ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Механико-математический факультет

Ф. И. Соловьева

ВВЕДЕНИЕ В ТЕОРИЮ КОДИРОВАНИЯ

Учебное пособие

Рекомендовано учебно-методическим Советом по математике и механике УМО по классическому университетскому образованию в качестве учебного пособия для студентов высших учебных заведений, обучающихся по специальности "010100 Математика"

Новосибирск

2006

УДК 519.725(075)
ББК з-811.4 я 73-1
С603

Соловьева Ф. И. Введение в теорию кодирования: Учебное пособие / Новосиб. гос. ун-т. Новосибирск, 2006. с. 127.

В настоящем учебном пособии изложены математические основы современной теории кодов, исправляющих ошибки в каналах связи с шумами. Пособие отражает содержание основной части лекций, читаемых автором в данное время в качестве основных курсов лекций по теории кодирования на механико-математическом факультете и факультете информационных технологий, а также читаемых в течение ряда лет в качестве специального годового курса "Введение в теорию кодирования" на механико-математическом факультете. Предназначено для студентов названных выше факультетов, а также может быть полезно студентам физического факультета, интересующимся математическими основами проблем передачи данных по каналам связи с помехами.

Рецензенты
проф., д-р техн. наук В. В. Зяблов, ИППИ РАН
д-р физ.-мат. наук С. А. Малюгин, ИМ СО РАН

Введение

Коды возникли в глубокой древности фактически с появлением системы знаков для записи звуков, слов, информации, которые позднее развились в различные языки. Каждый язык представляет собой сложную систему кодирования, включая в свою конструкцию алфавит, слова, грамматику. Язык позволяет в окружающем шуме передавать информацию по возможности быстро, надежно, с достаточно высокой степенью избыточности.

Позднее появились (еще до нашей эры) криптограммы (по-гречески криптограмма – тайнопись). Такими кодами пользовались для засекречивания сообщений. Уже в V в. до н. э. знаменитый греческий историк Геродот приводил примеры писем-криптограмм, понятных только одному адресату. Спартанцы имели специальный механический прибор, при помощи которого записывались сообщения–криптограммы, позволяющие сохранить тайну. Собственную секретную азбуку имел Юлий Цезарь (широко известный шифр Цезаря). В Средние века и эпоху Возрождения над изобретением тайных шифров работали многие выдающиеся умы, в том числе философ Фрэнсис Бэкон, математики Франсуа Виет, Джероламо Кардано. Криптографией занимались в монастырях, при дворах королей. Вместе с искусством шифрования сообщений развивалось и искусство их дешифрования. Многие оптимистично полагали, что вряд ли существует такая криптограмма, которую нельзя разгадать. И только в прошлом веке Клод Шеннон (1949 г.) показал, что существует совершенно секретный шифр – шифр Вернама, называемый также лентой однократного действия или шифром-блокнотом.

В настоящее время теория кодирования имеет важное широкое практическое применение как средство экономной, удобной, быстрой, а также надежной передачи сообщений по линиям связи с различного вида шумами (телефон, телеграф, радио, телевидение, компьютерная, космическая связи и т. д.). Подлинный взрыв развития теории связи начался в послевоенные годы, с 1948–1949 гг., с появлением классических работ Клода Шеннона и Норберта Винера. Труды Н. Винера были порождены исследованиями военного времени по автоматическому управлению огнем, труды К. Шеннона знаменитые "Математическая теория связи" и "Связь при наличии шума" – исследованиями по шифрованию сообщений и их передачи по секретным каналам связи. Математические модели Н. Винера и К. Шеннона довольно сильно различались: сигнал по Н. Винеру может обрабатываться после воздействия шумом, по К. Шеннону сигнал можно обрабатывать как до, так и после передачи по каналу связи с шумами. В силу этого и других различий, Винеровские труды легли в основу теории автоматического управления, Шенноновские труды оказались основополагающими для задач эффективного использования каналов связи. Таким образом, с 1949 г., с фундаментальных работ К. Шеннона, началось бурное развитие теории

кодирования как отдельной научной дисциплины, а также развитие таких тесно с нею связанных научных дисциплин, как сжатие информации и криптология.

В настоящем курсе лекций рассматриваются блоковые коды, предназначенные для исправления случайных ошибок в каналах связи с шумами, в основном будет изучаться модель двоичного симметричного канала связи, хотя многие результаты без труда могут быть обобщены для кодов над q-значными алфавитами. Мы только коснемся определений некоторых других типов ошибок. В курсе лекций будет изложена теория линейных кодов, в частности теория q-значных циклических кодов, имеющих широкое применение на практике для передачи сообщений в каналах связи с шумами, доказана известная теорема Шеннона о существовании хороших кодов в двоичных симметричных каналах связи с шумами, предложены различные методы построения кодов (среди них – свитчинговые и каскадные методы), а также рассмотрены известные классы кодов, таких как коды Рида – Маллера, Адамара, совершенные коды, коды Рида – Соломона, коды Юстесена, Препараты. При подготовке пособия были использованы источники [1–25].

Основные понятия и определения

Пусть E^n обозначает n-мерное метрическое пространство всех двоичных векторов длины n с метрикой Хэмминга (см. ниже примеры двумерных проекций E^n при малых n на рис. 1 и общепринятую модель E^n для произвольного n на рис. 2). Произвольное подмножество C пространства E^n называется *двоичным кодом* длины n, элементы кода называются *кодовыми словами*.

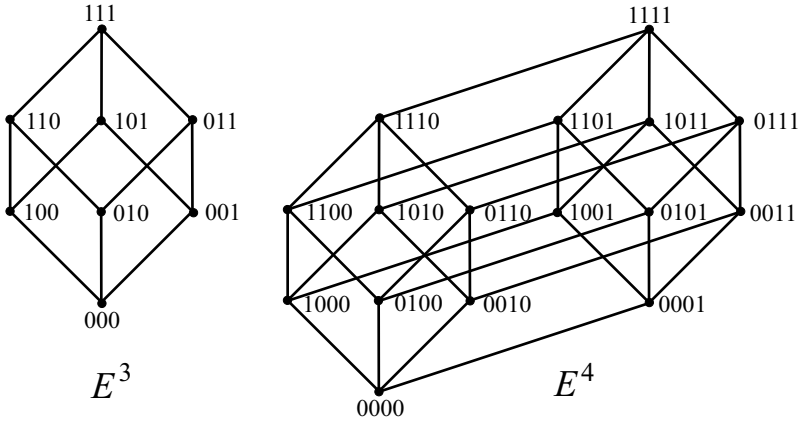

Рис. 1. Двумерные проекции E^3 и E^4

Хэммингово расстояние $d(x, y)$ между векторами $x, y \in E^n$ определяется как число координат, в которых эти векторы различаются. Нетрудно показать, что оно является метрикой. *Кодовое расстояние* равно минимальному расстоянию Хэмминга между различными кодовыми словами.

Вес Хэмминга $w(x)$ произвольного вектора $x \in E^n$ равен числу ненулевых координат x, т. е. $w(x) = d(x, \mathbf{0}^n)$, где $\mathbf{0}^n$ – нулевой вектор длины n. Обозначим через $\mathbf{1}^n$ единичный вектор длины n (иногда далее длины нулевого и единичного векторов указываться не будут, но из контекста всякий раз будет ясно, какова их длина). Множество

$$supp(x) = \{i \mid x_i = 1\}$$

называется *носителем* вектора x.

Известно, что группа автоморфизмов $\mathrm{Aut}(E^n)$ пространства E^n исчерпывается подстановками π на множестве координат и добавлением произвольного вектора $v \in E^n$, т. е. для группы автоморфизмов $\mathrm{Aut}(E^n)$ пространства E^n справедливо

$$\mathrm{Aut}(E^n) = E^n \leftthreetimes S_n = \{(v, \pi) \mid v + \pi(E^n) = (E^n), v \in E^n, \pi \in S_n\},$$

где $\leftthreetimes$ – полупрямое произведение, S_n – симметрическая группа подстановок длины n.

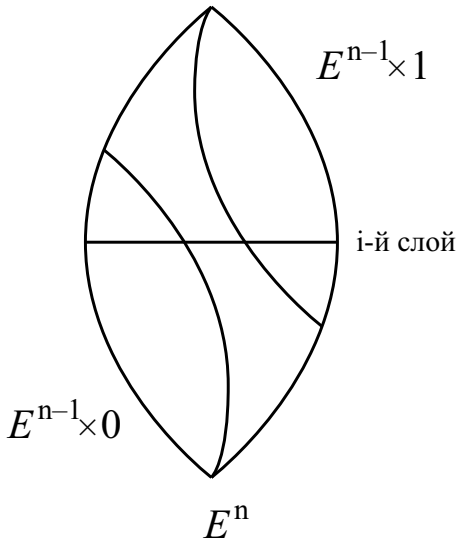

Рис. 2. Двумерная проекция E^n

Два кода – C и D – длины n называются *эквивалентными*, если существует автоморфизм (v, π) пространства E^n, отображающий один код в другой, т. е. $v + \pi(C) = D$.

Группа автоморфизмов $\mathrm{Aut}(C)$ произвольного кода C длины n состоит из всех автоморфизмов пространства E^n, переводящих код в себя, т. е.

$$\mathrm{Aut}(C) = \{(v, \pi) \mid v + \pi(C) = C\}.$$

Множество

$$\mathrm{Sym}(C) = \{\pi \in S_n \mid \pi(C) = C\}$$

называется *группой симметрий* кода C. Очевидно, что $\mathrm{Sym}(C)$ изоморфна подгруппе группы $\mathrm{Aut}(C)$.

Линейным (или групповым) кодом называется подмножество E^n, являющееся линейным подпространством (подгруппой) в E^n. Аналогично, *линейным q–значным кодом* называется линейное подпространство n–мерного метрического пространства E_q^n всех векторов длины n с метрикой Хэмминга над полем Галуа $GF(q)$, $q = p^k, q \geq 2$, где p–простое число (такой код может не являться групповым в E_q^n).

В дальнейшем параметры линейного кода C длины n с кодовым расстоянием d будем обозначать через $[n, k, d]$, где k – размерность кода; для нелинейного кода C параметры будем обозначать через $(n, |C|, d)$.

Линейный код длины n называется *циклическим*, если для любого кодового слова $(x_1, x_2, \ldots, x_n)$ слово $(x_2, \ldots, x_n, x_1)$ также является кодовым.

Для полноты изложения в дальнейшем некоторые из приведенных определений будут повторены.

Упражнение 1. Доказать, что расстояние Хэмминга является метрикой, а E^n — метрическим пространством:

а) $d(x, y) \geq 0$, причем $d(x, y) = 0 \Leftrightarrow x = y$ (аксиома тождества);

б) $d(x, y) = d(y, x)$ (аксиома симметрии);

в) $d(x, y) + d(y, z) \geq d(x, z)$ (аксиома треугольника) для $\forall x, y, z \in E^n$.

Упражнение 2. Найти число вершин и ребер в E^n, E_q^n.

Упражнение 3. Найти число вершин:
а) в сфере радиуса r в E^n, E_q^n;
б) в шаре радиуса r в E^n, E_q^n.

Двоичный симметричный канал связи

Пусть по каналу связи с шумом пересылаются двоичные сообщения из Новосибирска в Москву. Рассмотрим случай, когда входной алфавит A совпадает с выходным алфавитом B и равен $\{0, 1\}$. Пусть при посылке 0 принимается как 0, а 1 как 1, но иногда 0 может быть принят как 1 или 1 принята как 0. Пусть в среднем один из каждых 1 000 символов будет ошибочным. Это означает, что для каждого символа имеется вероятность $p = 1/1\,000$ того, что в канале связи произойдет ошибка, т. е. для переходных (условных) вероятностей $P(\beta|\alpha)$, где $\sum_{\alpha\in A} P(\beta|\alpha) = 1$, имеем

$$P(0|0) = P(1|1) = p \text{ и } P(1|0) = P(0|1) = 1 - p.$$

В данном случае переходные вероятности образуют симметричную матрицу и посему такая модель называется двоичным *симметричным каналом связи*. Сообщения должны передаваться как можно быстрее, экономнее и надежнее. Будем записывать сообщения в виде последовательностей из нулей и единиц. Закодируем эти сообщения в целях защиты их от ошибок, которые могут произойти в канале связи. Входная информация разбивается на блоки длины k. Каждый блок из k символов сообщения

$$u = (u_1, u_2, \ldots, u_k),\ u_i \in \{0, 1\}$$

преобразуется кодером в слово x длины n :

$$x = (x_1, x_2, \ldots, x_n),\ x_j \in \{0, 1\}, n > k.$$

Полученные слова образуют *двоичный код* и называются *кодовыми словами*. Схематично это представлено на рис. 3.

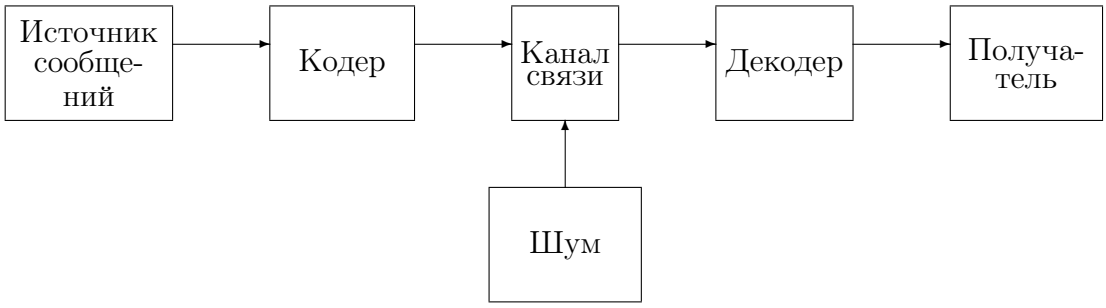

Рис. 3. Классическая схема системы связи по каналу связи с шумом

Так как канал с шумом, то принятый вектор y может отличаться от кодового слова x на вектор e, называемый вектором ошибок $y = x + e$, $e = (e_1, e_2, \ldots, e_n)$, где с вероятностью $(1 - p)$ имеем $e_i = 0$ (в этом случае i-й символ правильный) и с

вероятностью p имеем $e_i = 1$ (i-й символ исказился). В общем случае вероятность p может быть произвольным числом, удовлетворяющим неравенству $0 < p < 1/2$.

