alpha}_1$. Так как

$$\rho(\bar{\alpha}_1, \bar{\alpha}_j) = \| \bar{\alpha}_1 \oplus \bar{\alpha}_j \| = \| \bar{\alpha}_j \| = d,$$
$$j = 2, 3, \ldots, s,$$

то каждый ненулевой вектор из V имеет вес d. Далее, так как

$$\rho(\bar{\alpha}, \bar{\beta}) = \sum_{i=1}^{n} |\alpha_i - \beta_i| = \sum_{i=1}^{n} (\alpha_i - \beta_i)^2,$$

где $\bar{\alpha} = (\alpha_1\, \alpha_2 \ldots \alpha_n)$ и $\bar{\beta} = (\beta_1\, \beta_2 \ldots \beta_n)$, то

$$\rho(\bar{\alpha}, \bar{\beta}) = \| \bar{\alpha} \| + \| \bar{\beta} \| - 2(\bar{\alpha}, \bar{\beta}),$$

где $(\bar{\alpha}, \bar{\beta}) = \sum_{i=1}^{n} \alpha_i \beta_i$ — скалярное произведение векторов $\bar{\alpha}$ и $\bar{\beta}$. В частности, если $\bar{\alpha}_i, \bar{\alpha}_j \in V$, то

$$\rho(\bar{\alpha}_i, \bar{\alpha}_j) = 2d - 2(\bar{\alpha}_i, \bar{\alpha}_j) = d,$$

т. е.

$$(\bar{\alpha}_i, \bar{\alpha}_j) = \frac{d}{2}.$$

Покажем теперь, что векторы $\bar{\alpha}_2 \bar{\alpha}_3 \ldots \bar{\alpha}_s$ линейно независимы. Для этого составим матрицу Грама $A = \| (\bar{\alpha}_i \bar{\alpha}_j) \|$ и вычислим ее определитель. Имеем:

$$A = \|(\bar{\alpha}_i, \bar{\alpha}_j)\| = \begin{Vmatrix} d & \frac{d}{2} & \frac{d}{2} & \ldots & \frac{d}{2} \\ \frac{d}{2} & d & \frac{d}{2} & \ldots & \frac{d}{2} \\ & & & \ldots & \\ \frac{d}{2} & \frac{d}{2} & \frac{d}{2} & \ldots & d \end{Vmatrix}$$

Определитель матрицы A легко вычислить и получить

$$|A| = \frac{d^k (k+1)}{2^k}, \tag{15}$$

где $k = s - 1$.

Так как $|A| \neq 0$, то векторы $\bar{\alpha}_2, \bar{\alpha}_3, \ldots, \bar{\alpha}_s$ линейно независимы. Поэтому их число не превосходит размерности пространства E^n. Отсюда следует, что $s - 1 \leqslant n$, т. е. $s \leqslant n + 1$, и наше утверждение доказано.

Замечание. Нетрудно убедиться в том, что расстояние в эквидистантном коде обязательно должно быть четным числом (15).

Код V^*, двойственный к коду Хэмминга, носит название кода Макдональда. Он является одним из представителей класса эквидистантных кодов, которые можно получить с помощью так называемых матриц Адамара.

Матрицей Адамара называется квадратная матрица порядка n, состоящая из $+1$, -1, строки которой попарно

ортогональны. Простейшим примером такой матрицы является следующая 2×2-матрица:

$$H_2 = \begin{Vmatrix} 1 & 1 \\ +1 & -1 \end{Vmatrix}$$

Связь этих матриц с эквидистантными кодами описывается следующим утверждением.

Т е о р е м а. Если существует матрица Адамара порядка n, то существует эквидистантный код из n символов, содержащий ровно n точек.

Действительно, пусть H — матрица Адамара и $v_1 v_2 \ldots v_n$ — строки этой матрицы. Каждой строке v_i сопоставим двоичный набор $\bar{v}_i$ следующим образом: $+1$ заменяем на 0, а -1 — на 1. Таким образом, мы получаем код $V = \{\bar{v}_1 \bar{v}_2 \ldots \bar{v}_n\}$. Покажем теперь, что код V является эквидистантным.

Пусть $v_i = (v_i^1 v_i^2 \ldots v_i^n)$, $v_j = (v_j^1 v_j^2 \ldots v_j^n)$ и

$$\rho_k = \begin{cases} 1, \text{ если } v_i^k \neq v_j^k; \\ 0, \text{ если } v_i^k = v_j^k. \end{cases}$$

Тогда

$$(v_i, v_j) = \sum_{k=1}^{n} v_i^k v_j^k = \sum_{k=1}^{n} (-1)^{\rho_k} = n - 2\rho(\bar{v}_i, \bar{v}_j). \quad (16)$$

Так как строки v_i и v_j ортогональны, то

$$(v_i v_j) = 0.$$

Из (16) теперь получаем:

$$\rho(\bar{v}_i, \bar{v}_j) = \frac{n}{2}, \quad i, j = 1, 2, \ldots, n$$

что и доказывает эквидистантность кода V.

Так как расстояние в эквидистантном коде четно (см. предыдущее замечание), то число $\frac{n}{2}$ должно быть четным, т. е. n кратно четырем. Таким образом, порядок матрицы Адамара кратен четырем. В этом состоит необходимое условие существования матриц Адамара порядка n. Известна старая гипотеза о том, что приведенное условие является и достаточным. Но это предположение не доказано, и, более того, не видно никаких реальных шансов справиться с ним в ближайшее время.

Однако существует некоторое число методов, позволяющих конструировать матрицы Адамара. Простейший из этих методов позволяет строить матрицы Адамара порядка $2n$ из матриц Адамара порядка n. Сущность этого метода состоит в следующем.

Теорема. Если H_n-матрица Адамара порядка n, то матрица

$$H_{2n} = \begin{Vmatrix} H_n & H_n \\ H_n & -H_n \end{Vmatrix}$$

является матрицей Адамара порядка $2n$.

Применив эту теорему к матрице H_2, мы получаем матрицу

$$H_4 = \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & +1 \end{Vmatrix}$$

Этот метод позволяет для любого k построить матрицу Адамара порядка 2^k.

В настоящее время матрицы Адамара построены для всех n, кратных четырем и не превосходящих 200, кроме $n = 188$. Кроме этого, матрицы Адамара построены для многих бесконечных последовательностей значений n.

В случае матриц Адамара мы имеем еще один яркий пример использования старого понятия классической математики в области, на первый взгляд чрезвычайно далекой от той, которая привела к возникновению этого понятия.

IV. *Совершенные коды.* При обсуждении границы Хэмминга мы естественным образом пришли к понятию совершенного кода. Точки совершенного кода интерпретируются как центры шаров радиуса t, на которые разбиваются без пересечений вершины множества E^n. Первым встретившимся нам примером совершенного кода является тривиальный код, состоящий из двух «противоположных» точек куба E^n: $\bar{\alpha} = (0, 0, \ldots, 0)$, $\bar{\beta} = (1, 1, \ldots, 1)$. Этот код является совершенным кодом с расстоянием n при нечетном n.

Действительно, если точка $\bar{x} \in E^n$ имеет больше половины единичных координат, то она принадлежит шару радиуса $t = \frac{n-1}{2}$ с центром в точке $\bar{\beta}$. Если же $\bar{x}$ имеет меньше половины единичных координат, то она принадлежит такому же шару с центром в точке $\bar{\alpha}$. Таким образом,

куб E^n разбивается на два непересекающихся шара радиуса $\frac{n-1}{2}$ с центрами в точках $\bar{\alpha}$ и $\bar{\beta}$.

Следующим, уже нетривиальным, примером совершенного кода был код Хэмминга. Этот код существует для всех чисел n вида $n = 2^m - 1$; он имеет расстояние 3, т. е. исправляет одну ошибку. Число точек в коде Хэмминга равно:

$$\frac{2^n}{n+1} = \frac{2^{2^m-1}}{2^m} = 2^{2^m-m-1}.$$

Код Хэмминга является групповым $(2^m - 1, 2^m - m - 1)$-кодом. Остальные свойства этого кода были описаны выше.

Следующим и последним двоичным совершенным кодом является код Голея. Этот код является групповым (23, 12)-кодом с исправлением трех ошибок.

Отметим, что вместе с совершенным кодом V совершенным является и код $V \oplus \bar{\alpha}$ для любой точки $\bar{\alpha} \in E^n$. Возникает вопрос, существуют ли совершенные коды с расстоянием 3, которые не могут быть получены из кодов Хэмминга «сдвигом» на какую-нибудь точку $\bar{\alpha} \in E^n$. Ответ на этот вопрос оказывается положительным [4]. Сконструировано целое семейство так называемых «сильно негрупповых» совершенных кодов, каждые два из которых не переводятся друг в друга посредством преобразования «сдвига». Этот факт указывает на то обстоятельство, что даже такое «сверх симметричное» расположение, как совершенный код, может быть реализовано многими различными способами.

Перечисленными классами совершенных кодов исчерпываются все двоичные совершенные коды. Этот факт сравнительно недавно был установлен в работах [2] и [3].