Опишем некоторые другие, отличные от симметричных типы ошибок. Если для переходных вероятностей имеются запреты, то канал связи получается с односторонними ошибками, если в алфавит (входной и (или) выходной) включен пустой символ, то получаем канал связи со вставками или выпадениями. Рассмотрим подробнее несколько случаев.

1. Несимметричная ошибка типа $\{1 \to 0\}$ (ее называют также замещением вида $1 \to 0$) может возникнуть в ситуации, когда происходит замена 1 на 0, но не наоборот. Если наличию физического сигнала соответствует 1, а отсутствию – 0, то такие ошибки происходят в результате размыканий (обрывов) в канале, так как при размыкании сигнал может лишь исчезнуть.

2. Несимметричная ошибка типа $\{0 \to 1\}$ (замещение вида $0 \to 1$) аналогична предыдущей и происходит в результате замыканий в канале.

3. Ошибка стирания $\{0 \to z, 1 \to z\}$ возникает в случае, если сигнал, представляющий собой 0 или 1, искажается в канале так, что его нельзя интерпретировать ни как 0, ни как 1. Этот символ заменяется неопределенным символом, который обозначается через z.

4. Ошибка вида $0 \to \wedge$ ($1 \to \wedge$) состоит в удалении одного из символов кодового слова x, в результате чего длина слова уменьшается на единицу. Такие ошибки называют выпадениями символов (в выходной алфавит включен пустой символ).

5. Ошибка вида $\wedge \to 0$ ($\wedge \to 1$) состоит во вставке символа 0 (1) перед некоторым символом или после последнего символа слова x, в итоге длина кодового слова увеличится на единицу. Одиночные ошибки такого типа называют вставками символов.

Существуют также другие виды ошибок.

Глава 1

Линейные коды

1.1. Линейные коды

Напомним, что *линейным (или групповым) двоичным кодом* называется подмножество E^n, являющееся линейным подпространством (подгруппой) в E^n. Произвольный линейный код с параметрами $[n, k, d]$ можно задать различными способами:

аналитически, с помощью одной формулы (такой способ задания линейного кода не всегда может найтись);

посредством *кодовой матрицы* порядка $2^k \times n$ (строками матрицы являются кодовые слова);

порождающей матрицы порядка $k \times n$ (в строки записаны кодовые слова, образующие базу линейного кода);

проверочной матрицы – матрицы H такой, что для любого кодового слова $x = (x_1, x_2, \ldots, x_n)$ выполняется

$$H \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = Hx^T = \mathbf{0}^{n-k},$$

здесь H – матрица порядка $(n-k) \times n$. Последнее уравнение задает $(n-k)$ проверочных уравнений. Очевидно, что такое представление линейного кода также является аналитическим.

Следует отметить, что для данного линейного кода представление порождающей (проверочной) матрицей не единственно.

Опишем подробнее задание линейного кода посредством проверочной матрицы, имеющей канонический вид. Пусть от отправителя в кодер поступило сообщение $u = (u_1, u_2, \ldots, u_k)$. Сформируем кодовое слово $x = (x_1, x_2, \ldots, x_n)$. Положим первую часть кодового слова состоящей из символов самого сообщения (называемых *информационными символами*): $x_1 = u_1,\ x_2 = u_2, \ldots, x_k = u_k$. Далее следуют $n - k$ символов, называемых *проверочными* $x_{k+1}, \ldots, x_n$. Они выбираются таким образом, чтобы все кодовые слова удовлетворяли уравнению

$$Hx^T = \mathbf{0}^{n-k}.$$

Пусть матрица H имеет вид $[A_{n-k,k}|E_{n-k}]$, называемый *каноническим*, где $A_{n-k,k}$ – некоторая матрица порядка $(n-k)\times k$ из 0 и 1, E_{n-k} – единичная матрица порядка $n-k$. Все операции выполняются над полем Галуа $GF(2)$ характеристики 2.

Теорема 1. О связи проверочной и порождающей матриц. *Если проверочная матрица линейного кода задана в каноническом виде $H=[A_{n-k,k}|E_{n-k}]$, то порождающая матрица этого кода имеет вид $G=[E_k|-A^T_{n-k,k}]$. Верно обратное.*

Доказательство. Рассмотрим произвольное кодовое слово

$$x=(x_1,\ \ldots\ ,x_k,x_{k+1},\ \ldots,\ x_n),$$

где $x_{k+1},\ \ldots,\ x_n$ – проверочные символы, а $x_1,\ldots,x_k$ – информационные, т. е. информационный блок имеет вид

$$u=(u_1,\ \ldots,\ u_k),\ \text{где } x_1=u_1,x_2=u_2,\ \ldots,\ x_k=u_k,$$

что можно записать в матричном виде

$$\begin{pmatrix} x_1\\ \vdots\\ x_k\end{pmatrix}=E_k\begin{pmatrix} u_1\\ \vdots\\ u_k\end{pmatrix}. \tag{1.1}$$

Пусть

$$A=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1k}\\ a_{21} & a_{22} & \ldots & a_{2k}\\ \ldots & \ldots & \ldots & \ldots\\ a_{n-k} & a_{n-k,2} & \ldots & a_{n-k,k}\end{pmatrix}.$$

Тогда из определения проверочной матрицы имеем $Hx^T=\mathbf{0}$, т. е.

$$a_{i1}x_1+a_{i2}x_2+\ldots+a_{ik}x_k+x_{k+i}=0$$

для любого $i=1,\ldots,n-k$. Отсюда

$$x_{k+i}=-(a_{i1}x_1+a_{i2}x_2+\ldots+a_{ik}x_k),\ i=1,\ldots,n-k.$$

Таким образом,

$$\begin{pmatrix} x_{k+1}\\ \vdots\\ x_n\end{pmatrix}=-A_{n-k,k}\begin{pmatrix} x_1\\ \vdots\\ x_k\end{pmatrix}=-A_{n-k,k}\begin{pmatrix} u_1\\ \vdots\\ u_k\end{pmatrix}.$$

Из последнего и из уравнения (1.1.) имеем

$$x^T=\begin{pmatrix} E_k\\ -A_{n-k,k}\end{pmatrix}u^T.$$

Транспонируя, получаем: $x=u\,G$, где $G=[E_k|-A^T_{n-k,k}]$. ▲

Заметим, что теорема верна для q-значных кодов.

Упражнение 4. Найти число различных базисов в E^n.

Упражнение 5. Найти число различных линейных двоичных кодов длины n размерности k.

Упражнение 6. Показать, что в двоичном линейном коде либо каждый кодовый вектор имеет четный вес, либо половина кодовых векторов имеет четные веса и половина — нечетные.

Упражнение 7. Доказать, что ненулевой столбец кодовой матрицы линейного $[n, k]$-кода содержит ровно 2^{k-1} единиц.

Упражнение 8. Доказать, что для любого линейного кода справедливо $HG^T = \mathbf{0}$ для любых проверочной и порождающей матриц H и G этого кода соответственно.

1.2. Границы объемов кодов

Рассмотрим несколько известных границ мощностей кодов.

Теорема 2. Граница Хэмминга. *Для любого двоичного кода C длины n (не обязательно линейного) с кодовым расстоянием d выполняется неравенство*

$$|C| \leq \frac{2^n}{\sum_{i=0}^{[(d-1)/2]} C_n^i}.$$

Доказательство. Обозначим $t = [(d-1)/2]$. Поскольку кодовое расстояние равно d, то шары

$$S_t^n(x) = \{y \mid y \in E^n, d(y, x) \leq t]\}$$

радиуса t, описанные около кодовых слов x, не пересекаются. Очевидно, что все они имеют одинаковый объем, равный

$$|S_t^n(\mathbf{0}^n)| = C_n^0 + C_n^1 + \ldots + C_n^{[\frac{d-1}{2}]}.$$

Следовательно,

$$|C| \times |S_t^n(\mathbf{0}^n)| \leq 2^n. \tag{1.2}$$

Подставляя $|S_t^n(\mathbf{0}^n)|$ в неравенство (1.2.), получаем требуемое. ▲

Определение. Код называется *совершенным* или *плотно упакованным*, если

$$|C| = \frac{2^n}{\sum_{i=0}^{[(d-1)/2]} C_n^i},$$

т. е. имеет место плотная упаковка E^n шарами радиуса $[(d-1)/2]$.

Справедливы следующие утверждения.

Утверждение 1. *Кодовое расстояние $[n, k, d]$-линейного кода равно минимальному из весов его ненулевых кодовых слов.*

Теорема 3. О столбцах проверочной матрицы. *Если H – проверочная матрица кода длины n, то код имеет кодовое расстояние d тогда и только тогда, когда любые $d-1$ столбцов матрицы H линейно независимы и найдутся d линейно зависимых столбцов.*

Доказательство. Необходимость.

Вектор x веса ω принадлежит коду тогда и только тогда, когда

$$Hx^T = \mathbf{0}, \tag{1.3}$$

что эквивалентно линейной зависимости некоторых ω столбцов матрицы H. Обозначим i-й столбец матрицы H через h_i, т. е.

$$H = [h_1, h_2, \ldots, h_n].$$

Отсюда и из равенства (1.3) получаем

$$\sum_{i=1}^{n} h_i x_i = \mathbf{0},$$

откуда следует соотношение линейной зависимости $h_{i_1} + \ldots + h_{i_w} = \mathbf{0}$. По утверждению 1 кодовое расстояние кода равно минимальному из весов его ненулевых кодовых слов. По условию теоремы код имеет кодовое расстояние d, откуда получаем линейную зависимость некоторой совокупности d столбцов матрицы H.

Если существует $d-1$ линейно зависимых столбцов в матрице H, то найдется вектор веса $d-1$, принадлежащий коду C, противоречие.

Достаточность очевидна. ▲

Непосредственным следствием теоремы 3 является следующая верхняя граница объема кода.

Теорема 4. Граница Синглтона. *Для любого линейного $[n, k, d]$-кода выполняется $n - k \geq d - 1$.*

Код, достигающий границу Синглтона, называется *MDS-кодом*. Код, полученный из данного кода удалением одной или более координат во всех кодовых словах, называется *выколотым кодом*.

Теорема 5. Граница Синглтона для нелинейных q-значных кодов. *Для любого $(n, M, d)_q$-кода выполняется* $\log_q M \leq n - d + 1$.

Доказательство. Удаляя в $(n, M, d)_q$-коде последовательно любые $d-1$ координат, получим код длины $n - d + 1$ с кодовым расстоянием по крайней мере 1 и мощности M. ▲

Теорема 6. Граница Плоткина. *При $n < 2d$ для любого двоичного (n, M, d)-кода C справедливо неравенство*

$$M \leq 2[d/(2d - n)],$$

где M – мощность кода C.

Доказательство. Вычислим двумя способами сумму

$$S = \sum_{u \in C} \sum_{v \in C, v \neq u} d(u, v)$$

для различных кодовых слов u и v из C. Поскольку при $u \neq v$ расстояние $d(u, v) \geq d$, то сумма не меньше, чем $M(M-1)d$. С другой стороны, пусть A обозначает кодовую $(M \times n)$-матрицу, строками которой являются все кодовые слова. Предположим, что i-й столбец матрицы A содержит x_i нулей и $M - x_i$ единиц. Тогда вклад этого столбца в сумму S равен $2x_i(M - x_i)$. Суммируя по всем столбцам, получаем

$$S = \sum_{i=1}^{n} 2x_i(M - x_i).$$

При четном M максимум этого выражения достигается при $x_i = M/2$ для любого i, следовательно, эта сумма не превышает $nM^2/2$, т. е. имеем

$$M(M-1)\, d \leq nM^2/2,$$

отсюда

$$M \leq 2d/(2d - n).$$

Так как M четно, то

$$M \leq 2[d/(2d - n)].$$

При нечетном M эта сумма не превышает $n(M^2 - 1)/2$ и, следовательно,

$$M \leq n/(2d - n) = 2d/(2d - n) - 1.$$

Отсюда с учетом $[2x] \leq 2[x] + 1$ получаем

$$M \leq [2d/(2d - n)] - 1 \leq 2[d/(2d - n)].$$

▲

Теорема 7. Граница Варшамова–Гилберта. *Если выполняется неравенство*

$$1 + C_{n-1}^{1} + \cdots + C_{n-1}^{d-2} < 2^r,$$

то существует двоичный линейный код длины n с минимальным расстоянием по крайней мере d, имеющий не более чем r проверочных символов, т. е. $[n, k, d']$–код, где $k \geq n - r$, $d' \geq d$.

Доказательство. Теорема будет доказана, если построим $(r \times n)$-матрицу H такую, что любые ее $d - 1$ столбцов линейно независимы. Тогда, применяя теорему 3, получаем требуемое утверждение. В качестве первого столбца матрицы H возьмем любой ненулевой вектор длины r. Предположим, что выбрали i столбцов матрицы H так, что любые $d - 1$ из них линейно независимы. Имеем не более

$$C_i^1 + \ldots + C_i^{d-2}$$

различных ненулевых линейных комбинаций из этих i столбцов, содержащих $d-2$ или меньше столбцов. Если это число меньше, чем 2^r-1 (числа всех ненулевых векторов длины r), то мы можем добавить еще один столбец, не равный ни одной из всех этих линейных комбинаций. При этом любые $d-1$ столбцов новой матрицы размера $r \times (i+1)$ по-прежнему остаются линейно независимы. Будем выполнять эту процедуру до тех пор, пока выполняется неравенство

$$C_i^1 + \ldots + C_i^{d-2} < 2^r - 1.$$

▲

Упражнение 9. Обобщить для q-значных кодов границы:
а) Хэмминга;
б) Варшамова–Гилберта.

Упражнение 10. Доказать утверждение 1 и теорему 4.

1.3. Код Хэмминга и его свойства

1.3.1. Определение кода Хэмминга

Для построения линейного кода Хэмминга с m проверками на четность, исправляющего одну ошибку, воспользуемся теоремой 3: определим код посредством проверочной матрицы, столбцами которой являются **все** ненулевые векторы длины m. Очевидно, что любые два столбца этой матрицы линейно независимы и найдутся три линейно зависимых столбца, следовательно по теореме 3 кодовое расстояние равно 3 и значит код исправляет одну ошибку. Этот код называется *кодом Хэмминга*, далее будем его обозначать $\mathcal{H}^n$.

Параметры кода Хэмминга:

$$[n = 2^m - 1, k = n - \log(n+1), d = 3],$$

$m = \log(n+1)$ (здесь и всюду далее $\log(\cdot)$ является двоичным логарифмом, если не оговорено особо).

Утверждение 2. *Код Хэмминга $\mathcal{H}^n$ является совершенным кодом, исправляющим одну ошибку.*

Доказательство. Код $\mathcal{H}^n$ исправляет одну ошибку (по определению кода). По построению его мощность равна

$$|\mathcal{H}^n| = 2^{n-m} = \frac{2^n}{n+1},$$

где $m = \log(n+1)$. Следовательно, он достигает границы Хэмминга (см. теорему 2) и потому является совершенным. ▲

Утверждение 3. *Код Хэмминга единствен с точностью до изоморфизма.*

1.3.2. Примеры кодов Хэмминга длины 7

Рассмотрим три различных представления кода Хэмминга длины 7.

1. Код Хэмминга длины 7 задан проверочной матрицей в каноническом виде (см. [1], гл. 1), т. е. проверочная матрица имеет вид

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

2. Код Хэмминга длины 7 в циклическом виде. **Определение.** Линейный код C длины n называется *циклическим*, если для любого кодового слова $x = (x_1, x_2, \ldots, x_n)$ слово $(x_2, x_3, \ldots, x_n, x_1)$ принадлежит коду C.

Проверочная матрица кода Хэмминга длины 7 в циклическом представлении имеет вид:

$$H = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

3. Во многих случаях полезно определять код Хэмминга через проверочную матрицу, заданную в лексикографическом виде

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix},$$

т. е. столбцы проверочной матрицы записаны в лексикографическом порядке возрастания двоичных представлений натуральных чисел.

1.3.3. Декодирование кода Хэмминга

Код Хэмминга допускает самое простое декодирование. Рассмотрим представление кода Хэмминга $\mathcal{H}^n$ посредством проверочной матрицы, столбцы которой записаны в лексикографическом порядке:

$$H_m = [B(1), B(2), \ldots, B(n)],$$

здесь H_m – проверочная матрица кода $\mathcal{H}^n$, $m = \log(n+1)$, $B(i)$ – двоичное представление числа i. Используя эту проверочную матрицу H_m, можно определить код Хэмминга следующим образом:

$$\mathcal{H}^n = \{\, x = (x_1, \ldots, x_n) : x \in E^n, \quad \sum_{i=1}^{n} B(i) \cdot x_i = \mathbf{0}^m \}.$$

Определение. Вектор $S_y = Hy^T$, где H — проверочная матрица линейного кода, называется *синдромом* вектора y.

Пусть в канале связи при передаче вектора x произошла одна ошибка в i-й координате и получен вектор y. Воспользовавшись тем, что для любого кодового слова x кода $\mathcal{H}^n$ выполняется $Hx^T = \mathbf{0}^m$, найдем синдром вектора y:

$$S_y = Hy^T = Hx^T + He_i^T = He_i^T = B(i).$$

Синдром S_y равен столбцу, номер которого i является номером ошибочной координаты вектора y. Здесь e_i – двоичный вектор длины n с единицей только в i-й координате.

Позднее (см. гл. 4, подразд. 4.5.1.) будет рассмотрен пример кода Хэмминга над полем Галуа $GF(q)$ для произвольного $q = p^m$, где p – любое простое число.

Упражнение 11. Доказать утверждение 3.

1.4. Способы построения новых кодов

Здесь рассмотрим несколько общих способов построения кодов, как линейных, так и нелинейных. В случае построения линейных кодов их параметры будут заключены в квадратные скобки, в случае нелинейных кодов – в круглые скобки.

1. Метод комбинирования кодов

Теорема 8. *Пусть G_1, G_2 – порождающие матрицы для $[n_1, k, d_1]$ и $[n_2, k, d_2]$ – кодов соответственно. Тогда коды с порождающими матрицами*

$$\begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix} \text{ и } \begin{pmatrix} G_1 \mid G_2 \end{pmatrix}$$

представляют собой $[n_1 + n_2, 2k, min\{d_1, d_2\}]$- и $[n_1 + n_2, k, d]$-коды соответственно, причем $d \geqslant d_1 + d_2$.

2. Добавление общей проверки на четность

Из двоичного кода C с параметрами $(n, |C|, d)$, где d нечетно, образуем новый код C' с параметрами $(n + 1, |C|, d + 1)$, называемый *расширенным кодом* добавляя 0 в конце каждого кодового слова C четного веса и добавляя 1 в конце каждого кодового слова нечетного веса. Все кодовые слова кода C' имеют, очевидно, четный вес. Код C' удовлетворяет общему уравнению проверки на четность

$$x_1 + x_2 + \ldots + x_{n+1} = 0.$$

Для линейных кодов проверочная матрица кода C' имеет вид

$$\begin{pmatrix} 1 & \ldots & 1 \\ & & 0 \\ & H & \vdots \\ & & 0 \end{pmatrix},$$

где H – проверочная матрица исходного кода C. Такой код называется *расширенным кодом.*

3. Выкалывание кодовых координат

Выкалывание кодовых координат представляет собой удаление одной и более координат во всех кодовых словах. Если исходный код C имел параметры: $(n, |C|, d)$, то код C', полученный выкалыванием одной координаты из C, имеет следующие параметры: $(n-1, |C'|, d')$, где $|C'| \leq |C|$, $d-1 \leq d' \leq d$ (заметим, что $|C| = |C'|$, если $d > 1$).

4. Код с выбрасыванием

Код с выбрасыванием получается из исходного удалением всех слов нечетного (или четного) веса. Из кода с параметрами (n, k, d) получается код (n, k', d'), где $k' \leq k$, $d' \geq d$ и часто $d' > d$ (например, если d — нечетное и удалены все слова нечетного веса).

Вообще говоря, можно рассматривать коды с выбрасыванием слов некоторого веса, например, малого или большого.

5. Пополнение кода

Пополнение кода C с параметрами $[n, k, d]$ добавлением новых слов представляет собой следующее: если в E^n найдется вектор a такой, что $d(C, a) \geq d$, то добавим к исходному коду множество $C + a$. При этом мы получим код той же длины n, размерности $k+1$.

Упражнение 12. Оценить кодовое расстояние полученного кода.

6. Укорочение кода

Укорочение кода состоит в следующем:

а) выбираем все кодовые слова, у которых координата i равна 0 (либо 1). Как правило, выбирается более мощная часть кодовой матрицы с фиксированной координатой i, если таковой нет, как, например, в линейных кодах, то выбираются все кодовые слова, у которых координата i равна 0;

б) удаляем эту координату в выбранных словах.

Из кода C с параметрами $(n, |C|, d)$ получается $(n-1, |C'|, d')$-код C', где $|C'| \geq |C|/2$, $d' \geq d$.

Упражнение 13. Построить расширенный код Хэмминга длины 8 добавлением общей проверки на четность. Доказать, что код имеет расстояние 4, обнаруживает 2 ошибки и исправляет одну ошибку.

Конструкция Плоткина

Рассмотрим еще один эффективный способ построения кодов, который позволяет, имея в качестве стартовых коды малых длин с оптимальными или близкими к оптимальным параметрами, строить бесконечные серии кодов с такими же хорошими параметрами. В дальнейшем нам потребуется следующее нетрудно доказываемое утверждение.

Утверждение 4. *Для любых векторов x и y из E^n справедливо*

$$w(x+y) \geq w(x) - w(y).$$

Теорема 9 (Плоткин М., 1960, см. [1]). *Пусть C и D — двоичные (n, M_1, d_1) и (n, M_2, d_2)-коды соответственно. Тогда множество*

$$C^{2n} = \{(x, x+y) : x \in C, y \in D\}$$

является $(2n, M_1 \cdot M_2, d = min\{2d_1, d_2\})$-кодом.

Доказательство. Пусть $u = (x, x+y)$, $v = (x', x'+y')$ — произвольные различные кодовые слова кода C^{2n}, где $x, x' \in C$, $y, y' \in D$.

Если $y = y'$, то

$$d(u,v) = d((x,x),(x',x')) = 2d(x,x') \geq 2d_1.$$

Пусть $y \neq y'$, тогда, используя утверждение 4, получаем

$$d(u,v) = w(x-x') + w(x+y-x'-y') =$$

$$= w(x-x') + w((y-y') + (x-x')) \geq$$

$$\geq w(x-x') + w(y-y') - w(x-x') = w(y-y') = d_2.$$

▲

Упражнение 14. Доказать теорему 8.

Упражнение 15. Найти порождающую матрицу расширенного кода Хэмминга длины 16, построенного с помощью конструкции Плоткина. Какие коды для этого нужно использовать?

1.5. Теорема Глаголева

В этом разделе множество строк порождающей матрицы кода будем называть *базовым множеством*. Для доказательства теоремы Глаголева потребуется следующий несложно доказываемый факт (см. также п. 5 предыдущего раздела).

Лемма 1. *Если G — линейный код с кодовым расстоянием d и если найдется такой вектор x, что $d(G, x) \geq d$, то множество $G \cup (G + x)$ является линейным кодом с кодовым расстоянием d.*

Теорема 10 (Глаголев В. В., 1971). *Для любого двоичного линейного $[n, k, d]$-кода C существует линейный код C' с теми же параметрами такой, что его базовое множество состоит из кодовых слов минимального веса d.*

Доказательство. В качестве базового множества рассмотрим множество

$$T_d \cup T_{d+1} \cup \ldots \cup T_{d+p}$$

кода C. Здесь T_d — максимальное линейно независимое множество кодовых слов веса d; T_{d+1} — множество кодовых слов веса $d+1$, которое может быть выбрано в коде C так, что $T_d \cup T_{d+1}$ — максимальное линейно независимое множество кодовых слов веса не более $d+1$. Аналогично выбираем остальные множества вплоть до T_{d+p} кодовых слов веса $d+p$ для некоторого p. Таким образом код C совпадает с линейной оболочкой множества $T_d \cup T_{d+1} \cup \ldots \cup T_{d+p}$, т. е.

$$C =< T_d \cup T_{d+1} \cup \ldots \cup T_{d+p} > .$$

Рассмотрим произвольный вектор y из T_{d+1}. Докажем, что расстояние между y и любым кодовым словом из T_d больше d. Пусть это неверно и найдется вектор $z \in T_d$ такой, что $d(y,z) = d$. Тогда $w(y+z) = d$ и в силу линейности кода C имеем $y+z \in C$. С другой стороны, в силу линейной независимости множества $T_d \cup T_{d+1}$ справедливо $y+z \notin T_d$. Следовательно, получили подмножество $T_d \cup (y+z)$ в коде C, которое является линейно независимым множеством кодовых слов веса d и имеет мощность больше мощности множества T_d, что противоречит выбору множества T_d. Следовательно,

$$d(T_d, y) \geq d+1. \tag{1.4}$$

Возьмем любой вектор y' веса d, предшествующий вектору y, т. е. $y' \prec y$. Это означает, что все единичные координаты вектора y' находятся среди единичных координат кодового слова y. Используя неравенство (1.4), получаем

$$d(T_d, y) > d(T_d, y') \geq d.$$

Рассмотрим множество $T_d \cup (T_d + y')$. Согласно лемме 1, его линейная оболочка является линейным кодом с кодовым расстоянием d. Далее аналогичным образом в тени некоторого кодового слова множества T_{d+1} найдем вектор y'' и рассмотрим множество $T_d \cup \{y', y''\}$, которое позволяет построить новый линейный код с расстоянием d и т. д., переходя от множества T_{d+1} к множеству T_{d+2} и далее до множества T_{d+p}, не более чем за k шагов построим линейный $[n,k,d]$-код C' с базовым множеством, состоящим из кодовых слов минимального веса d. ▲

Замечание. В 1992 г. Ю. Симонис получил аналогичный результат для q-значных линейных кодов над $GF(q)$.

Следующее утверждение вытекает из теоремы Глаголева и утверждения 3 о единственности кода Хэмминга.

Следствие 1. *Для произвольного кода Хэмминга существует базовое множество, состоящее из кодовых слов веса* 3.

Упражнение 16. Доказать лемму 1.