Оказалось, что решающую роль в проблеме существования совершенных кодов играют корни полинома $\varphi_t(n-1, x)$. Действительно, в случае групповых совершенных кодов из формулы (10) получаем:

$$\sum_{i=0}^{n} b_i \varphi_t(n-1, i-1)(1+z)^{n-i}(1-z)^i = 2^{n-k}(1+z)^n,$$

так как в совершенном коде $C(s) = C_n^s$.

Аналогично формулам (12) и (13) отсюда получаем:

$$\sum_{i=1}^{n} b_i \varphi_t (n-1, i-1)(1+z)^{n-i} (1-z)^i = 0. \qquad (18)$$

Из этого равенства вытекают следующие соотношения:

$$b_i \varphi_t (n-1, i-1) = 0, \quad i = 1, 2, \ldots, n. \qquad (19)$$

Таким образом, справедливо следующее утверждение.

Т е о р е м а. Для того чтобы групповой (n, k)-код V с исправлением t ошибок был совершенным, необходимо и достаточно выполнение следующей альтернативы: либо число $i \leqslant n$ есть корень уравнения $\varphi_t (n-1, i) = 0$, либо в двойственном коде нет векторов веса $i-1$.

Д о к а з а т е л ь с т в о. Необходимость условия теоремы следует из равенств (19). Докажем теперь его достаточность.

Итак, пусть V — (n, k)-код с исправлением t ошибок и $\{b_0 b_1 \ldots b_n\}$ — спектр двойственного кода V^*. Тогда из условия теоремы следуют равенства:

$$b_i \varphi_t (n-1, i-1) = 0.$$

Из этих уравнений получаем:

$$\sum_{i=0}^{n} b_i \varphi_t (n-1, i-1)(1+z)^{n-i} (1-z)^i = b_0 \varphi_t (n-1, -1) \times$$

$$\times (1+z)^n. \qquad (20)$$

Сравнивая (17) и (20) при $z = 0$, получаем:

$$\varphi_t (n-1, -1) = 2^{n-k}. \qquad (21)$$

Но

$$\varphi_t (n-1, -1) = \sum_{i=0}^{t} C_n^i,$$

т. е.

$$2^k = 2^n \Big| \sum_{i=0}^{t} C_n^i,$$

откуда и следует совершенность кода V.

Ранее американским математиком Ллойдом была доказана следующая теорема, выражающая необходимое условие существования совершенного кода с исправлением t ошибок, тоже в терминах, связанных с корнями полинома $\varphi_t (n-1, x)$.

Теорема Ллойда. Если существует совершенный код длины n с исправлением t ошибок, то полином $\varphi_t(n-1, x)$ имеет t различных целых корней среди чисел $1, 2, \ldots, n-1$.

Подробное изучение полинома $\varphi_t(n-1, x)$ совместно с некоторыми другими необходимыми условиями существования совершенных кодов позволяет доказать следующее утверждение [2], [3].

Теорема. Единственными нетривиальными совершенными двоичными кодами являются коды Хэмминга и код Голея.

V. *Циклические коды.* Следующим очень важным и богатым классом кодов, исправляющих ошибки, являются циклические коды. Построение циклических кодов основано на следующем соответствии между двоичными векторами длины n и полиномами степени не выше $n-1$ с коэффициентами 0 и 1:

$$\bar{\alpha} = (\alpha_1 \alpha_2, \ldots, \alpha_n) \rightleftarrows f_{\bar{\alpha}}(x) = \alpha_1 + \alpha_2 x \ldots + \alpha_n x^{n-1}.$$

В этом множестве полиномов с коэффициентами 0,1 вводятся обычная операция сложения по mod 2 и операция умножения по mod $(x^n - 1)$. Последнее означает следующее. Перемножив два полинома обычным способом, мы, используя равенство $x^n - 1$, преобразуем полученное произведение в полином степени не выше чем $n-1$.

Пример 10. Пусть $n = 4$, $f(x) = 1 + x^2 + x^3$, $g(x) = x + x^3$. Тогда

$$f(x)g(x) = x + x^3 + x^3 + x^5 + x^4 + x^6 = x + x^4 + x^5 + x^6 = x + 1 + x \cdot x^4 + x^2 \cdot x^4 = 1 + x + x + x^2 = 1 + x^2 \pmod{x^4 + 1}.$$

Введение такого умножения можно также трактовать как определение некоторой новой операции в множестве E^n. Но при этом смысл этой операции оказался бы несколько завуалированным. Множество полиномов с введенными операциями сложения и умножения образуют алгебраическую структуру, носящую название кольца и полиномов над полем из двух элементов $\{0,1\}$. Это кольцо мы будем обозначать через A_n.

Класс циклических кодов задается следующим образом. Берется некоторый полином $f(x) \in A_n$ и рассматриваются все полиномы из A_n, кратные $f(x)$, т. е. все те полиномы, которые можно представить в виде $f(x)g(x)$ при $g(x) \in A_n$. Если полученному множеству полиномов $A(f)$ сопоставить соответствующее ему множество двоичных наборов, то V

и будет циклическим кодом, порожденным полиномом $f(x)$. При этом полином $f(x)$ выбирается среди делителей полинома $x^n + 1$.

Пример 11. Пусть $n = 4$ и $f(x) = 1 + x^2$. Нетрудно видеть, что $1 + x^4 = (1 + x^2)^2$, т. е. полином $f(x)$ является делителем $1 + x^4$. Множество $A(f)$ является подпространством A_4, так как сумма любых двух полиномов, делящихся на $f(x)$, тоже делится на $f(x)$. Это подпространство имеет размерность 2, так как в него входят полиномы 3-й и 4-й степеней. Чтобы его описать, достаточно найти два линейно независимых полинома, делящихся на $f(x)$. В качестве таких полиномов можно взять, например, следующие: сам полином $f(x)$ и полином $f_1(x)$, где

$$f_1(x) = x \cdot f(x) = x(1 + x^2) = x + x^3.$$

Теперь код V, соответствующий множеству $A(f)$, можно задать следующей порождающей матрицей:

$$G = \begin{Vmatrix} f(x) \\ f_1(x) \end{Vmatrix} = \begin{Vmatrix} 1010 \\ 0101 \end{Vmatrix}$$

Название циклических эти коды носят потому, что вместе с каждой последовательностью $\bar{\alpha} = (\alpha_1 \alpha_2 \ldots \alpha_n)$ они содержат ее циклический сдвиг $\bar{\alpha}' = (\alpha_n \alpha_1 \ldots \alpha_{n-1})$. Это легко следует из того, что вместе с полиномом $g(x)$ они содержат и полином $x \cdot g(x)$, а умножение полинома $g(x)$ на x как раз и соответствует циклическому сдвигу последовательности, соответствующей $g(x)$, на единицу вправо. Если $g(x) = f_1(x) = x + x^3$, как в предыдущем примере, то

$$g(x) \to (0\ 1\ 0\ 1)$$

и полиному $x \cdot g(x) = x \cdot (x + x^3) = x^2 + x^4 = 1 + x^2$ соответствует последовательность

$$x \cdot g(x) \to (1\ 0\ 1\ 0),$$

которая как раз и является сдвигом последовательности, соответствующей $g(x)$.

Можно показать, что справедливо и обратное утверждение, т. е. любой групповой код V, который вместе с каждой точкой содержит и ее «сдвиг», состоит из всех полиномов, кратных некоторому делителю полинома $x^n + 1$.

Эквивалентным методом задания циклического кода является способ его задания с помощью корней, порождающего код полинома. Например, код V предыдущего примера можно задать следующим образом: полином $g(x)$

принадлежит коду V тогда и только тогда, когда 1 является его корнем кратности два. Действительно, 1 является корнем полинома $f(x)$ кратности 2, т. е.

$$f(x) = 1 + x^2 = (1 + x)^2.$$

Связь между обоими способами задания циклических кодов устанавливается следующим образом.

Пусть код V задан тем свойством, что все его полиномы имеют корнями различные элементы $\alpha_1\alpha_2 \ldots \alpha_r$. Следует при этом отметить, что сами корни $\alpha_1\alpha_2 \ldots \alpha_r$ чаще всего не задаются в «явном» виде, а также используются один или несколько полиномов, среди корней которых содержатся элементы $\alpha_1\alpha_2 \ldots \alpha_r$.

Пусть $m_i(x)$ — минимальная функция для корня α_i, т. е. полином минимальной степени с коэффициентами 0,1, корнем которого является α_i. Тогда порождающим полиномом кода V является следующий полином $f(x)$:

$$f(x) = \text{н. о. к.}(m_1(x), m_2(x) \ldots m_r(x)).$$

Здесь н. о. к.— наименьшее общее кратное полиномов $m_i(x)$ $(i = 1, 2, \ldots, r)$, т. е. полином минимальной степени с коэффициентами 0 и 1, делящийся на каждый из полиномов $m_i(x)$ $(i = 1, 2, \ldots, r)$. При этом длина или число разрядов кода V определяется следующим образом.