Глава 2

Декодирование

2.1. Декодирование двоичных кодов

Пусть сообщение $u = (u_1, \ldots, u_k)$ закодировано кодовым словом $x = (x_1, \ldots, x_n)$, которое передается по каналу связи с шумом. Принятый вектор $y = (y_1, \ldots, y_n)$ может отличаться от x. Введем вектор ошибок

$$e = y - x = (e_1, \ldots, e_n),$$

где $e_i = 0$ с вероятностью $1 - p$, $e_i = 1$ с вероятностью p (исказился i-й символ), p — произвольное число, удовлетворяющее $0 < p < 1/2$.

Задача декодера — решить на основании принятого слова y, какое сообщение u или, что, как правило, удобнее, какое кодовое слово x было передано. Если декодер найдет слово e, то легко вычислить кодовое слово $x = y - e$. Но декодер часто не может определить в точности, чему равен вектор ошибок e. В этом случае его стратегия заключается в выборе наиболее вероятного вектора ошибок e.

Опишем декодирование в ближайшее кодовое слово или декодирование по максимуму правдоподобия (для двоичного симметричного канала связи эти методы декодирования совпадают, поскольку метрика Хэмминга согласована с двоичным симметричным каналом). Поскольку ошибки происходят независимо с вероятностью p, то для вектора ошибок e длины n имеем

$$\begin{aligned} &P\{e = (000\ldots0)\} = (1-p)^n, \\ &P\{e = (010\ldots0)\} = p \cdot (1-p)^{n-1}, \\ &\ldots \\ &P\{e = v, w(v) = k\} = p^k \cdot (1-p)^{n-k}, \end{aligned}$$

где e — вектор ошибки длины n.

Так как $p < 1/2$, то $1 - p > p$ и, очевидно, справедливо

$$(1-p)^n > p \cdot (1-p)^{n-1} > \ldots > p^k \cdot (1-p)^{n-k} > \ldots \ .$$

Другими словами, фиксированный вектор ошибок веса 1 более вероятен, чем вектор ошибок веса 2, и т. д. Отсюда напрашивается стратегия декодирования: y декодируется в ближайшее кодовое слово x или, что то же самое, выбирается вектор ошибок e наименьшего веса. Такая процедура декодирования называется декодированием *по*

максимуму правдоподобия (т. е. в выборе наиболее вероятного вектора ошибок e для данного принятого вектора y) или *декодированием в ближайшее кодовое слово*.

Для кодов, мощность которых невелика, допустимо декодирование полным перебором, но, очевидно, что эта стратегия не является эффективной. Одной из главных целей теории корректирующих кодов является конструирование кодов с эффективными процедурами декодирования.

2.2. Декодирование линейных кодов

2.2.1. Стандартное расположение. Синдром

Для линейных кодов могут применяться особые процедуры декодирования с использованием синдромов и таблицы стандартного расположения.

Пусть C — линейный двоичный $[n, k]$-код (все дальнейшие рассуждения в этом пункте справедливы для линейных кодов над полем из $q \geq 2$ элементов). Для любого вектора a множество

$$a + C = \{a + x \mid x \in C\}$$

называется *смежным классом* (или сдвигом) кода C. Каждый смежный класс содержит 2^k векторов, два вектора a, b принадлежат одному и тому же смежному классу тогда и только тогда, когда $a - b \in C$. Любые два смежных класса либо не пересекаются либо совпадают (частичное пересечение невозможно). Этот факт легко доказывается от противного.

Таким образом, n-мерный единичный куб E^n можно разбить на классы смежности по линейному коду C:

$$E^n = C \cup (a^1 + C) \cup \ldots \cup (a^m + C),$$

где $m = 2^{n-k} - 1$.

Вектор y, принятый декодером, попадает в некоторый i-й класс смежности

$$y \in a^i + C,$$

т. е. $y = a^i + x$ для некоторого $x \in C$. Если было передано слово x', то вектор ошибок вычисляется как

$$e = y - x' = a^i + x - x' = a^i + x'' \in a^i + C,$$

т. е. вектор ошибок принадлежит тому же i-му классу смежности. Таким образом, если получили вектор y из i-го класса смежности, то возможными векторами ошибок будут все векторы i-го смежного класса. Отсюда вытекает следующая *стратегия декодера*: из смежного класса, содержащего вектор y, выбирается вектор e наименьшего веса. Если таких векторов несколько, то выбирается любой из них. Далее производится декодирование y как

$$x = y - e.$$

Вектор минимального веса из смежного класса называется *лидером смежного класса*.

Опишем таблицу, называемую *стандартным расположением для кода*. Стандартное расположение — удобный способ описания работы декодера. В первую строку

таблицы помещаются сообщения, во вторую — кодовые векторы, причем в первом столбце стоит нулевой вектор: $x^1 = \mathbf{0}, x^2, \ldots, x^{2^k}$. В третью строку под нулевым вектором помещается лидер a_1 некоторого класса смежности по коду C и строка заполняется таким образом, чтобы под кодовым словом x^i стояло слово $a^1 + x^i$. Следующие строки заполняются аналогично. Процесс продолжается до тех пор, пока не исчерпаются все векторы из E^n.

Сообщение	$\mathbf{u^1}$	$\mathbf{u^2}$	$\cdots$	$\mathbf{u^{2^k}}$	**Синдром**
Код	$x^1 = \mathbf{0}$	x^2	$\cdots$	x^{2^k}	S_0
Первый смежный класс	a^1	$a^1 + x^2$	$\cdots$	$a^1 + x^{2^k}$	S_{a^1}
Второй смежный класс	a^2	$a^2 + x^2$	$\cdots$	$a^2 + x^{2^k}$	S_{a^2}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$m = (2^{n-k} - 1)$-й смежный класс	a^m	$a^m + x^2$	$\cdots$	$a^m + x^{2^k}$	S_{a^m}

По построению в таблице стандартного расположения в строках находятся классы смежности с лидерами классов смежности, расположенными в первом столбце. В последнем столбце записываются синдромы лидеров. Сначала рассмотрим процесс декодирования без использования столбца синдромов.

Пример 1. Рассмотрим пример линейного $[4, 2]$-кода (см. [1], с. 26) с порождающей матрицей

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Таблица стандартного расположения без столбца синдромов для этого кода имеет вид

Сообщение	**00**	**10**	**01**	**11**
Код	0000	1011	0101	1110
Первый смежный класс	1000	0011	1101	0110
Второй смежный класс	0100	1111	0001	1010
Третий смежный класс	0010	1001	0111	1100

Каким образом действует декодер?

• Ищет в таблице полученное на выходе канала связи слово y, например, $y = (1101)$.

• Принимает решение, что вектор ошибок e — это лидер класса смежности, содержащего вектор y, т. е. $e = (1000)$.

• Далее вектор y декодируется в вектор $x = y - e = (1101) + (1000) = (0101)$ и делается вывод, что исходное сообщение было равно (01).

Очевидно, что декодирование, использующее стандартное расположение, является декодированием по максимуму правдоподобия.

Если линейный код имеет небольшие параметры, то таблица стандартного расположения очень удобна для процесса декодирования. Но в большинстве случаев такая процедура неэффективна, так как требует большого объема вычислений. Например,

если взять линейный код длины 100 и размерности 70, то таблица декодирования содержит 2^{70} столбцов и 2^{30} строк, т. е. имеет очень большие размеры.

Процесс декодирования может быть упрощен с помощью синдромов. Для декодирования достаточно выписать лидеры смежных классов и соответствующие им синдромы. После того как получен вектор y, вычисляется его синдром. Далее отыскивается лидер смежного класса a^i с тем же синдромом и вычитание этого лидера из вектора y

$$x = y - a^i$$

дает предположительно посланный вектор x.

2.2.2. Свойства синдрома

1. Если проверочная матрица имеет $(n-k)$ строк, то синдром S_y произвольного вектора $y \in E^n$ является вектором длины $(n-k)$.

2. Поскольку по определению линейного кода вектор y является кодовым тогда и только тогда, когда $Hy^T = \mathbf{0}$, то справедливо

Утверждение 5. *Синдром S_y вектора y равен $\mathbf{0}$ тогда и только тогда, когда y является кодовым вектором.*

3. Справедливо

Утверждение 6. *Для двоичного линейного кода синдром S_y принятого вектора y равен сумме тех столбцов проверочной матрицы H, где произошли ошибки.*

Доказательство. Пусть получен вектор $y = x + e$, где x — кодовый вектор. Тогда, по определению синдрома и из утверждения 5, имеем

$$S_y = Hy^T = H(x+e)^T = Hx^T + He^T = He^T.$$

Пусть e имеет "1" в координатах с номерами i_1, i_2, …, i_s, т. е.

$$supp(e) = \{i_1, i_2, \ldots, i_s\}$$

Это означает, что произошли ошибки в i_1, i_2, …, i_s координатах. Таким образом, имеем

$$He^T = \sum_{k=1}^{s} e_{i_k} h_{i_k} = h_{i_1} + h_{i_2} + \ldots + h_{i_s},$$

где h_{i_k} — это i_k-й столбец матрицы H. Следовательно,

$$S_y = \sum_{k=1}^{s} h_{i_k},$$

т. е. действительно синдром выделяет те позиции вектора, где произошли ошибки. ▲

4. Имеет место взаимно однозначное соответствие между синдромами и смежными классами.

Утверждение 7. *Два вектора u и v принадлежат одному и тому же смежному классу тогда и только тогда, когда их синдромы равны.*

Доказательство. Два элемента группы u и v принадлежат одному и тому же смежному классу по данной подгруппе тогда и только тогда, когда $u-v$ принадлежит этой подгруппе. В нашем случае подгруппой является код C, т. е. $u - v \in C$. Тогда по определению линейного кода выполняется

$$H(u - v)^T = \mathbf{0}.$$

Отсюда $Hv^T = Hu^T$. ▲

Таким образом, синдром содержит всю информацию, которую имеет декодер об ошибках: по принятому слову y вычисляется синдром S_y. По утверждению 6 синдром равен сумме тех столбцов проверочной матрицы, где произошли ошибки. По синдрому ищется лидер смежного класса, то есть вектор ошибок и далее — кодовое слово.

Рассмотрим вычисление синдрома для приведенного выше примера. Сначала по порождающей матрице G найдем проверочную матрицу H:

$$G = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}\right) \longrightarrow H = \left(\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right).$$

Затем вычислим синдром для смежного класса с лидером (1000), он равен $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, т. е. первому столбцу матрицы H. Далее вычислим все синдромы и запишем их в таблицу синдромов.

Сообщение	**Лидер**	**Синдром**
Код	0000	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
Первый смежный класс	1000	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Второй смежный класс	0100	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
Третий смежный класс	0010	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Пусть получено слово $y = (1100)$. Вычислим его синдром:

$$Hy^T = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

он равен третьему столбцу проверочной матрицы H. Ему отвечает лидер смежного класса (0010), он же является искомым вектором ошибок, т. е.

$$x = y - (0010) = (1100) \oplus (0010) = (1110).$$

Отсюда делаем вывод, что было передано кодовое слово (1110).

Упражнение 17. Построить таблицу стандартного расположения для кода с проверочной матрицей

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

и декодировать два произвольных вектора, один непосредственно с помощью таблицы стандартного расположения, другой вектор — с помощью таблицы синдромов.

2.3. Вероятность ошибки декодирования

Определение. *Вероятностью ошибки декодирования* или *вероятностью ошибки на слово* $P_{\text{ош}}$ для данной схемы декодирования называется вероятность появления некодового слова на выходе декодера.

Пусть дан линейный код длины n мощности M с кодовыми словами

$$\{x^1, \ldots, x^M\},$$

где кодовые слова используются с равной вероятностью. Обозначим вероятность того, что на выходе декодера получено слово $y \neq x^i$ при переданном x^i через

$$P(y \neq x^i | x^i).$$

Тогда

$$P_{\text{ош}} = \frac{1}{M} \sum_{i=1}^{M} P(y \neq x^i | x^i),$$

т. е. вероятность ошибки на слово равна средней вероятности неправильного декодирования.

При декодировании, использующем стандартное расположение, правильное декодирование имеем тогда и только тогда, когда вектор ошибок совпадает с лидером смежного класса. Иначе говоря, вероятность ошибки в этом случае равна

$$P_{\text{ош}} = P\{e \neq \text{лидер смежного класса}\}.$$

Если имеем α_i лидеров смежных классов веса i, $i = 0, \ldots, n$, то вероятность правильного декодирования $P_{\text{пр.дек.}}$ равна

$$P_{\text{пр.дек.}} = \sum_{i=0}^{n} \alpha_i p^i (1-p)^{n-i}, \tag{2.1}$$

поскольку вероятность правильного декодирования кодового слова с некоторым вектором ошибок v веса i равна

$$p^i (1-p)^{n-i}$$

(см. разд. 2.1.).