Можно показать, что для любого корня α_i существует такое число $m \geqslant 1$, что выполнено равенство:

$$\alpha_i^m = 1.$$

Наименьшее из этих чисел m называется порядком элемента α_i и обозначается через n_i. Длина n кода V выражается следующей формулой:

$$n = \text{н. о. к. } (n_1, n_2, \ldots, n_r).$$

Наиболее часто встречающимся на практике случаем является такой случай, когда каждый из корней α_i является некоторой степенью одного и того же корня α, т. е.

$$\{\alpha_1\alpha_2 \ldots \alpha_r\} = \{\alpha^{u_1}\alpha^{u_2} \ldots \alpha^{u_r}\}.$$

В этом случае построение кода V ведется по схеме, суть которой можно продемонстрировать на следующем примере.

Пример 12. Пусть α — корень уравнения

$$x^4 + x + 1 = 0$$

и код V задается совокупностью корней $(\alpha_1\alpha^3)$, т. е. любой

«кодовый многочлен» имеет корни α и α^3. Можно доказать, что минимальной функцией для корня α является сам многочлен $f(x) = x^4 + x + 1$. Нам надо найти минимальную функцию для корня $\alpha_1 = \alpha^3$ и длину кода V.

Из тождества $(a + b)^2 = a^2 + b^2$, справедливого в поле из двух элементов 0,1, легко вывести следующее утверждение: если β — корень некоторого полинома $g(x) \in A_n$, то β^2 также является корнем этого полинома.

Действительно, пусть $g(x) = \alpha_1 + \alpha_2 x \ldots + \alpha_n x^{n-1}$, где $\alpha_i = 0,1$. Тогда

$$g(\beta^2) = \sum_{i=1}^{n} \alpha_i (\beta^2)^{i-1} = \sum_{i=1}^{n} \alpha_i^2 (\beta^2)^{i-1} = \sum_{i=1}^{n} (\alpha_i \beta^{i-1})^2 =$$

$$= \left(\sum_{i=1}^{n} \alpha_i \beta^{i-1} \right)^2 = 0,$$

т. е. β^2 — корень $g(x)$ и т. д.

Таким образом, полином $f(x)$ вместе с корнем α имеет также корнями следующие элементы $\alpha^2, \alpha^4, \alpha^8 \ldots$ и т. д.

В дальнейшем нам понадобится следующая таблица, в которой степени корня α выражаются в виде полиномов от α степени не выше 3. Эта таблица составлена исходя из тождества:

$$\alpha^4 = \alpha + 1,$$

которое является следствием того, что α есть корень многочлена $x^4 + x + 1 = 0$.

Имеем:

$\alpha^4 = \alpha + 1;$

$\alpha^5 = \alpha \cdot (\alpha + 1) = \alpha^2 + \alpha;$

$\alpha^6 = \alpha \cdot \alpha^5 = \alpha^3 + \alpha^2;$

$\alpha^7 = \alpha \cdot \alpha^6 = \alpha^4 + \alpha^3 = \alpha^3 + \alpha + 1;$

$\alpha^8 = \alpha \cdot \alpha^7 = \alpha^4 + \alpha^2 + \alpha = 1 + \alpha + \alpha^2 + \alpha = \alpha^2 + 1;$

$\alpha^9 = \alpha + \alpha^3;$

$\alpha^{10} = \alpha^2 + \alpha + 1;$

$\alpha^{11} = \alpha^3 + \alpha^2 + \alpha;$

$\alpha^{12} = \alpha^3 + \alpha^2 + \alpha + 1;$

$\alpha^{13} = \alpha^3 + \alpha^2 + 1;$

$\alpha^{14} = \alpha^3 + 1;$

$\alpha^{15} = 1.$

На степенях от α^5 до α^8 показана сама процедура составления таблицы. Степени от α^9 до α^{15} читатель легко вычислит сам, пользуясь тем же принципом.

Из этой таблицы, в частности, видно, что порядок корня α равен 15, т. е. $n_0 = 15$, так как ни одна степень α, меньшая чем 15, не равна 1. Далее,

$$\alpha^{15} = (\alpha^3)^5 = \alpha_1^5 = 1;$$
$$\alpha_1^2 = (\alpha^3)^2 = \alpha^6 \neq 1;$$
$$\alpha_1^3 = (\alpha^3)^3 = \alpha^9 \neq 1;$$
$$\alpha_1^4 = (\alpha^3)^4 = \alpha^{12} \neq 1,$$

т. е. порядок корня α_1 равен 5. Отсюда

$$n = \text{н. о. к.}(15,5) = 15,$$

т. е. длина кода V равна 15.

Пусть $g(x)$ — минимальная функция для $\alpha_1 = \alpha^3$. Тогда $g(x)$ вместе с α_1 имеет корни: $\alpha_1^2 = \alpha^6$, $\alpha_1^4 = \alpha^{12}$, $\alpha_1^8 = \alpha^{24} = \alpha^{15} \cdot \alpha^9 = \alpha^9$. Поэтому

$$g(x) = (x - \alpha^3)(x - \alpha^6)(x - \alpha^9)(x - \alpha^{12}).$$

Раскрыв скобки, мы получим следующее выражение:

$$g(x) = x^4 + x^3 \cdot (\alpha^3 + \alpha^6 + \alpha^9 + \alpha^{12}) + x^2(\alpha^9 + \alpha^{12} + \alpha^{18} + \alpha^{21}) + x(\alpha^{18} + \alpha^{21} + \alpha^{24} + \alpha^{27}) + \alpha^{30}).$$

При помощи таблицы нетрудно убедиться в том, что каждая сумма в круглых скобках равна 1, т. е.

$$g(x) = x^4 + x^3 + x^2 + x + 1.$$

Так как полиномы $f(x)$ и $g(x)$ не имеют общих множителей, то код V порождается следующим полиномом:

$$m(x) = f(x)\, g(x) = x^8 + x^7 + x^6 + x^4 + 1.$$

Степень полинома $m(x)$ равна 8, т. е. размерность кода V равна $n - 8 = 15 - 8 = 7$, и код V является циклическим (15,7)-кодом.

В качестве порождающего множества кода можно взять точки, соответствующие следующим полиномам:

$$m(x),\ x \cdot m(x),\ x^2 \cdot m(x),\ \ldots,\ x^6 m(x),$$

т. е. порождающую матрицу G кода V можно выбрать в следующей форме:

$$G = \left\| \begin{array}{l} 100010111000000 \\ 010001011100000 \\ 001000101110000 \\ 000100010111000 \\ 000010001011100 \\ 000001000101110 \\ 000000100010111 \end{array} \right\|$$

Коды Боуза — Рой-Чоудхури (*бчх*-коды) задаются следующим образом: полином $f(x)$ принадлежит *бчх*-коду V тогда и только тогда, когда элементы

$$\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{2t}$$

являются его корнями. Здесь α — некоторый специальный (примитивный) корень полинома:

$$g(x) = x^{2^m-1} - 1.$$

Параметры *бчх* кода V определяются следующей теоремой, принадлежащей творцам *бчх*-кодов американским математикам Боузу и Рой-Чоудхури.

Т е о р е м а. Код V имеет мощность $|V| \geqslant \frac{2^{2^m-1}}{2^{m \cdot t}} = \frac{2^n}{2^{m \cdot t}}$

и исправляет не менее t ошибок.

Длина кода V равна $n = 2^m - 1$. Так как при больших m $m \approx \log_2 n$, то выражение для мощности кода V можно записать в следующем виде:

$$|V| \gtrsim \frac{2^n}{n^t}. \qquad (22)$$

Если t — постоянное число, то верхнюю границу Хэмминга для максимальной мощности $A(n, 2t+1)$ кода с исправлением t ошибок можно записать в следующей асимптотической форме:

$$A(n, 2t+1) \lesssim t! \frac{2^n}{n^t}. \qquad (23)$$

Оценки (22) и (23) позволяют определить при постоянном t порядок функции $A(n, 2t+1)$:

$$A(n, 2t+1) \asymp \frac{2^n}{n^t}.$$

Запись означает: с точностью до постоянной, не зависящей от n.

Построенный нами в предыдущем примере код является *бух*-кодом с исправлением двух ошибок. Действительно, каждый полином кода V имеет корни α и α^3. Но если полином $f(x)$ имеет корень α, то, как мы убедились выше, он имеет корень α^2 и α^4. Таким образом, каждый полином кода имеет следующую последовательность корней:

$$\alpha, \alpha^2, \alpha^3, \alpha^4$$

и по теореме Боуза — Рой-Чоудхури исправляет не менее двух ошибок.

Коды Хэмминга также являются частным случаем *бух*-кодов. Они определяются тем условием, что каждый кодовый полином имеет корень α, где α — это примитивный корень многочлена

$$f(x) = x^{2^m-1} + 1.$$

Для *бух*-кодов разработаны специальные алгоритмы декодирования, имеющие сравнительно небольшую трудоемкость.

Мы рассмотрели лишь небольшое число из имеющихся в настоящее время конструктивных кодов. Одно перечисление названий этих кодов составило бы внушительный список: коды Рида — Маллера, коды Рида — Соломона, коды Препарата, коды Файра, коды Васильева, коды Гоппа и т. д. Одни из этих кодов обладают простой схемой декодирования, другие исправляют большое число ошибок, третьи хорошо приспособлены к исправлению определенных типов ошибок и т. д. Многие из этих кодов обладают интересной математической структурой, что делает их изучение важным и самим по себе. Вообще разработка новых классов конструктивных кодов и методов их декодирования представляет важную и бурно развивающуюся ветвь теории кодирования.

§ 4. ВЕРОЯТНОСТНЫЕ КРИТЕРИИ КАЧЕСТВА КОДОВ

Всюду выше мы рассматривали ситуацию, когда для числа ошибок, происходящих в канале, можно указать нетривиальную верхнюю границу, т. е. когда число ошибок в последовательности длины не превосходит заданного числа t.

Однако уже в простейшем *д. с. к.* ситуация несколько другая. Там ошибки могут происходить независимо в каждом символе и вероятность искажения двоичного символа равна p, а вероятность безошибочной передачи символа равна $q = 1 - p$. При этом предполагается, что $q > p$. При этих условиях нетрудно вычислить вероятность того, что какие-нибудь i-символов в последовательности длины n будут искажены. Эта вероятность $p\ (i)$ равна следующему числу:

$$p\ (i) = C_n^i p^i\ q^{n-i}. \tag{24}$$

Действительно, вероятность того, что искажены данные i-символов, равна $p^i q^{n-i}$. А так как эти i-символов могут быть выбраны C_n^i способами, то мы приходим к формуле (24).

В условиях неопределенного числа ошибок важнейшими характеристиками кодов являются следующие: средняя вероятность правильного декодирования и средняя вероятность необнаружения ошибки. К рассмотрению этих характеристик мы и переходим.

1. *Вероятность правильного декодирования.* Как мы уже отмечали в самом начале, процесс исправления ошибок на приемном конце, или, другими словами, процедуру декодирования, удобно описывать с помощью таблицы декодирования. Как только задан код $V = \{\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_s\}$ и задана таблица декодирования A, можно выписать вероятность правильного декодирования, которую мы обозначим через $P\ (V, A)$.

Эта вероятность определяется следующим образом. Обозначим через $P_A\ (\bar{\alpha}_i)$ — вероятность правильного декодирования по таблице A кодового слова $\bar{\alpha}_i$. Тогда

$$P(V, A) = \frac{1}{s} \sum_{i=1}^{s} P_A \cdot (\bar{\alpha}_i). \tag{25}$$

Саму вероятность $P_A\ (\bar{\alpha}_i)$ легко вычислить по таблице декодирования. Сделаем эти вычисления для таблицы на стр. 6. Имеем:

$$\bar{\alpha}_1 = 1\ 1\ 0\ 0\ 0$$
$$\bar{\alpha}_2 = 0\ 0\ 1\ 1\ 0$$
$$\bar{\alpha}_3 = 1\ 0\ 0\ 1\ 1$$
$$\bar{\alpha}_4 = 0\ 1\ 1\ 0\ 1$$

Кодовое слово $\bar{\alpha}_1$ будет правильно декодировано тогда и только тогда, когда оно под действием шума перейдет в одно из слов, имеющихся в первом столбце таблицы декодирования. Вычислим вероятность этого события. Обозначим слова в первом столбце следующим образом: $\bar{\alpha}_{11}$, $\bar{\alpha}_{12}$, $\bar{\alpha}_{13}$, $\bar{\alpha}_{14}$, $\bar{\alpha}_{15}$, $\bar{\alpha}_{16}$, $\bar{\alpha}_{17}$, $\bar{\alpha}_{18}$, где $\bar{\alpha}_{11} = \bar{\alpha}_1$. Вероятность перехода $\bar{\alpha}_{11}$ в $\bar{\alpha}_{1k}$ будем обозначать через $p(\bar{\alpha}_{11} \to \bar{\alpha}_{1k})$. Имеем:

$$p(\bar{\alpha}_{11} \to \bar{\alpha}_{11}) = q^5; \quad p(\bar{\alpha}_{11} \to \bar{\alpha}_{15}) = pq^4;$$
$$p(\bar{\alpha}_{11} \to \bar{\alpha}_{12}) = pq^4; \quad p(\bar{\alpha}_{11} \to \bar{\alpha}_{16}) = pq^4;$$
$$p(\bar{\alpha}_{11} \to \bar{\alpha}_{13}) = pq^4; \quad p(\bar{\alpha}_{11} \to \bar{\alpha}_{17}) = p^2q^3;$$
$$p(\bar{\alpha}_{11} \to \bar{\alpha}_{14}) = pq^4. \quad p(\bar{\alpha}_{11} \to \bar{\alpha}_{18}) = p^2q^3.$$

Отсюда

$$P_A(\bar{\alpha}_1) = q^5 + 5pq^4 + 2p^2q^3.$$

Аналогично для остальных кодовых слов получаем:

$$P_A(\bar{\alpha}_2) = q^5 + 5pq^4 + 2p^2q^3;$$
$$P_A(\bar{\alpha}_3) = q^5 + 5pq^4 + 2p^2q^3;$$
$$P_A(\bar{\alpha}_4) = q^5 + 5pq^4 + 2p^2q^3.$$

Поэтому

$$P(V, A) = q^5 + 5pq^4 + 2p^2q^3.$$

Функцию $P(V, A)$ можно выразить как функцию от расстояния Хэмминга между кодовыми словами и словами в столбцах таблицы декодирования. Обозначим через $A_k(i)$ — число точек в k-м столбце таблицы декодирования, находящихся от $\bar{\alpha}_k$ на расстоянии i. Тогда

$$P(V, A) = \frac{1}{s} \sum_{k=1}^{s} \sum_{i=0}^{n} A_k(i)\, p^i\, q^{n-i}. \tag{26}$$

Действительно,

$$P_A(\bar{\alpha}_k) = \sum_{i=0}^{n} A_k(i)\, p^i\, q^{n-i}. \tag{27}$$

так как если $\rho(\bar{\alpha}, \bar{\beta}) = i$, то $p(\bar{\alpha} \to \bar{\beta}) = p^i q^{n-i}$. В нашем примере:

$$A_1(0) = 1; \quad A_2(0) = 1; \quad A_3(0) = 1; \quad A_4(0) = 1;$$
$$A_1(1) = 5; \quad A_2(1) = 5; \quad A_3(1) = 5; \quad A_4(1) = 5;$$
$$A_1(2) = 2; \quad A_2(2) = 2; \quad A_3(2) = 2; \quad A_4(2) = 2$$

и $A_k(i) = 0$ для $i \geqslant 3$.

В связи с функцией $P(V, A)$ возникают две экстремальные задачи.

Задача 1. При заданном коде V выбрать такую таблицу декодирования A, чтобы вероятность правильного декодирования $P(V, A)$ была максимальной.

Задача 2. При заданном числе s выбрать код V и схему декодирования A таким образом, чтобы функция $P(V, A)$ приняла наибольшее значение.

Задача 1 решается так называемым методом максимального правдоподобия. По этому методу таблица декодирования составляется следующим образом. Под каждым кодовым словом $\overline{\alpha}_i$ выписываются те точки множества E^n, которые находятся к $\overline{\alpha}_i$ ближе, чем к остальным кодовым словам. Если при этом некоторая точка $\overline{\beta} \in E^n$ равноудалена от кодовых слов $\overline{\alpha}_i$ и $\overline{\alpha}_j$, то ее произвольно записывают в i-й или j-й столбец.

Можно доказать, что таблица, составленная по методу максимального правдоподобия, дает решение задачи 1.

На геометрическом языке составление таблицы декодирования по методу максимального правдоподобия означает разбиение вершин E^n на области притяжения или «области Дирихле», «центрами» которых служат кодовые слова.

Нетрудно убедиться в том, что рассмотренная нами выше таблица декодирования составлена по методу максимального правдоподобия. Действительно, в каждом ее столбце находятся все точки шара радиуса единица с центром в кодовом слове и еще некоторые точки, находящиеся на расстоянии два от кодового слова. Таким образом, «областями Дирихле» при заданном разбиении являются шары радиуса единица с некоторыми точками, находящимися на расстоянии два от центров этих шаров.

Декодирование по максимуму правдоподобия еще называют декодированием в ближайшую точку, так как при получении на выходе некоторой точки $\overline{\alpha}$ мы принимаем решение, что была передана ближайшая к $\overline{\alpha}$ кодовая точка.

Следует отметить, что, несмотря на простоту описания и естественность декодирования по максимуму правдоподобия, этот способ на практике является трудно реализуемым. Дело в том, что необходимо вычислить расстояние от полученной точки до каждого кодового слова. Если же

этих слов много, то эта процедура становится трудоемкой, а в ряде случаев просто невыполнимой. Кроме того, большинство кодов задается не перечислением всех кодовых слов, а некоторым способом их порождения (проверочная матрица, порождающий многочлен и т. д.). В этом случае эта трудность еще более усугубляется. Поэтому на практике часто используют несколько «худшие», но более просто реализуемые методы декодирования.