Так как стандартное расположение обеспечивает декодирование по максимуму правдоподобия, то любая другая схема будет иметь вероятность правильного декодирования меньше, чем (2.1), а следовательно, и вероятность ошибки произвольная

схема декодирования будет иметь не меньше, чем вероятность ошибки при декодировании, использующем стандартное расположение, поскольку

$$P_{\text{ош}} = 1 - P_{\text{пр.дек.}} = 1 - \sum_{i=0}^{n} \alpha_i p^i (1-p)^{n-i}. \tag{2.2}$$

Для кода, приведенного выше, имеем $\alpha_0 = 1$, $\alpha_1 = 3$. Таким образом, при $p = \frac{1}{100}$ получим

$$P_{\text{ош}} = 1 - (1-p)^4 - 3p(1-p)^3 = 0,0103\ldots .$$

Рассмотрим линейный код с минимальным расстоянием $d = 2t+1$. По теореме 3 он исправляет t ошибок и, следовательно, каждый вектор веса не больше t является лидером смежного класса, т. е.

$$\alpha_i = C_n^i,\ 0 \le i \le t.$$

Если вероятность p ошибки в канале очень мала, то $(1-p) \approx 1$ и, следовательно,

$$p^{i+1}(1-p)^{n-i-1} = o(p^i(1-p)^{n-i}).$$

Отбрасывая в равенстве (2.2) все члены, начиная с $i > t$, получаем формулу, аппроксимирующую формулу (2.2) достаточно точно:

$$P_{\text{ош}} \sim 1 - \sum_{i=0}^{t} C_n^i p^i (1-p)^{n-i}. \tag{2.3}$$

Аналогично в случае кодового расстояния $d = 2t+2$ получаем следующую аппроксимацию формулы (2.2):

$$P_{\text{ош}} \sim 1 - \sum_{i=0}^{t} C_n^i p^i (1-p)^{n-i} - \alpha_{t+1} p^{t+1} (1-p)^{n-t-1}. \tag{2.4}$$

Если $\alpha_i = 0$ при $i > t = \lfloor (d-1)/2 \rfloor$, то формула (2.3) становится точной; при $i > t+1$ становится точной формула (2.4). В первом случае код называется *совершенным*, во втором — *квазисовершенным*.

Геометрически это означает, что в первом случае имеем разбиение пространства E^n на непересекающиеся шары радиуса t (так как код может исправлять не больше, чем t ошибок). Во втором случае, поскольку код исправляет все ошибки веса не больше t, некоторые ошибки веса $t+1$ и не может исправлять ни одной ошибки веса больше, чем $t+1$, имеем покрытие пространства E^n шарами радиуса $t+1$. При этом шары радиуса $t+1$ могут пересекаться.

Более тонкой мерой качества декодирования является *вероятность ошибки на символ*.

Пусть имеем линейный код мощности $M = 2^k$ с кодовыми словами $x^1, \ldots, x^M$, где первые k символов $x_1^i, \ldots, x_k^i$ в каждом слове являются информационными. Пусть $y = (y_1, \ldots, y_n)$ — слово на выходе декодера.

Определение. *Вероятность ошибки на символ* $P_{\text{симв.}}$ равна средней вероятности того, что после декодирования информационный символ является ошибочным:

$$P_{\text{симв.}} = \frac{1}{kM} \sum_{j=1}^{k} \sum_{i=1}^{M} P\{y_j \ne x_j^i | x^i \text{ было послано}\}.$$

Вернемся к декодированию стандартным расположением. Так как все сообщения и ошибки в любом символе независимы и равновероятны (т. е. рассматривается симметричный канал), то достаточно рассмотреть произвольный кодовый вектор $x = (x_1, \ldots, x_n)$. Пусть получен вектор y, тогда

$$P_{\text{симв.}} = \frac{1}{k} \sum_{j=1}^{k} P\{y_j \neq x_j^i\}.$$

Пусть $f(e)$ — число неправильных информационных символов после декодирования при условии, что e — вектор ошибок, тогда

$$P_{\text{симв.}} = \frac{1}{k} \sum_{e} f(e) P\{e\}. \tag{2.5}$$

Рассмотрим наш пример: изучая таблицу стандартного расположения, приходим в выводу, что $f(e) = 0$ для всех векторов ошибок первого столбца, так как в первом столбце стоят лидеры смежных классов; $f(e) = 1$ для всех векторов ошибок как второго, так и третьего столбцов; $f(e) = 2$ для всех векторов четвертого столбца. Отсюда по формуле (2.5) при $p = 1/100$ получаем

$$P_{\text{симв.}} = \frac{1}{2}\Big[1 \cdot \big(p^4 + 3p^3(1-p) + 3p^2(1-p)^2 + p(1-p)^3\big) + 2 \cdot \big(p^3(1-p) + 3p^2(1-p)^2\big)\Big] =$$

$$= 0{,}00530\ldots .$$

Глава 3

Теорема Шеннона

3.1. Необходимые понятия

Рассмотрим двоичный симметричный канал связи. Пусть для каждого символа имеется вероятность $0 < p < \frac{1}{2}$ того, что при передаче его по каналу связи произойдет ошибка. Пусть C — двоичный код, содержащий M равновероятных кодовых слов $x^1, \ldots, x^M$ длины n, в котором каждое слово встречается с равной вероятностью. Напомним, что *вероятностью ошибки на слово* или вероятностью ошибки P_C для данной схемы декодирования называется вероятность появления неправильного кодового слова на выходе декодера. Пусть P_i — вероятность неправильного декодирования при условии, что передано кодовое слово x^i. Тогда

$$P_C := \frac{1}{M} \sum_{i=1}^{M} P_i,$$

где P_i зависит от вероятности p. Рассмотрим совокупность $\widetilde{\mathfrak{L}}$ всех двоичных кодов длины n мощности M и определим

$$P^*(M, n, p) := \min_{C \in \widetilde{\mathfrak{L}}} \{P_C\}. \tag{3.1}$$

Определение. *Функция энтропии* $\mathcal{H}(x)$ определяется равенством

$$\mathcal{H}(x) = -x \log x - (1 - x) \log(1 - x)$$

при $0 < x < 1$, при $x = 0$ и $x = 1$ полагают $\mathcal{H}(0) = \mathcal{H}(1) = 0$.

Отметим, что $\log x$ здесь и далее рассматривается по основанию 2.

Определение. *Пропускная способность* двоичного симметричного канала с вероятностью $0 \leq p \leq \frac{1}{2}$ равна

$$C(p) = 1 - \mathcal{H}(p) = 1 + p \log p + (1 - p) \log(1 - p).$$

Определение. *Скоростью* (n, M, d)-кода называется величина $(\log M)/n$.

Теорема 11 (Шеннон К., 1948). *Пусть R — любое число, удовлетворяющее условию $0 < R < C(p)$, и пусть $M_n = 2^{[n \cdot R]}$. Тогда*

$$P^*(M_n, n, p) \to 0 \text{ при } n \to \infty.$$

Теорему Шеннона можно переформулировать следующим образом: для любой сколь угодно малой величины $\varepsilon > 0$ и любого $0 < R < C(p)$ существует двоичный код C длины n, мощности M и скорости R такой, что вероятность ошибки декодирования $P_C < \varepsilon$, где M определяется из соотношения $R = (\log M)/n$.

Или, говоря неформально, для достаточно больших n существует хороший код длины n, со скоростью, сколь угодно близкой к пропускной способности канала связи.

Прежде чем доказать теорему, рассмотрим несколько важных свойств энтропии по Шеннону и приведем необходимые понятия и утверждения, которые будут использоваться в доказательстве теоремы Шеннона.

3.2. Свойства энтропии

Рассмотрим бернуллиевский источник A: буквы входного алфавита $a_i, i = 1, \ldots, k$ появляются независимо с независимыми вероятностями p_i, удовлетворяющими условию

$$\sum_{i=1}^{k} p_i = 1.$$

Энтропия $\mathcal{H}(A)$ источника A по Шеннону определяется следующим образом:

$$\mathcal{H}(A) = -\sum_{i=1}^{k} p_i \log p_i.$$

В определенной выше энтропии $\mathcal{H}(x)$ имеем два исхода с вероятностями p и $1 - p$ соответственно.

Рассмотрим несколько важных свойств энтропии.

Утверждение 8. *Справедливо $\mathcal{H}(A) \geq 0$. Энтропия неотрицательна и равна нулю тогда и только тогда, когда одна из вероятностей равна единице, а остальные равны нулю.*

Доказательство. Действительно, из $0 \leq p_i \leq 1$ имеем $\frac{1}{p_i} \geq 1$ и $\log \frac{1}{p_i} \geq 0$, т. е. $-\log p_i \geq 0$, отсюда $-p_i \log p_i \geq 0$. Поскольку, по определению, $-p_i \log p_i = 0$ при $p_i = 0$, то для любого $p_i \geq 0$ выполняется $-p_i \log p_i \geq 0$ и, следовательно,

$$\mathcal{H}(A) = -\sum_{i=1}^{k} p_i \log p_i \geq 0.$$

При $\mathcal{H}(A) = 0$ каждое слагаемое равно нулю, а значит, либо $p_i = 0$, либо $\log p_i = 0$, т. е. $p_i = 1$. Так как $\sum_{i=1}^{k} p_i = 1$, то среди вероятностей p_i принять значение 1 может лишь одна, остальные равны нулю. Таким образом, неопределенность события равна нулю тогда и только тогда, когда исход события заранее известен, в остальных случаях энтропия положительна. ▲

Рассмотрим произвольную непрерывную выпуклую вверх функцию $f(x)$, заданную на положительной полуоси. Для произвольных положительных чисел β_i,

$i = 1, \ldots, k$ таких, что $\sum_{i=1}^{k} \beta_i = 1$ и любых $x_1, \ldots, x_k$ из участка выпуклости функции $f(x)$ выполняется *неравенство Йенсена*

$$\sum_{i=1}^{k} \beta_i f(x_i) \leq f\left(\sum_{i=1}^{k} \beta_i x_i\right),$$

причем равенство имеет место тогда и только тогда, когда

$$x_1 = \ldots = x_k.$$

Утверждение 9. *Для источника A, определенного выше, справедливо* $\mathcal{H}(A) \leq \log k$.

Доказательство. Рассмотрим функцию $f(x) = \log x$. Она выпукла вверх при $x > 0$, поскольку ее вторая производная меньше нуля. Положим $\beta_i = p_i$. Тогда с учетом $\sum_{i=1}^{k} p_i = 1$ и неравенства Йенсена имеем

$$\mathcal{H}(A) = -\sum_{i=1}^{k} p_i \log p_i = \sum_{i=1}^{k} p_i \log\left(\frac{1}{p_i}\right) \leq \log\left(\sum_{i=1}^{k} p_i \frac{1}{p_i}\right) = \log k.$$

▲

Утверждение 10. *Для произвольных $q_1, \ldots, q_k$ таких, что $q_i \geq 0$ и $\sum_{i=1}^{k} q_i = 1$ справедливо*

$$-\sum_{i=1}^{k} p_i \log q_i \geq -\sum_{i=1}^{k} p_i \log p_i,$$

причем равенство достигается только при $q_i = p_i$ (таким образом значение энтропии минимизирует функцию $f(q_1, \ldots, q_k) = -\sum_{i=1}^{k} p_i \log q_i$).

Доказательство. Рассмотрим доказательство этого свойства в случае, когда все p_i положительны (при $p_i = 0$ для некоторого $i = 1, \ldots, k$, доказательство аналогично с некоторыми модификациями). Воспользуемся неравенством Йенсена при $\beta_i = p_i$, $x_i = q_i/p_i$, $i = 1, \ldots, k$ для функции $f(x) = \log x$:

$$\sum_{i=1}^{k} p_i \log \frac{q_i}{p_i} \leq \log\left(\sum_{i=1}^{k} p_i \cdot \frac{q_i}{p_i}\right) = \log\left(\sum_{i=1}^{k} q_i\right) = \log 1 = 0.$$

Отсюда имеем

$$\sum_{i=1}^{k} p_i \log q_i - \sum_{i=1}^{k} p_i \log p_i \leq 0,$$

откуда вытекает требуемое. Неравенство Йенсена переходит в равенство только тогда, когда $x_1 = \ldots = x_k$ или в нашем случае при $q_1/p_1 = \ldots = q_k/p_k$, т. е. когда векторы $(q_1, \ldots, q_k)$ и $(p_1, \ldots, p_k)$ пропорциональны и, следовательно, в силу

$$\sum_{i=1}^{k} q_i = \sum_{i=1}^{k} p_i = 1$$

имеем $q_i = p_i$. ▲

Утверждение 11. *Для любых случайных опытов A и B справедливо*

$$\mathcal{H}(AB) \le \mathcal{H}(A) + \mathcal{H}(B),$$

причем равенство достигается только когда опыты A и B независимы. Для бернуллиевских источников справедливо

$$\mathcal{H}(A^N) = N \cdot \mathcal{H}(A)$$

для любого натурального N.

Здесь AB обозначает произведение источников A и B : пусть

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_k \\ p_1 & p_2 & \dots & p_k \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 & b_2 & \dots & b_k \\ q_1 & q_2 & \dots & q_k \end{pmatrix},$$

тогда

$$AB = \begin{pmatrix} (a_i,\, b_j) \\ p_i q_j \end{pmatrix},$$

где $i, j \in \{1, 2, \dots, k\}$; A^N — произведение N копий источника A.

В дальнейшем нас будет интересовать случай событий с двумя исходами, поскольку основная модель изучаемого нами канала связи — двоичный симметричный канал. В этой ситуации имеем функцию $\mathcal{H}(x)$ одного аргумента. В силу утверждения 8, она равна нулю при $x = 0$ и $x = 1$, согласно утверждения 9, достигает максимума, равного единице, в точке $x = 1/2$. Кроме того, для нее выполняется соотношение

$$\mathcal{H}(x) = \mathcal{H}(1 - x),$$

следовательно, функция $\mathcal{H}(x)$ симметрична относительно прямой $x = 1/2$, см. ее график на рис. 4.