Что касается задачи 2, то ее решение еще далеко не завершено, несмотря на большие усилия специалистов.

Много результатов теории кодирования связано с изучением функции $P(s, n)$ — максимума вероятности правильного декодирования на множестве всех кодов, содержащих ровно s-точек из E^n. Для этой функции получен ряд нижних и верхних оценок, показывающих предел наших возможностей (верхние оценки) и что можно сделать, используя имеющиеся методы (нижние оценки).

Как показывает метод максимума правдоподобия, наилучшим для декодирования будет такое расположение кодовых слов в E^n, при котором шары некоторого радиуса t с центрами в кодовых словах будут, не пересекаясь, покрывать все множество E^n. Другими словами, наилучшим расположением является совершенный код. Хотя совершенных кодов и мало, но оценка вероятности правильного декодирования для совершенного кода является верхней границей для функции $P(s, n)$.

Немного уточнив эти рассуждения, можно получить следующую верхнюю оценку:

$$P(s, n) \leqslant \sum_{i=0}^{t} C_n^i p^i q^{n-i}, \tag{28}$$

где t — наименьшее целое число, такое, что

$$2^n \left| \sum_{i=0}^{t} C_n^i \leqslant s. \right. \tag{29}$$

Действительно, поменяв в правой части (26) порядок суммирования, мы получим следующее выражение:

$$P(V, A) = \frac{1}{s} \sum_{i=0}^{n} A(i)\, p^i\, q^{n-i}, \tag{30}$$

где

$$A(i) = \sum_{k=1}^{s} A_k(i).$$

Содержательно величина $A(i)$ выражает собой число точек в E^n, находящихся на расстоянии i от некоторого кодового слова и декодирующихся в это кодовое слово. Ясно, что $A(i) \leqslant s \cdot C_n^i$. С другой стороны, функция $p^i q^{n-i}$ убывает с ростом i, так как $p < q$, поэтому

$$P(V, A) \leqslant \frac{1}{s} \sum_{i=0}^{t} s \cdot C_n^i p^i q^{n-i} = \sum_{i=0}^{t} C_n^i p^i q^{n-i},$$

где число t определяется условием (29).

Нижней оценкой для функции $P(s, n)$ безусловно является значение $P(V, A)$ для любого кода V, содержащего s точек, и для любого метода декодирования A. Однако вычисление функции $P(V, A)$ даже для кодов с относительно простой структурой является довольно сложной задачей. Гораздо проще и полезнее, как впервые заметил американский математик К. Шеннон, вычислить «среднее значение» функции $P(V, A)$ по всем кодам, содержащим ровно s точек. Это среднее значение $\overline{P}(s, n)$ и будет являться нижней оценкой функции $P(V, A)$, так как если среднее значение по всем кодам V функции $P(V, A)$ равно $\overline{P}(s, n)$, то найдется хотя бы один код V_0, такой, что $P(V_0, A) \geqslant \overline{P}(s, n)$.

Перейдем теперь к вычислению $\overline{P}(s, n)$, где

$$\overline{P}(s, n) = \frac{1}{C_{2^n}^s} \sum_{V \subseteq E^n} P(V, A). \qquad (31)$$

Суммирование в (31) ведется по всем подмножествам V, содержащим ровно s точек. Способ декодирования A — максимальное правдоподобие.

Для вычисления $\overline{P}(s, n)$ нам придется ввести несколько обозначений. Пусть $V_1, V_2, \ldots, V_{C_{2^n}^s}$ — все подмножества мощности s из E^n. Обозначим через $\xi_k(\overline{\alpha}, \overline{\beta})$ следующую функцию:

$$\xi_k(\overline{\alpha}, \overline{\beta}) = \begin{cases} 1, \text{ если точка } \overline{\beta} \text{ декодируется в } \overline{\alpha} \in V_k \\ 0 — \text{в противном случае} \end{cases}$$

Введем далее характеристическую функцию кода V_k и расстояния i:

$$\eta_k(\bar{\alpha}) = \begin{cases} 1, & \text{если } \bar{\alpha} \in V_k \\ 0 & \text{— в противном случае} \end{cases}$$

$$\delta^i(\bar{\alpha}, \bar{\beta}) = \begin{cases} 1, & \text{если } \rho(\bar{\alpha}, \bar{\beta}) = i \\ 0 & \text{— в противном случае} \end{cases}$$

С помощью введенных функций можно следующим образом выразить величину $A(i)$ для кода V_k:

$$A(i) = \sum_{\bar{\alpha},\, \bar{\beta} \in E^n} \eta_k(\bar{\alpha})\, \delta^i(\bar{\alpha}, \bar{\beta})\, \xi_k(\bar{\alpha}, \bar{\beta}). \tag{32}$$

Действительно, отдельное слагаемое в (32) равно единице, если все три сомножителя равны единице, т. е. если точка $\bar{\beta} \in E^n$, находящаяся от некоторой точки $\bar{\alpha} \in V_k$ на расстоянии i, декодируется в эту точку. И вся сумма (31) равна общему числу таких точек $\bar{\beta}$, т. е. равна $A(i)$. Из (30), (31) и (32) получаем:

$$\bar{P}(s, n) = \frac{1}{C^s_{2^n}} \sum_{k=1}^{C^s_{2^n}} \sum_{i=0}^{n} p^i q^{n-i} \sum_{\bar{\alpha},\, \bar{\beta} \in E^n} \eta_k(\bar{\alpha})\, \delta^i(\bar{\alpha}, \bar{\beta}) \times \times \xi_k(\bar{\alpha}, \bar{\beta}). \tag{33}$$

Меняя в (33) порядок суммирования, получаем:

$$\bar{P}(s, n) = \frac{1}{s \cdot C^s_{2^n}} \sum_{i=0}^{n} p^i q^{n-i} \sum_{\bar{\alpha},\, \bar{\beta} \in E^n} \delta^i(\bar{\alpha}, \bar{\beta}) \times \times \sum_{k=1}^{C^s_{2^n}} \xi_k(\bar{\alpha}, \bar{\beta})\, \eta_k(\bar{\alpha}). \tag{34}$$

Вычислим теперь самую внутреннюю сумму в (34).

Эта сумма $R(\bar{\alpha}, \bar{\beta})$ равна числу кодов V_k, таких, что $\bar{\alpha} \in V_k$ и точка $\bar{\beta} \in E^n$ декодируется в точку $\bar{\alpha}$. Заметим, что точка $\bar{\beta}$ декодируется в точку $\bar{\alpha}$ тогда и только тогда, когда в коде V_k нет точек, расположенных ближе к $\bar{\beta}$, чем точка $\bar{\alpha}$. Другими словами, $\bar{\beta}$ декодируется в $\bar{\alpha}$ тогда и только тогда, когда в шаре $S(\bar{\alpha}, \bar{\beta})$ радиуса $\rho(\bar{\alpha}, \bar{\beta})$ с центром в точке $\bar{\beta}$ нет других точек кода V_k. Поэтому если

точки $\bar{\alpha}$, $\bar{\beta}$ фиксированы, то остальные $(s-1)$-точек кода в V_h мы можем выбирать произвольным способом из множества $\Gamma = E^n - S(\bar{\alpha}, \bar{\beta})$. Так как число точек в этом множестве

$$|\Gamma| = 2^n - S_n^{\rho(\bar{\alpha}, \bar{\beta})},$$

где $S_n^{\rho(\bar{\alpha}, \bar{\beta})} = \sum_{i=0}^{\rho(\bar{\alpha}, \bar{\beta})} C_n^i$ — число точек в шаре $S(\bar{\alpha}, \bar{\beta})$, то число вариантов выбора выражается следующей формулой:

$$R(\bar{\alpha}, \bar{\beta}) = C_{2^n - S_n}^{\rho(\bar{\alpha}, \bar{\beta})}. \tag{35}$$

Из (34) и (35) теперь получаем:

$$\bar{P}(s, n) = \frac{1}{s \cdot C_{2^n}^s} \sum_{i=0}^{n} p^i q^{n-i} \sum_{\bar{\alpha}, \bar{\beta} \in E^n} \delta^i(\bar{\alpha}, \bar{\beta}) C_{2^n - S_n^{\rho}(\bar{\alpha}, \bar{\beta})}^{s-1}. \tag{36}$$

Вычислим теперь внутреннюю сумму в (36). Имеем:

$$R(i) = \sum_{\bar{\alpha}, \bar{\beta} \in E^n} \delta^i(\bar{\alpha}, \bar{\beta}) C_{2^n - S_n^{\rho}(\bar{\alpha}, \bar{\beta})}^{s-1} = K_n^i C_{2^n - n}^{s-1 - i}. \tag{37}$$

Здесь K_n^i — число пар точек $(\bar{\alpha}, \bar{\beta}) \in E^n$, таких, что $\rho(\bar{\alpha}, \bar{\beta}) = i$. Так как на расстоянии i от данной точки $\bar{\alpha} \in E^n$ находится ровно C_n^i точек, то число искомых пар равно $2^n C_n^i$, т. е.

$$K_n = 2^n C_n^i.$$

С учетом этого равенства из (36) получаем:

$$\bar{P}(s, n) = \frac{2^n}{s \cdot C_{2^n}^s} \sum_{i=0}^{n} C_n^i p^i q^{n-i} C_{2^n - S_n^i}^{s-1}.$$

Так как

$$\frac{1}{C_{2^n}^s} = \frac{s}{2^n} \cdot \frac{1}{C_{2^n - 1}^{s-1}},$$

то окончательно получаем:

$$\bar{P}(s, n) = \sum_{i=0}^{n} C_n^i \frac{C_{2^n - S_n^i}^{s-1}}{C_{2^n - 1}^{s-1}} p^i q^{n-i}. \tag{38}$$

Подробное аналитическое исследование границ (28) и (38) приводит к следующей замечательной теореме, принадлежащей К. Шеннону. Эта теорема гласит, что если

скорость передачи, которая измеряется величиной $R = \frac{\log_2 s}{n}$, не превосходит пропускную способность канала, которая в случае *д. с. к.* равна

$$c = 1 - H(p) = 1 - p \log_2 p - (1 - p) \log_2 (1 - p),$$

то существует код V, такой, что вероятность ошибки при расшифровке любой полученной на приемном конце последовательности длины n из кода V будет меньше любого заданного положительного числа ε. При этом ε зависит от n, и для уменьшения ε нужно увеличивать длину n передаваемых кодовых слов. Если же скорость передачи R больше пропускной способности, то таких кодов, вообще говоря, уже не существует.

Теорема Шеннона является одним из центральных результатов теории информации. Различные уточнения и обобщения этой теоремы для все более широких классов каналов связи составляют значительную часть этой теории.

Следует также отметить, что теорема Шеннона лишь утверждает существование некоторого кода с вероятностью правильного декодирования, близкой к единице в классе всех кодов с заданной скоростью передачи. Как «эффективно» построить такой код, в этой теореме не указывается. Такое обстоятельство связано с самим методом доказательства этой теоремы — методом случайного кодирования.

Действительно, мы вычислили среднее значение вероятности правильного декодирования на целом классе кодов и нашли, что эта вероятность как угодно мало уклоняется от единицы, если скорость передачи не превосходит пропускной способности *д. с. к.* Отсюда сразу и следует, что в этом классе существует код V, который обладает корректирующей способностью, не худшей, чем среднее по ансамблю. Из вычисления $\overline{P}(s, n)$ вытекает даже больше: подавляющая часть кодов из этого класса имеет вероятность правильного декодирования, не меньшую, чем $\overline{P}(s, n)$. Однако как конструктивно выделить хотя бы один такой код, остается неясным.

II. *Вероятность обнаружения ошибки.* Предположим теперь, что передача сообщений ведется по *д. с. к.* с так называемой бесшумной обратной связью. Содержательно это означает следующее: получив на приемном конце какое-нибудь сообщение и обнаружив в нем ошибки, мы можем попросить продублировать это сообщение и, таким обра-

зом, через некоторое время исправить ошибки. Для этого способа коррекции ошибок выгодно использовать коды, которые обладают способностью хорошо обнаруживать ошибки. Точная постановка задачи состоит в следующем.

Пусть задан код $V=\{\bar{\alpha}_1\bar{\alpha}_2 \ldots \bar{\alpha}_s\}\subseteq E^n$. Если по д. с. к. передавалось слово $\bar{\alpha}_1\in V$, а было получено слово $\bar{\beta}\notin V$, то мы говорим, что при передаче $\bar{\alpha}_1$ ошибки обнаружены. Если же $\bar{\beta}\in V$, то на приемном конце, естественно, принимаем решение, что и передавалось слово $\bar{\beta}$, т. е. ошибки не обнаруживаем. Найдем теперь вероятность $p(\bar{\alpha}_1)$ того, что при передаче $\bar{\alpha}_1$ ошибки не обнаруживаются. Это событие равносильно одному из следующих событий: $\bar{\alpha}_1\to\bar{\alpha}_k$ ($k=2, 3, \ldots, s$). Так как

$$p(\alpha_1\to\alpha_k)= p^{\rho(\alpha_1\alpha_k)}q^{n-\rho(\bar{\alpha}_1\bar{\alpha}_k)},$$

то искомая вероятность равна следующему числу:

$$p(\bar{\alpha}_1)=\sum_{k=2}^{s} p^{\rho(\bar{\alpha}_1,\ \bar{\alpha}_k)}\, q^{n-\rho(\bar{\alpha}_1,\ \bar{\alpha}_k)}\ .$$

Аналогично и для любого $\bar{\alpha}_i$ имеем:

$$p(\bar{\alpha}_i)=\sum_{k\neq i} p^{\rho(\bar{\alpha}_i,\ \bar{\alpha}_k)}\, q^{n-\rho(\bar{\alpha}_i,\ \bar{\alpha}_k)}\ .$$

Средняя вероятность необнаружения ошибки (с. в. н. о). определяется следующим образом:

$$P_H(V, q)=\frac{1}{s}\sum_{i=1}^{s} p(\alpha_i)=\frac{1}{s}\sum_{i=1}^{s}\sum_{i\neq k} p^{\rho(\bar{\alpha}_i,\ \bar{\alpha}_k)}\, q^{n-\rho(\bar{\alpha}_i,\ \bar{\alpha}_k)}. \tag{39}$$

Теперь задача состоит в том, чтобы найти код V, такой, чтобы величина $P_H(V, q)$ принимала наименьшее значение в классе всех кодов с заданным числом точек s. Это наименьшее значение обозначают через $P_H(n, s, q)$.

Если обозначить через A_i число пар точек с расстоянием i в множестве V, то (39) можно переписать в следующем виде:

$$P_H(V, q)=\frac{1}{s}\sum_{i=1}^{n} A_i p^i\, q^{n-i}. \tag{40}$$

В случае, когда код V является групповым, выражение (40) еще упрощается:

$$P_H(V, q) = \sum_{i=1}^{n} a_i p^i q^{n-i}. \tag{41}$$

Здесь a_i — число точек веса i в коде V.

Равенство (41) показывает, что знание спектра группового кода полностью определяет величину $P_H(V, q)$. Если $f(V, z)$ — производящая функция весов кода V, т. е.

$$f(V, z) = \sum_{i=1}^{n} a_i z^i,$$

то

$$P_H(V, q) = q^n \left[f\left(V \frac{p}{q} \right) - 1 \right]. \tag{42}$$

С помощью (42) легко вычислить *с. в. н. о.* для некоторых групповых кодов, спектры которых были найдены нами раньше.

1. Счетчик четности. Так мы определили код V, все точки которого имеют четный вес. Число точек в этом коде равно 2^{n-1}, и спектр его определяется следующими условиями:

$$a_{2i} = C_n^{2i};$$
$$a_{2i+1} = 0.$$

Таким образом,

$$f(V, z) = \sum_{i=0}^{\left(\frac{n}{2}\right)} C_n^{2i} z^{2i} = \frac{(1+z)^n + (1-z)^n}{2}. \tag{43}$$

Из (42) и (43) получаем:

$$P_H(V, q) = \frac{1}{2} + \frac{(q-p)^n - 2q^n}{2}.$$

2. Код Хэмминга. Для кода Хэмминга мы нашли вид производящей функции:

$$f(z) = \frac{1}{n+1} \left[(1+z)^n + n(1+z)^{\frac{n-1}{2}} (1-z)^{\frac{n+1}{2}} \right].$$

Отсюда для *с. в. н. о.* $P_H(V, q)$ кода Хэмминга получаем выражение:

$$P_H(V, q) = \frac{1}{n+1} - \frac{n(q-p)^{\frac{n+1}{2}} - q^n}{n}. \qquad (44)$$

3. Код Макдональда. Этот код является двойственным к коду Хэмминга. Все ненулевые точки этого кода имеют одинаковый вес: $d = \frac{n+1}{2}$. Поэтому *с. в. н. о.* $P_H^*(q)$ этого кода выражается следующим образом:

$$P_H^*(q) = n \cdot p^{\frac{n+1}{2}} q^{\frac{n-1}{2}}. \qquad (45)$$

Вычисление $P_H(V, q)$ для отдельных кодов представляет собой задачу более простую, чем вычисление вероятности правильного декодирования, но тем не менее еще достаточно сложную. Это же замечание относится и к функции $P_H(n, s, q)$, о которой известно больше, чем о функции $P(n, s)$, но тоже далеко не все.