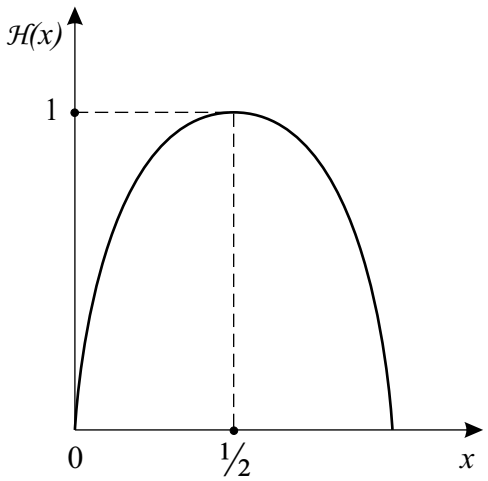

Рис. 4. График функции $\mathcal{H}(x)$

3.3. Необходимые комбинаторно-вероятностные утверждения

I. Неравенство Чебышева. При передаче информации по двоичному симметричному каналу связи число ошибок в полученном слове является бернуллиевской случайной величиной τ, принимающей значения $0, 1, \ldots, n$, математическое ожидание $\mathcal{E}(\tau)$ которой равно np, а дисперсия $\mathcal{D}(\tau)$ равна $np(1-p)$. Если в кодовом слове $x = (x_1, \ldots, x_n)$ произошло t ошибок, то, в силу того что имеем *биномиальное распределение*, вероятность получить вектор ошибок e веса t равна

$$P(e = v, w(v) = t) = p^t(1-p)^{n-t},$$

т. е. зависит только от n и t.

Теорема 12. Неравенство Чебышева. *Если ξ — случайная величина с математическим ожиданием $\mathcal{E}(\xi)$ и дисперсией $\mathcal{D}(\xi)$, тогда для любого $\varepsilon > 0$ справедливо*

$$P\{|\xi - \mathcal{E}(\xi)| \geq \varepsilon\} \leq \frac{\mathcal{D}(\xi)}{\varepsilon^2}.$$

Выберем произвольное $\varepsilon > 0$. Для случайной величины τ обозначим через b следующую величину:

$$b = \left(\frac{\mathcal{D}(\tau)}{\varepsilon/2}\right)^{1/2}.$$

Тогда, используя неравенство Чебышева, получаем

$$P\{|\tau - \mathcal{E}(\tau)| \geq b\} \leq \frac{\mathcal{D}(\tau)}{b^2} = \frac{\varepsilon}{2}.$$

Откуда следует

$$P\{\tau > \rho\} \leq \frac{\varepsilon}{2}, \tag{3.2}$$

где

$$\rho = [\mathcal{E}(\tau) + b] = \left[np + \left(\frac{np(1-p)}{\varepsilon/2}\right)^{1/2}\right].$$

Неравенство (3.2) означает: вероятность того, что в результате произошедших в канале τ ошибок полученное на приемном конце слово y находится от переданного кодового слова x на расстояние большее, чем ρ, мала (не превышает $\varepsilon/2$).

Зафиксируем $\varepsilon > 0$, тогда для достаточно больших n величина ρ не превосходит $(n/2)$, поскольку $p < 1/2$.

II. Объем шара радиуса $[pn]$**.** Рассмотрим шар радиуса $[pn]$ с центром в некоторой вершине $x \in E^n$:

$$B_{[pn]}(x) = \{y \,|\, d(x, y) \leq [pn]\}.$$

Оценим его объем

$$|B_{[pn]}(x)| = \sum_{i=0}^{[pn]} C_n^i$$

с помощью функции энтропии $\mathcal{H}(p)$.

Лемма 2 *Пусть* $0 \le p \le \frac{1}{2}$. *Тогда справедлива оценка*

$$\sum_{i=0}^{[pn]} C_n^i \le 2^{n\mathcal{H}(p)}.$$

Доказательство. Имеем

$$1 = (p + (1-p))^n \ge \sum_{i=0}^{[pn]} C_n^i p^i (1-p)^{n-i} \ge \sum_{i=0}^{[pn]} C_n^i p^{pn} (1-p)^{n-np} =$$

$$= \sum_{i=0}^{[pn]} C_n^i 2^{\log(1-p)^n (\frac{p}{1-p})^{pn}} = \sum_{i=0}^{[pn]} C_n^i 2^{n\log(1-p) + pn\log(\frac{p}{1-p})} =$$

$$= \sum_{i=0}^{[pn]} C_n^i 2^{n(p\log p + (1-p)\log(1-p))} = 2^{-n\mathcal{H}(p)} \sum_{i=0}^{[pn]} C_n^i.$$

Отсюда $\sum\limits_{i=0}^{[pn]} C_n^i \le 2^{n\mathcal{H}(p)}$.

▲

III. Объем шара радиуса $\rho = [\mathcal{E}(\tau) + b]$**.** Оценим объем шара радиуса $\rho = [\mathcal{E}(\tau) + b]$ с центром в некоторой вершине, используя функцию энтропии $\mathcal{H}(p)$.

Лемма 3 *Пусть* $0 \le p \le \frac{1}{2}$ *и* $\rho = [\mathcal{E}(\tau) + b]$, *где* $b = \left(\frac{\mathcal{D}(\tau)}{\varepsilon/2}\right)^{1/2}$. *Тогда*

$$\frac{1}{n}\log|B_\rho(x)| \le \mathcal{H}(p) - O\Big(\frac{1}{\sqrt{n}}\Big) \text{ при } n \to \infty.$$

Доказательство. По лемме 2 имеем

$$\frac{1}{n}\log|B_\rho(x)| \le \mathcal{H}\Big(\frac{\rho}{n}\Big) = -\frac{\rho}{n}\log\frac{\rho}{n} - \Big(1 - \frac{\rho}{n}\Big)\log\Big(1 - \frac{\rho}{n}\Big) =$$

$$= -\frac{[np+b]}{n}\log\frac{[np+b]}{n} - \Big(1 - \frac{[np+b]}{n}\Big)\log\Big(1 - \frac{[np+b]}{n}\Big) =$$

$$= -p\log p - (1-p)\log(1-p) - O\Big(\frac{b}{n}\Big) = \mathcal{H}(p) - O\Big(\frac{1}{\sqrt{n}}\Big) \text{ при } n \to \infty,$$

что доказывает лемму. ▲

3.4. Доказательство теоремы Шеннона

Введем функцию $f(y, x)$. Пусть $x, y \in E^n$, тогда

$$f(y,x) = \begin{cases} 0, & \text{если } d(y,x) > \rho, \\ 1, & \text{если } d(y,x) \le \rho. \end{cases}$$

Функция $f(y,x)$ — характеристическая функция принадлежности вектора y шару с центром в вершине x, т. е.

$$f(y,x)=\begin{cases}0, & \text{если } y\notin B_\rho(x),\\ 1, & \text{если } y\in B_\rho(x).\end{cases}$$

Доказательство теоремы Шеннона. Выберем сколь угодно малую величину $\varepsilon>0$. Рассмотрим случайный двоичный код длины n мощности M, т. е. выберем случайным образом кодовые слова $x^1,\dots,x^M$. Декодируем полученный вектор y следующим образом: если существует в точности одно кодовое слово x^i такое, что

$$d(x^i,y)\le\rho,$$

то y декодируем в x^i, в противном случае регистрируем ошибку или, если должны произвести декодирование в любом случае, всегда декодируем в x^1.

Пусть P_i, как и выше, вероятность того, что на выходе декодера получено слово, отличное от переданного слова x^i. Для P_i имеем следующую оценку сверху:

$$P_i\le\sum_{y\in E^n}P(y|x^i)\bigl[1-f(y,x^i)+\sum_{j\ne i}f(y,x^j)\bigr]=$$

$$=\sum_{y\in E^n}P(y|x^i)(1-f(y,x^i))+\sum_{y\in E^n}\sum_{j\ne i}P(y|x^i)f(y,x^j),\tag{3.3}$$

здесь выражение $\bigl[1-f(y,x^i)+\sum\limits_{j\ne i}f(y,x^j)\bigr]$ равно нулю тогда и только тогда, когда найдется единственное кодовое слово x^i такое, что $d(x^i,y)\le\rho$, в противном случае

$$\bigl[1-f(y,x^i)+\sum_{j\ne i}f(y,x^j)\bigr]\ge 1.$$

Первая сумма в неравенстве (3.3) равна вероятности того, что полученное на приемном конце слово находится на расстоянии большем ρ от переданного кодового слова x. Согласно неравенству (3.2), оно не превышает $\varepsilon/2$. Таким образом,

$$P_C\le\frac{\varepsilon}{2}+\frac{1}{M}\sum_{i=1}^{M}\sum_{y\in E^n}\sum_{j\ne i}P(y|x^i)f(y,x^j).$$

Основная идея дальнейшего доказательства состоит в том, что величина

$$P^*(M,n,p):=\min_{C\in\widetilde{\mathcal{L}}}\{P_C\}$$

меньше, чем ожидаемое значение, т. е. меньше математического ожидания P_C над ансамблем $\widetilde{\mathcal{L}}$ всех возможных кодов C длины n, мощности M, взятых случайно. Отсюда имеем

$$P^*(M,n,p)\le\frac{\varepsilon}{2}+\frac{1}{M}\sum_{i=1}^{M}\sum_{y\in E^n}\sum_{j=1,j\ne i}^{M}\mathcal{E}(f(y,x^j))\mathcal{E}(P(y|x^i))=$$

$$= \frac{\varepsilon}{2} + \frac{1}{M}\sum_{i=1}^{M}\sum_{y\in E^n}\sum_{j=1,j\neq i}^{M}\frac{|B_\rho|}{2^n}\mathcal{E}(P(y|x^i)) =$$

$$= \frac{\varepsilon}{2} + \frac{|B_\rho|}{M\cdot 2^n}\sum_{i=1}^{M}\sum_{y\in E^n}\sum_{j=1,j\neq i}^{M}\mathcal{E}(P(y|x^j)) =$$

$$= \frac{\varepsilon}{2} + \frac{|B_\rho|}{M\cdot 2^n}\sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M}\mathcal{E}\left(\sum_{y\in E^n}P(y|x^j)\right) =$$

$$= \frac{\varepsilon}{2} + \frac{|B_\rho|\cdot M\cdot(M-1)}{M\cdot 2^n} \leq \frac{\varepsilon}{2} + M\frac{|B_\rho|}{2^n}.$$

Таким образом, $P^*(M,n,p) - \frac{\varepsilon}{2} \leq M\cdot\frac{1}{2^n}\cdot|B_\rho|$. Логарифмируя обе части, применяя лемму 3 и деля на n, получаем

$$\frac{1}{n}\log(P^*(M,n,p) - \frac{\varepsilon}{2}) \leq \frac{1}{n}\log M - (1-\mathcal{H}(p)) - O\Big(\frac{1}{\sqrt{n}}\Big).$$

Подставляя $M = M_n = 2^{[R\cdot n]}$ в правую часть (вспомним, что по условию число R сколь угодно близко к пропускной способности $C(p) = 1-\mathcal{H}(p)$), получаем

$$\frac{1}{n}\log(P^*(M,n,p) - \frac{\varepsilon}{2}) < -\beta < 0,$$

где β — константа, равная $C(p) - R$. Отсюда $P^*(M,n,p) < \frac{\varepsilon}{2} + 2^{-\beta n}$. Начиная с некоторого n будет выполняться $2^{-\beta n} < \frac{\varepsilon}{2}$, и, следовательно, $P^*(M,n,p) < \varepsilon$. Таким образом,

$$P^*(M_n,n,p) \to 0 \text{ при } n\to\infty.$$

Теорема доказана.

Глава 4

Свитчинговые методы

4.1. Коды Васильева

В 1959 г. Г. С. Шапиро и Д. С. Злотник предположили, что не существует совершенных кодов, не эквивалентных коду Хэмминга. В 1962 г. Ю. Л. Васильев опроверг эту гипотезу, предложив богатый класс неэквивалентных совершенных двоичных кодов. Рассмотрим этот итеративный способ построения для кодов с произвольными кодовыми расстояниями. Беря в качестве исходных кодов коды со специфическими параметрами, можно получить такие хорошие коды, как совершенные или коды Рида-Маллера.

Рассмотрим произвольные двоичные коды B и C длины n с кодовыми расстояниями d_1 и d_2 соответственно, где d_1 нечетно. Пусть λ — произвольная функция из кода C в множество $\{0,1\}$, $|x| = x_1 + \ldots + x_n (\mathrm{mod}\, 2)$, где $x = (x_1, \ldots, x_n)$.

Теорема 13 *Множество*

$$C^{2n+1} = \{(x + y, |x| + \lambda(y), x) \mid x \in B, y \in C\} \tag{4.1}$$

является двоичным кодом длины $2n+1$, *мощности* $|B| \cdot |C|$ *с кодовым расстоянием* $d \geq min\{2d_1 + 1, d_2\}$.

Доказательство. Проверим параметры построенного кода, а именно его длину, мощность кода и кодовое расстояние.

1. Легко видеть, что $2n + 1$ является длиной кода.

2. Поскольку x и y независимо пробегают множества B и C соответственно, мощность кода C^{2n+1}, очевидно, равна

$$|C^{2n+1}| = |B| \cdot |C|.$$

3. Проверим, что кодовое расстояние равно $d \geq \min\{2d_1 + 1, d_2\}$. Рассмотрим два произвольных различных кодовых слова:

$$u = (x + y, |x| + \lambda(y), x),$$

$$v = (x' + y', |x'| + \lambda(y'), x').$$

Возможны случаи.

3a. Если $y = y'$ и $x \neq x'$, то, с учетом того что $d(x, x') \geq d_1$ и d_1 нечетно, получаем

$$d(u, v) = d((x, |x|, x), (x', |x'|, x')) \geq 2d_1 + 1.$$

3b. Пусть $y \neq y'$ и $x = x'$. Векторы y, y' принадлежат коду C^n, следовательно, $d(y, y') \geq d_2$ и получаем

$$d(u, v) \geq d(y, y') \geq d_2.$$

3c. Если $y \neq y'$ и $x \neq x'$, то аналогично доказательству кодового расстояния в теореме Плоткина (см. теорему 9), используя предложение 4, получаем

$$d(u, v) \geq d(x, x') + d(x + y, x' + y') =$$

$$= w(x - x') + w(x + y - x' - y') =$$

$$= w(x - x') + w((y - y') + (x - x')) \geq$$

$$\geq w(x - x') + w(y - y') - w(x - x') = w(y - y') \geq d_2.$$

Теорема доказана. ▲

Рассмотрим применение этой конструкции для построения совершенных двоичных кодов.