Нижнюю оценку для *с. в. н. о.* любого кода $V \subseteq E^n$, состоящего из s точек, можно просто получить, используя одно хорошо известное в математическом анализе неравенство. Это неравенство Иенсена, справедливое для выпуклых функций:

$$\frac{\sum_{i=1}^{n} f(x_i)}{n} \geqslant f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right). \qquad (46)$$

Так как функция $\left(\frac{p}{q}\right)^m$ является выпуклой, то из (39) и (46) получаем: $P_H(V, q) =$

$$= \frac{q^n}{s} \sum_{i \neq j} \left(\frac{p}{q}\right)^{\rho(\bar{\alpha}_i, \bar{\alpha}_j)} \geqslant \frac{q^n}{s} s(s-1)\left(\frac{p}{q}\right)^{\sum_{i \neq j} \rho(\bar{\alpha}_i, \bar{\alpha}_j)} =$$

$$= (s-1) q^n \left(\frac{p}{q}\right)^{\sum_{i \neq j} \rho(\bar{\alpha}_i, \bar{\alpha}_j)}. \qquad (47)$$

Используя теперь ранее установленное нами неравенство (3)

$$\sum_{i \neq j} \rho(\bar{\alpha}_i, \bar{\alpha}_j) \leqslant \frac{n \cdot s^2}{2},$$

получаем:

$$P_H(V, q) \geqslant (s-1) \cdot q^n \left(\frac{p}{q}\right)^{\frac{n}{2} \cdot \frac{s}{s-1}}. \qquad (48)$$

При $s=n+1$ из (48) получаем:

$$P_H(V, q) \geqslant n \cdot p^{\frac{n+1}{2}} q^{\frac{n-1}{2}}. \qquad (49)$$

Из (45) и (49) следует, что эквидистантный код Макдональда является оптимальным в смысле *с. в. н. о.* для любой вероятности безошибочной передачи символа q. Аналогично легко доказывается, что оптимальным в смысле *с. в. н. о.* является любой групповой код, который содержится в коде Макдональда.

Используя оптимальность кода Макдональда, можно доказать оптимальность кода Хэмминга. Действительно, из формулы Мак-Вильямс (8) следует такое соотношение:

$$P(V, q) = 2^k q^n P\left(V^*, \frac{1}{2q}\right) + \frac{1}{2^{n-k}} - q^n,. \qquad (50)$$

Здесь $|V|=2^k$ и $|V^*|=2^{n-k}$.

Из формулы (50) вытекает, что если код V^* оптимален в канале с вероятностью безошибочной передачи символа $q_1 = \frac{1}{2q}$, то код V оптимален в канале с вероятностью безошибочной передачи символа, равной q.

Так как код Макдональда V^* оптимален для всех q, то и код Хэмминга обладает этим же свойством. Более того, оптимальным в смысле *с. в. н. о.* является любой групповой код, который содержит код Хэмминга. Таким образом, код Хэмминга является оптимальным во всех рассматриваемых нами смыслах.

Так же, как и для вероятности правильного декодирования, для *с. в. н. о.* можно ввести понятие среднего значения $\overline{p}(n, s, q)$ и вычислить это среднее значение по ансамблю всех подмножеств мощности s из E^n.

Если воспользоваться обозначениями, введенными в предыдущем разделе, то можно записать следующую формулу:

$$P_H = (V_k, q) = \frac{1}{s} \sum_{i=1}^{n} A_i^k p^i q^{n-i},$$

где A_i^k — число пар точек в множестве V_k с расстоянием i.

Величину A_i^k можно выразить следующим образом:

$$A_i^k = \sum_{\bar{\alpha}, \bar{\beta}} \eta_k(\bar{\alpha}) \eta_k(\bar{\beta}) \delta^i(\bar{\alpha}, \bar{\beta}).$$

Теперь для среднего значения $\overline{P}(n,\ s,\ q)$ можно написать следующую формулу:

$$\overline{P}\,(n,\ s,\ q) = \frac{1}{C^s_{2^n}} \sum_{k=1}^{C^s_{2^n}} P\,(V_k, q) =$$

$$= \frac{1}{s \cdot C^s_{2^n}} \sum_{k=1}^{C^s_{2^n}} \sum_{i=1}^{n} \sum_{\overline{\alpha},\,\overline{\beta}} \eta_k\,(\overline{\alpha})\,\eta_k\,(\overline{\beta})\,\delta^i\,(\overline{\alpha},\ \overline{\beta})\,p^i q^{n-i}\,. \qquad (51)$$

Меняя в (51) порядок суммирования, получаем:

$$\overline{P}\,(n,\ s,\ q) = \frac{1}{s \cdot C^s_{2^n}} \sum_{i=1}^{n} p^i q^{n-i} \sum_{\overline{\alpha},\,\overline{\beta}} \delta^i\,(\overline{\alpha},\ \overline{\beta}) \sum_{k=1}^{C^s_{2^n}} \eta_k(\overline{\alpha})\,\eta_k(\overline{\beta})\,. \qquad (52)$$

Самая внутренняя сумма в (52) равна числу подмножеств V_k мощности s, которое содержит заданную пару точек α, β. Поэтому эта сумма равна $C^{s-2}_{2^n-2}$. Отсюда

$$\overline{P}\,(n,\ s,\ q) = \frac{C^{s-2}_{2^n-2}}{C^s_{2^n}} \sum_{i=1}^{n} piq^{n-i} \sum_{\overline{\alpha},\,\overline{\beta}} \delta^i\,(\overline{\alpha},\ \overline{\beta}). \qquad (53)$$

Внутренняя сумма в (53) равна числу пар точек E^n с расстоянием i, т. е. K^i_n. Но мы раньше уже нашли, что

$$K^i_n = 2^n C^i_n.$$

Отсюда

$$\overline{P}\,(n,\ s,\ q) = \frac{2^n C^{s-2}_{2^n-2}}{C^s_{2^n}} \sum_{i=1}^{n} C^i_n p^i q^{n-i}\,. \qquad (54)$$

Так как

$$\frac{2^n C^{s-2}_{2^n-2}}{C^s_{2^n}} = \frac{s-1}{2^n-1}$$

и

$$\sum_{i=1}^{n} C^i_n p^i q^{n-i} = (p+q)^n - q^n = 1 - q^n,$$

то из (54) окончательно получаем:

$$\overline{P}(n, s, q) = \frac{s-1}{2^n-1}(1-q^n). \qquad (55)$$

Равенство (55) показывает, что если число точек $s(n)$ удовлетворяет предельному соотношению

$$\lim_{n\to\infty} \frac{s(n)}{2^n} = 0, \qquad (56)$$

т. е. составляет малую долю от числа всех точек E^n, то возможно обнаруживать любую комбинацию ошибок с вероятностью, как угодно близкой к единице.

Можно показать, что если условие (56) не выполнено, то ситуация меняется. Другими словами, справедливо следующее утверждение, которое можно наглядно сформулировать на языке групповых кодов: если число проверочных символов стремится к бесконечности с ростом длины блока, то с вероятностью единица можно обнаруживать любую комбинацию ошибок; если же число проверочных символов ограничено сверху, т. е. $n-k<c$, то *с. в. н. о.* никакого группового кода не может быть сделана меньше, чем $1/2^c$.

§ 5. ТЕОРИЯ КОДИРОВАНИЯ И ДРУГИЕ РАЗДЕЛЫ МАТЕМАТИКИ

Теория кодирования как сложившаяся область исследования имеет самостоятельные результаты и методы, позволяющие ее считать вполне автономной областью науки. Однако по характеру рассматриваемых в ней проблем эта теория может быть отнесена к таким более широким разделам математики, как комбинаторный анализ и дискретная геометрия.

Действительно, одна из основных задач теории кодирования — задача о плотнейшей упаковке шаров в E^n — имеет истоки в теории чисел, алгебре и кристаллографии. Эта задача в случае плоскости, т. е. E^2, может быть сформулирована следующим наглядным образом.

Требуется расположить на большом столе максимальное число одинаковых монет.

Ясно, что решение этой задачи существенно зависит от величины и формы стола. И поэтому трудно надеяться на существование общего рецепта для нахождения оптимального размещения. Однако если столом является вся плос-

кость, то задача приобретает вполне определенный смысл, состоящий в нахождении размещения с наибольшей плотностью.

Точный результат, относящийся к этой задаче, состоит в следующем. Пусть $n(R)$ — число кругов радиуса единица, которые можно разместить без пересечения внутри круга радиуса R и

$$d = \lim_{R\to\infty} \frac{n(R)}{R^2}.$$

Тогда

$$d \leqslant \frac{\pi}{\sqrt{12}} = 0{,}9069 \ldots$$

Этот результат, справедливость которого предполагал еще И. Кеплер, можно сформулировать следующим наглядным образом: непересекающимися кругами можно покрыть не более 90, 69% площади всей плоскости. Аналогично для нижней грани D плотности системы равных кругов, целиком покрывающих плоскость, справедлива следующая оценка:

$$D \leqslant \frac{\pi}{\sqrt{27}} = 1{,}209 \ldots$$

При этом плотнейшая упаковка кругов получается следующим образом: вся плоскость разбивается на шестиугольники (шестиугольный паркет), и за центр кругов берутся вершины этих шестиугольников. Таким образом, в плотнейшей упаковке кругов на плоскости каждый круг окружают шесть других.

Следует отметить, что полное доказательство этого утверждения было получено лишь в 50-х годах нашего столетия, т. е. через три столетия после того, как И. Кеплер предположил его справедливость. Решение задачи о плотнейшей упаковке шаров в 3-мерном пространстве не известно до сих пор, хотя тем же И. Кеплером также было высказано предположение о том, что здесь решение задачи дается так называемой «кубореберной» упаковкой (подробней обо всех этих вопросах можно прочесть в книге И. М. Яглома «Элементарная геометрия прежде и теперь». М., «Знание», 1972. Для тех, кто захочет основательно изучить проблемы упаковки, можно рекомендовать книгу Л. Ф. Тота «Расположения на плоскости, на сфере и в пространстве»).

Таким образом, задача о плотнейшей упаковке шаров в пространстве размерности $n \geqslant 3$, несмотря на солидный возраст, еще очень далека от окончательного решения.

I. *Теория кодирования и алгебра.* Многими своими успехами теория кодирования обязана тесной связи с такой классической областью математики, как алгебра. Основные понятия алгебры, такие, как группа, кольцо, поле, идеал, играют центральную роль в уже выделившемся специальном разделе теории кодирования — алгебраической теории кодирования. Основной задачей алгебраической теории кодирования является разработка конструктивных методов построения различных классов кодов, обладающих теми или иными корректирующими свойствами. С простейшими методами построения одного из класса кодов — циклических кодов — мы уже познакомились выше. В задачу этой теории входит также и разработка методов декодирования, имеющих небольшую сложность реализации.

Хотя задачи, рассматриваемые в алгебраической теории кодирования, не принадлежат к традиционным задачам абстрактной алгебры, а сами объекты изучения, например циклические коды, имеют очень простую структуру, в этой теории имеется очень большое число нетривиальных задач, которые заставляют по-новому взглянуть на некоторые уже, казалось бы, устоявшиеся классические теории.

В теории кодирования используется также ряд известных алгебраических операций для получения новых кодов из уже построенных.

Имеется еще ряд алгебраических конструкций. позволяющих путем формальных операций расширять классы имеющихся кодов и имеющих применение в теории кодирования.

II. *Теория кодирования и комбинаторный анализ.* Многие задачи теории кодирования могут быть, по существу, сформулированы как задачи комбинаторного анализа и теории графов в терминах, знакомых любому специалисту по этим дисциплинам.

Например, задача о построении максимального по числу точек кода с исправлением t ошибок может быть сформулирована как задача о нахождении максимального по мощности внутренне устойчивого множества в графе $(E^n)^{2t}$, где под графом $(E^n)^{2t}$ понимается граф, полученный из E^n соединением всех вершин, расстояние Хэмминга между которыми не превосходит $2t$. При этом имеет место очевидное равенство:

$$A(n,\ 2t+1)=\alpha[(E^n)^{2t}].$$

Здесь $\alpha(G)$ — число внутренней устойчивости графа G. Но на самом деле такое сведение мало что дает, так как для произвольных графов не известны ни хорошие оценки для $\alpha(G)$, ни простые алгоритмы для нахождения максимальных по мощности внутренне устойчивых множеств.

Более глубокие связи существуют между кодами, исправляющими ошибки, и тактическими конфигурациями.

Имеются также интересные связи между теорией кодирования и целочисленным линейным программированием, основной задачей которого является минимизация (максимизация) линейной целевой функции при линейных ограничениях на целочисленные переменные. Попытки сформулировать основную задачу теории кодирования на языке целочисленного линейного программирования и применить имеющийся аппарат делались неоднократно. Однако они, по-видимому, были слишком прямолинейны и не приводили к новым существенным результатам. Лишь сравнительно недавно французскому ученому П. Дилсарту удалось скомбинировать теорию двойственности линейного программирования и некоторые конструкции в теории корректирующих кодов (например, преобразование Мак-Вильямс), получить единообразным методом много верхних оценок для корректирующих возможностей кодов и ряд других интересных результатов.

Велико влияние идей и результатов теории кодирования на теорию надежности различных классов дискретных устройств. Так, коды с исправлением ошибок используются при построении асимптотически оптимальных самокорректирующихся схем из функциональных элементов.

Много результатов теории кодирования используется при исследовании надежности автоматов. В частности, эквидистантные коды были использованы Ю. Л. Сагаловичем для помехоустойчивого кодирования состояний автомата.

Многие идеи и методы теории корректирующих кодов используются в различных комбинаторных рассуждениях, конструкциях, доказательствах.

III. *Проблемы теории кодирования.* Как уже отмечалось выше, одними из основных задач теории корректирующих кодов являются задача о «плотнейшей» упаковке шаров заданного радиуса в E^n и задача о размещении в вершинах заданного числа точек таким образом, чтобы самые «близкие» из этих точек находились друг от друга на возможно более далеком расстоянии. Этим задачам посвящено значи-

тельное число работ, имеется большое число результатов, характеризующих поведение функций $A(n, d)$ и $d(s, n)$, определенных выше, но тем не менее полное решение остается пока недоступным.

Решение задачи о плотной упаковке идет двумя путями. С одной стороны, пытаются уменьшить верхние границы для $A(n, d)$ и $d(s, n)$, а с другой стороны — прямыми построениями или с помощью иных соображений доказывается существование упаковок более плотных, чем все уже известные. Следует отметить, что проблемы «плотнейшей» упаковки являются одними из самых «старых» и уважаемых проблем теории корректирующих кодов и каждый новый существенный результат в этом направлении воспринимается как серьезный успех.

Следующим, интенсивно изучающимся, классом задач теории корректирующих кодов являются задачи конструирования новых классов кодов, обладающих теми или иными полезными свойствами.

Первым из таких свойств является простота кодирования, т. е. простой способ перечисления кодовых точек. Одним из простых способов задания кода является задание с помощью порождающей или проверочной матрицы. Еще более простым способом задания характеризуются циклические коды. Для задания циклического кода, как мы видели раньше, достаточно задания одного полинома, который порождает весь код.

Вторым, практически важным, свойством кода является простота декодирования, т. е. простота способа восстановления исходного сообщения по полученному. Пример такого простого способа представляет уже рассмотренный нами алгоритм декодирования кода Хэмминга, предложенный самим автором. Относительно большая сложность процедуры декодирования кодов Боуза-Чоудхури служит препятствием для широкого практического применения этих кодов. Разработка новых эффективных способов кодирования и декодирования для уже построенных кодов и конструирование новых классов кодов с простыми алгоритмами кодирования и декодирования составляют существенный аспект теории корректирующих кодов, особенно важный с точки зрения их практического использования.

Следующим важным направлением в теории корректирующих кодов является построение специальных кодов для различных классов каналов. Дело в том, что встречающиеся на практике виды ошибок отнюдь не ограничиваются теми,

модель которых представляет рассмотренный нами двоичный симметричный канал.

Однако самое интересное заключается в том, что во многих своих существенных аспектах они очень сходны с теорией корректирующих кодов для двоичного симметричного канала. Это обстоятельство указывает на правильность выбора двоичного симметричного канала в качестве одной из основных моделей исследования и оправдывает то большое внимание, которое этому каналу уделяется.

ЛИТЕРАТУРА

Питерсон У. Коды, исправляющие ошибки. М., «Мир», 1964.

Зиновьев В. А. и Леонтьев В. К. О совершенных кодах. — «Проблемы передачи информации», 1972, 8, № 1, с. 26—35.

Tictäväinen A., Perko A. There are no Unkonown Perfect Binary Codes. Anhales Universitatis Turkuensis, 1971, 148, sor. A.

Васильев Ю. Л. О нелинейных плотно упакованных кодах. — В сб.: «Проблемы кибернетики», 8, М., «Наука», 1962, с. 337—339.

Яглом А. М. и Яглом И. М. Вероятность и информация. М., «Наука», 1973.

Виктор Константинович ЛЕОНТЬЕВ

ТЕОРИЯ КОДИРОВАНИЯ

Редактор В. И. Ковалев.
Главный отраслевой редактор И. Г. Вирко. Младший редактор Е. В. Новикова. Художник Л. П. Романенко. Худож. редактор В. Н. Конюхов. Техн. редактор Т. В. Самсонова. Корректор В. В. Канончкина.

Т-10462. Индекс заказа 74306 Сдано в набор 16.IV.1977 г. Подписано к печати 30.V.1977 г. Формат бумаги 84×108$^{1}/_{32}$. Бумага тип. № 3. Бум. л. 1,0. Печ. л. 2,0. Усл. печ. л. 3,36. Уч.-изд. л. 3,19. Тираж 43 950 экз. Издательство «Знание». 101835, Москва, Центр, проезд Серова, д. 4. Заказ 857. Цена 11 коп.
Чеховский полиграфический комбинат Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств, полиграфии и книжной торговли г. Чехов Московской области

11 коп.

Индекс 70096