Теорема 14 (Васильев Ю. Л., 1962). *Пусть $C^{(n-1)/2}$ — произвольный совершенный код длины $(n-1)/2 = 2^m - 1, m \geq 2$ и λ — произвольная функция из кода $C^{(n-1)/2}$ в множество $\{0, 1\}$. Множество*

$$V^n = \{(x + y, |x| + \lambda(y), x) : x \in E^{(n-1)/2}, y \in C^{(n-1)/2}\} \tag{4.2}$$

является совершенным двоичным кодом длины n с кодовым расстоянием 3.

Доказательство этой теоремы аналогично доказательству предыдущей. Код (4.2) будем далее называть *кодом Васильева*. Приведем несколько важных следствий, вытекающих из этой теоремы.

Следствие 2. *При $\lambda \equiv 0$ конструкция Васильева, примененная к коду Хэмминга $\mathcal{H}^{(n-1)/2}$ длины $(n-1)/2$, дает код Хэмминга длины n:*

$$\mathcal{H}^n = \{\, (x + y, |x|, x) : x \in E^{(n-1)/2}, y \in \mathcal{H}^{(n-1)/2}\}.$$

Следствие 3. *Если $\lambda(y) + \lambda(y') \neq \lambda(y + y')$ для некоторых $y, y' \in C^{(n-1)/2}$, то код Васильева длины n является нелинейным.*

Поскольку функция λ произвольна, то, принимая во внимание предыдущие итеративные шаги, т. е. подставляя в формулу (4.2) снова произвольный код Васильева длины $(n-1)/2$, затем произвольный код Васильева длины $(n-3)/4$ и т. д., получаем следующее утверждение.

Следствие 4. *Число D_n различных кодов Васильева длины n удовлетворяет следующей нижней оценке:*

$$D_n \geq 2^{2^{\frac{n+1}{2}-\log(n+1)}} \cdot 2^{2^{\frac{n+5}{4}-\log(n+1)}} \cdot 2^{2^{\frac{n+17}{8}-\log(n+1)}} \dots$$

для достаточно больших n.

Зная нижнюю оценку числа различных кодов длины n, легко вычислить нижнюю оценку числа неэквивалентных кодов с теми же параметрами. Для этого достаточно разделить число различных кодов на число различных автоморфизмов E^n, равное $2^n \cdot n!$, здесь 2^n — число различных сдвигов на векторы из E^n и $n!$ — число различных подстановок на n координатах. Нетрудно из следствия 4 получить следующее утверждение.

Следствие 5. *Для числа N_n неэквивалентных кодов Васильева длины n справедливо*

$$N_n \geq 2^{2^{\frac{n+1}{2}-\log(n+1)}} \cdot 2^{2^{\frac{n+5}{4}-\log(n+1)}}$$

при достаточно больших n.

Эта оценка до 1996 г. оставалась лучшей нижней оценкой числа неэквивалентных совершенных кодов, несмотря на многочисленные усилия многих исследователей.

Для $n = 7$ существует единственный совершенный код длины 7, для $n = 15$ существует 11 неэквивалентных кодов Васильева длины 15 и, по крайней мере, 963 неэквивалентных кода, полученных каскадной конструкцией (см. [13]; о каскадных конструкциях см. следующую главу). Следует отметить, что классификация совершенных кодов даже длины 15 до сих пор не найдена. Обзоры результатов по совершенным кодам могут быть найдены в работах [13, 19].

Упражнение 18. Доказать следствие 4.

Упражнение 19. Доказать следствие 5, используя формулу Стирлинга

$$n^n e^{-n}\sqrt{2n\pi} \leq n! \leq n^n e^{1-n}\sqrt{2n\pi}. \tag{4.3}$$

4.2. Конструкция Моллара

Рассмотрим конструкцию Моллара для двоичных кодов. Пусть P^t и C^m — произвольные двоичные коды длин t и m соответственно с кодовыми расстояниями не менее 3, содержащие нулевые векторы. Пусть

$$x = (x_{11}, x_{12}, \dots, x_{1m}, x_{21}, \dots, x_{2m}, \dots, x_{t1}, \dots, x_{tm}) \in E^{tm}.$$

В этом разделе будем использовать следующую матричную запись для вектора x: i-я строка матрицы X равна $x_{i1}\ x_{i2}\ \dots\ x_{im}$, где $i = 1, \dots, t$. Функции $p_1(x)$ и $p_2(x)$, определенные следующим образом:

$$p_1(x) = (\sigma_1, \sigma_2, \dots, \sigma_t) \in E^t, \ \ p_2(x) = (\sigma'_1, \sigma'_2, \dots, \sigma'_m) \in E^m,$$

где $\sigma_i = \sum_{j=1}^{m} x_{ij}$ и $\sigma'_j = \sum_{i=1}^{t} x_{ij}$, называются *обобщенными проверками на четность*. Пусть f — произвольная функция из P^t в E^m.

Теорема 15 (Моллар М., 1986). *Множество*

$$C^n = \{\,(x, y + p_1(x), z + p_2(x) + f(y)) \mid x \in E^{tm}, y \in P^t, z \in C^m\}$$

является двоичным кодом длины $n = tm + t + m$*, с кодовым расстоянием* 3.

Доказательство. Легко проверить, что код C^n имеет длину $n = tm + t + m$ и мощность

$$|C^n| = |E^{tm}| \cdot |P^t| \cdot |C^m|.$$

Пусть

$$u = (x, y + p_1(x), z + p_2(x) + f(y)),$$
$$u' = (x', y' + p_1(x'), z' + p_2(x') + f(y'))$$

— два произвольных вектора из кода C^n. Покажем, что $d(u, u') \geq 3$.

Возможны приведенные ниже случаи.

1. При $x = x'$ имеем $p_1(x) = p_1(x')$, $p_2(x) = p_2(x')$ и, следовательно, при $y \neq y'$ имеем

$$d(u, u') \geq d(y, y') \geq 3.$$

Аналогично при $z \neq z'$, $y = y'$. В случае $y = y'$, $z = z'$ убеждаемся, что нулевой вектор принадлежит коду C^n.

2. Если $d(x, x') = 1$, то

$$d(p_1(x), p_1(x')) = d(p_2(x), p_2(x')) = 1.$$

Пусть $y \neq y'$, тогда

$$d(y + p_1(x), y' + p_1(x')) \geq 2$$

и имеем $d(u, u') \geq 3$. Пусть $y = y'$, тогда

$$d(z + p_2(x) + f(y), z' + p_2(x') + f(y')) \geq 2$$

и снова имеем $d(u, u') \geq 3$.

3. В случае $d(x, x') = 2$ расстояния $d(p_1(x), p_1(x'))$ и $d(p_2(x), p_2(x'))$ равны 0 или 2, но одновременно оба не могут быть равны нулю. Отсюда получаем, что равенства

$$y + p_1(x) = y' + p_1(x'),$$
$$z + p_2(x) + f(y) = z' + p_2(x') + f(y')$$

не выполняются одновременно и, следовательно, $d(u, u') \geq 3$. ▲

Замечания

1. В случае, когда P^t и C^m — произвольные совершенные двоичные коды длин $t = 2^r - 1$ и $m = 2^s - 1$ соответственно, код C^n является совершенным двоичным кодом.

2. При $m = 1$, $t = 2^r - 1$ конструкция Моллара совпадает с конструкцией Васильева.

3. Существуют совершенные коды Моллара, неэквивалентные совершенным кодам Васильева.

4.3. Общая идея метода свитчинга

Основная идея метода свитчинга состоит в следующем: в произвольном двоичном коде C длины n рассмотрим некоторое подмножество M кодовых слов. Если найдется в E^n подмножество M', отличное от множества M, и множество $C' = (C \setminus M) \cup M'$ является двоичным кодом с параметрами, совпадающими с параметрами кода C, то будем говорить, что код C' получен из кода C *свитчингом* множества M на множество M'. Результирующий код отличен или неэквивалентен исходному.

Удобно рассмотреть общую идею метода свитчинга на примере совершенных кодов. Пусть дано подмножество M в пространстве E^n. Множество M' получено из множества M инвертированием i-й координаты, $i \in N = \{1, 2, \ldots, n\}$, всех слов M. Обозначим его $M' = M + i$. Множество M называется *i-компонентой* кода C, если $K(M) = K(M + i)$, где $K(M)$ — окрестность порядка 1 множества M. Легко видеть, что код $C' = (C \setminus M) \cup (M + i)$ также является совершенным кодом. Будем говорить, что код C' получен из кода C *свитчингом* i-компоненты M, (рис. 5).

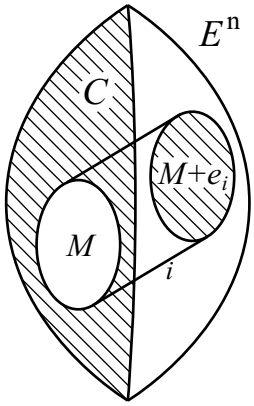

Рис. 5. Свитчинг i-компоненты

Рассмотрим основную идею более общего свитчингового подхода построения кодов (также на примере совершенных кодов), называемого методом *α-компонент*. Пусть $\alpha \subseteq N$. Множество M назовем *α-компонентой* кода C, если для всех $i \in \alpha$ множество M является, в свою очередь, i-компонентой C. Сначала для каждой α-компоненты выбираем свое направление i из множества направлений α и заменяем (делаем "свитчинг") произвольное число i-компонент в каждой α-компоненте на новые i-компоненты такой же мощности. Затем производим замену полученных новых α-компонент на другие α-компоненты, делая свитчинги по неиспользованным из множества α направлениям. В итоге результирующий код остается по-прежнему совершенным, но отличным или даже неэквивалентным исходному. Метод α-компонент оказался особенно подходящим в применении его к коду Хэмминга, поскольку позволяет, разрушая групповую структуру кода Хэмминга, тем не менее следить за структурой нелинейного совершенного кода, получаемого вследствие серии свитчингов.

Впервые свитчинговый способ построения кодов был предложен Ю. Л. Васильевым. Можно показать, что конструкция Моллара также является свитчинговой конструкцией. В 1996 г. С. В. Августиновичем и Ф. И. Соловьевой был предложен способ построения совершенных двоичных кодов посредством метода α-компонент (примененного к коду Хэмминга), который после 30-летнего перерыва дал первое существенное улучшение нижней оценки числа неэквивалентных совершенных кодов.

С помощью этого подхода была решена серия проблем, касающихся структуры совершенных кодов: например, обнаружены совершенные коды с тривиальной группой автоморфизмов, доказано существование несистематических совершенных кодов, кодов полного ранга. Ф. И. Соловьевой доказано существование совершенных двоичных кодов с i-компонентами различной мощности и структуры. Метод α-компонент получил дальнейшее развитие в работах С. А. Малюгина, Д. С. Кротова, С. В. Августиновича (см. подробнее [13]).

4.4. Некоторые свойства совершенных кодов

4.4.1. Дистанционная инвариантность

Код называется *дистанционно инвариантным*, если число $A_i(n)$ всех кодовых слов на расстоянии i от фиксированного кодового слова не зависит от выбора этого кодового слова.

В 1957 г. С. П. Ллойд и независимо в 1959 г. Г. С. Шапиро и Д. С. Злотник доказали, что произвольный совершенный код является дистанционно инвариантным. В этом и следующем параграфах приведем с доказательствами несколько красивых теорем о свойствах совершенных кодов с кодовым расстоянием 3, принадлежащих Г. С. Шапиро и Д. С. Злотнику.

Теорема 16 (Шапиро Г. С., Злотник Д. С., 1959). *Пусть C — произвольный совершенный код с кодовым расстоянием 3. Число кодовых слов на расстоянии k от данного кодового слова $x \in C$ не зависит от выбора этого кодового слова и от выбора кода и зависит только от расстояния k.*

Доказательство. Обозначим число кодовых слов на расстоянии k от кодового слова x через A_k. Без ограничения общности рассмотрим код, содержащий вектор $x = \mathbf{0^n}$, где n — длина кода C. Построим систему линейных уравнений для A_k, $k = 0, \ldots, n$. Все числа A_k однозначно будут вычислены из этой системы.

Рассмотрим k-й слой E^n_k (все векторы веса k) в кубе E^n. Согласно свойству плотной упакованности кода C, все векторы из E^n_k разобьются на следующие три подмножества:

1) кодовые слова веса k. Их число в точности равно A_k;

2) векторы, которые принадлежат сферам, окружающим все кодовые слова из E^n_{k-1}, имеется $(n-k+1)\cdot A_{k-1}$ таких векторов;

3) векторы, принадлежащие сферам с центрами в кодовых словах из E^n_{k+1}, имеется $(k+1)\cdot A_{k+1}$ таких векторов.

Отсюда получаем следующую систему из $n+1$ линейных уравнений с $n+1$ неизвестными:

$$A_0 = 1, A_1 = A_2 = 0,$$

$$\left(\begin{array}{c} n \\ k \end{array}\right) = (k+1)A_{k+1} + A_k + (n-k+1)A_{k-1},$$

$$k = 2, 3, \ldots, n,$$

здесь числа A_k с отрицательными индексами полагаются равными нулю (см. рис. 6).

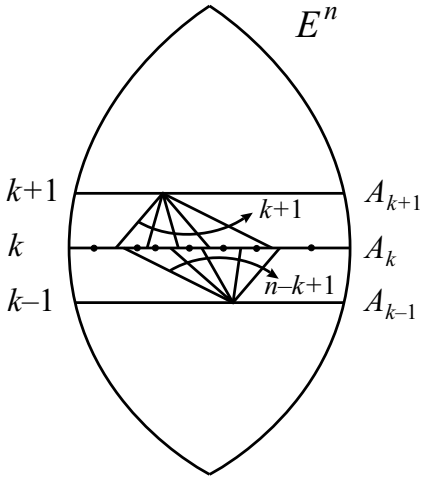

Рис. 6. Иллюстрация к теореме 16

Используя производящую функцию

$$A_0 + A_1 t + \dots A_n t^n,$$

можно найти точный вид A_k, $k = 0, 1, \dots, n$, а именно:

$$A_{2k} = \frac{1}{n+1}\left(\begin{pmatrix} n \\ 2k \end{pmatrix} + (-1)^k n \begin{pmatrix} (n-1)/2 \\ k \end{pmatrix}\right),$$

$$A_{2k+1} = \frac{1}{n+1}\left(\begin{pmatrix} n \\ 2k+1 \end{pmatrix} + (-1)^{k+1} n \begin{pmatrix} (n-1)/2 \\ k \end{pmatrix}\right)$$

и тем самым убедиться, что система имеет единственное решение. ▲

Складывая последние два равенства, получаем

$$A_{2k} + A_{2k+1} = \frac{\begin{pmatrix} n \\ 2k \end{pmatrix} + \begin{pmatrix} n \\ 2k+1 \end{pmatrix}}{n+1},$$

что дает возможность сделать представленный ниже вывод.

Следствие 6. *Произвольный совершенный код длины n, содержащий нулевой вектор, имеет равномерное распределение по парам соседних слоев E^n_{2k} и E^n_{2k+1}, $k = 0, \dots, (n-1)/2$ куба E^n.*

Из этой теоремы вытекают также следующие полезные свойства.

Следствие 7. *Для каждого кодового слова $x \in C$, где C — произвольный совершенный код длины n, антиподальный ему вектор $x + \mathbf{1^n}$ также принадлежит коду C.*

Это свойство антиподальности оказалось чрезвычайно важным для исследования многих нетривиальных свойств совершенных двоичных кодов.

Следствие 8. *Число кодовых слов веса $(n-1)/2$ произвольного совершенного кода длины n, содержащего нулевой вектор, равно*

$$A_{(n-1)/2} = \frac{1}{n+1}\left(\begin{pmatrix} n \\ (n-1)/2 \end{pmatrix} + n \begin{pmatrix} (n-1)/2 \\ (n-3)/4 \end{pmatrix}\right).$$

4.4.2. О существовании совершенных кодов

Здесь рассмотрим несколько теорем о существовании совершенных кодов. Доказательство следующей очень важной теоремы весьма нетривиально и может быть найдено в работе [1].

Теорема 17. О существовании совершенных кодов (Зиновьев В. А., Леонтьев В. К., Тиетвайнен А., 1972). *Нетривиальный совершенный код над любым полем Галуа $GF(q)$ должен иметь те же самые параметры, что и один из кодов Хэмминга или Голея, т. е.:*

1) q-значный $(n = (q^m - 1)/(q - 1), n - m, 3)$-код;

2) двоичный $[23, 12, 7]$-код Голея;

3) троичный $[11, 6, 5]_3$-код Голея.

Оба кода Голея единственны с точностью до эквивалентности и существует много неэквивалентных совершенных q-значных кодов, $q \geq 2$ (см. также следствие 5). Под тривиальным совершенным кодом понимается код, состоящий либо из одного кодового слова, либо из двух антиподальных (в случае, если n нечетно). В 1949 г. М. Ж. И. Голей построил двоичный совершенный $[23, 12, 7]$-код.

Теорема 18 (Шапиро Г. С., Злотник Д. С., 1959). *Единственными совершенными двоичными кодами с расстоянием 7 является код с параметрами кода Голея длины 23 и тривиальный код длины 7.*

Доказательство. Если существует совершенный двоичный код длины n, размерности k, с кодовым расстоянием 7, то

$$2^n : \left(1 + \left(\begin{array}{c} n \\ 1 \end{array}\right) + \left(\begin{array}{c} n \\ 2 \end{array}\right) + \left(\begin{array}{c} n \\ 3 \end{array}\right)\right) = 2^k$$

и, следовательно,

$$1 + \left(\begin{array}{c} n \\ 1 \end{array}\right) + \left(\begin{array}{c} n \\ 2 \end{array}\right) + \left(\begin{array}{c} n \\ 3 \end{array}\right) = 2^r,$$

где $r = n - k$. Умножая на 6 и преобразовывая левую часть последнего равенства, получаем

$$(n^2 - n + 6)(n + 1) = 3 \cdot 2^{r+1}.$$

Следовательно, первый или второй сомножители в левой части этого равенства делятся на 3.

Имеется два случая.

1. Пусть $3|(n^2 - n + 6)$. Тогда

$$n + 1 = 2^l, \;\; n^2 - n + 6 = 3 \cdot 2^{r-l+1}$$

и, значит, выполняется

$$(2^l - 1)^2 - (2^l - 1) + 6 = 3 \cdot 2^{r-l+1}$$

и

$$2^{2l} - 3 \cdot 2^l + 8 = 3 \cdot 2^{r-l+1}. \tag{4.4}$$

При $l = 3$ имеем тривиальный код длины $n = 7$. Пусть $l > 3$ и $n > 7$. Из равенства (4.4) имеем

$$2^3(2^{2l-3} - 3\cdot 2^{l-3} + 1) = 3\cdot 2^{r-l+1}.$$

Здесь первый сомножитель в левой части сравним с нулем по модулю 2, второй — с единицей по тому же модулю. Анализируя сомножители правой части, приходим к заключению, что $2^3 = 2^{r-l+1}$ и, следовательно, $r - l + 1 = 3$. Отсюда

$$n^2 - n + 6 = 3\cdot 2^3,$$

что противоречит $n > 7$ (n должно быть целым числом).

2. Пусть $3|(n + 1)$. В этом случае имеем

$$n + 1 = 3\cdot 2^l,\ n^2 - n + 6 = 2^{r-l+1},$$

откуда

$$(3\cdot 2^l - 1)^2 - (3\cdot 2^l - 1) + 6 = 2^{r-l+1}$$

и

$$9\cdot 2^{2l} - 9\cdot 2^l + 8 = 2^{r-l+1},$$

$$2^3(9\cdot 2^{2l-3} - 9\cdot 2^{l-3} + 1) = 2^{r-l+1},$$

$$9\cdot 2^{2l-3} - 9\cdot 2^{l-3} + 1 = 2^{r-l-2},$$

$$9\cdot 2^{2l-3} - 9\cdot 2^{l-3} = 2^{r-l-2} - 1.$$

Для равенства должно выполняться $2^{l-3} = 1$. Отсюда $l = 3$ и $n+1 = 3\cdot 2^3 = 24$. Иными словами, получаем возможность для существования кода длины 23, с кодовым расстоянием 7, что завершает доказательство теоремы. ▲

В следующей теореме показано неконструктивным методом, что количество совершенных двоичных кодов с расстоянием более 4 конечно. Доказательство теоремы является следствием одного глубокого результата Зигеля из теории чисел.

Лемма 4 (Зигель). *Пусть $f(x)$ — многочлен, принимающий целые значения при целых значениях переменной x. Если $f(x)$ не является степенью линейного многочлена, умноженного на константу, то наибольший простой делитель числа $f(n)$ неограниченно возрастает при $n \to \infty$.*

Теорема 19 (Шапиро Г. С., Злотник Д. С., 1959). *Количество совершенных двоичных кодов длины n, с кодовым расстоянием $d \geq 5$ конечно.*

Доказательство. Чтобы доказать теорему, мы должны убедиться, что многочлен $f(x)$, определенный как

$$f(x) = 1 + \left(\begin{array}{c} x \\ 1 \end{array}\right) + \ldots + \left(\begin{array}{c} x \\ t \end{array}\right), \tag{4.5}$$

не является степенью линейного многочлена при $t \geq 2$ (здесь правая часть равенства (4.5) является числом векторов в шаре радиуса t в x-мерном кубе E^x), где

$d = 2t + 1$ — кодовое расстояние. Тогда, согласно лемме 4, из равенства (4.5) получаем, что число $f(n)$ имеет простой множитель, больший двух при n достаточно большом, и, следовательно, не может быть равным степени двойки; значит, $2^n/f(n) \neq 2^k$ ни для какого k. Это означает, что граница Хэмминга не может быть достигнута и не существует совершенного кода длины n с расстоянием d.

Докажем теорему от противного. Пусть

$$f(x) = a(b + cx)^t, \tag{4.6}$$

где a, b, c — некоторые рациональные числа. Вычислим $f(0)$ из последнего соотношения:

$$f(0) = 1 = ab^t,$$

т. е. $f(x)$ можно переписать в виде

$$f(x) = (1 + r\cdot x)^t, \tag{4.7}$$

где $r = c/b$ — рациональное число.

Подставляя $x = 1$ в равенство (4.5), получаем

$$f(1) = 1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2.$$

С другой стороны, из представления (4.7) имеем

$$f(1) = (1 + r)^t.$$

Таким образом, $(1 + r)^t = 2$, откуда $1 + r = \sqrt[t]{2}$ рационально. Это противоречие доказывает теорему. ▲

4.4.3. Верхняя оценка числа совершенных кодов

К сожалению, имеется только верхняя оценка числа совершенных двоичных кодов, близкая к тривиальной, хотя доказательство этого факта далеко от тривиального и основано на одном важном свойстве совершенных кодов, к описанию которого приступим.

Лемма 5 (Августинович С. В., 1995). *Произвольный совершенный двоичный код длины n, содержащий нулевой вектор, однозначно определяется множеством своих кодовых слов веса $(n-1)/2$.*

Доказательство. Как и прежде, обозначим все двоичные векторы веса k через E_k^n и рассмотрим множество $X_{\frac{n-1}{2}} = C \cap E^n_{\frac{n-1}{2}}$ всех кодовых слов веса $(n-1)/2$ в совершенном коде C, содержащем $\mathbf{0}^n$. Прежде всего, следует заметить, что если известно множество $X_{\frac{n-1}{2}}$, то, согласно следствию 7, множество $\overline{X}_{\frac{n-1}{2}}$ является подмножеством кода C, где $\overline{X}_{\frac{n-1}{2}}$ — множество всех антиподальных слов множеству слов $X_{\frac{n-1}{2}}$.

Пусть существует по крайней мере два продолжения множества $X_{\frac{n-1}{2}} \cup \overline{X}_{\frac{n-1}{2}}$ до совершенных кодов:

$$C = A \cup X_{\frac{n-1}{2}} \cup \overline{X}_{\frac{n-1}{2}} \cup \overline{A}, \tag{4.8}$$

$$C' = B \cup X_{\frac{n-1}{2}} \cup \overline{X}_{\frac{n-1}{2}} \cup \overline{B},$$

где $A \neq B$.

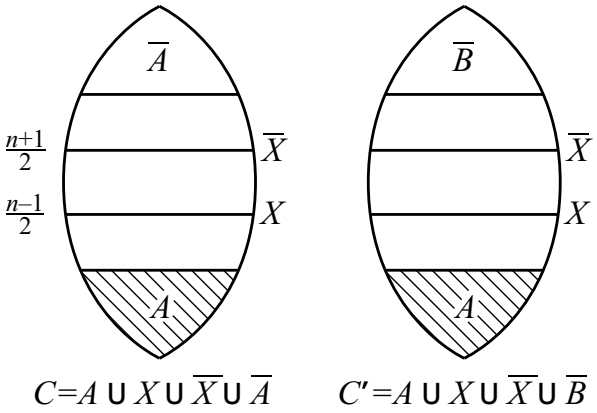

Рис. 7. Иллюстрация к предложению 5

Легко видно, что $d(A, \overline{B}) \geq 3$. Используя этот факт, можно построить совершенный код

$$D = A \cup X_{\frac{n-1}{2}} \cup \overline{X}_{\frac{n-1}{2}} \cup \overline{B} \tag{4.9}$$

(рис. 7).

Из $A \neq B$ имеем $\overline{A} \neq \overline{B}$ и, следовательно, найдется вектор $y \in \overline{B}$ такой, что $y \notin \overline{A}$. Отсюда, из соотношения 4.9 и свойства антиподальности совершенного кода (см. следствие 7), получаем $\overline{y} \notin A$. Снова в силу антиподальности совершенного кода из равенства 4.9 имеем $\overline{y} \in D$, следовательно, $\overline{y} \in A$, противоречие. ▲

Используя это свойство, можно доказать верхнюю оценку N_n числа различных совершенных двоичных кодов длины n.

Теорема 20 (Августинович С. В., 1995). *Число N_n различных совершенных двоичных кодов длины n удовлетворяет неравенству*

$$N_n \leq 2^{2^{n-\frac{3}{2}\log n + \log\log(en)}}.$$

Доказательство. Из леммы 5 легко получить следующую верхнюю оценку числа различных совершенных двоичных кодов длины n:

$$N_n \leq \begin{pmatrix} |E^n_{(n-1)/2}| \\ |E^n_{(n-1)/2} \cap C^n| \end{pmatrix}.$$

Используя дважды формулу Стирлинга

$$n^n e^{-n}\sqrt{2\pi n} \leq n! \leq n^n e^{1-n}\sqrt{2\pi n}$$

и следствие 8, получаем

$$N_n \leq \begin{pmatrix} |E^n_{(n-1)/2}| \\ A_{\frac{n-1}{2}} \end{pmatrix} \leq \begin{pmatrix} 2^n/\sqrt{n} \\ 2^n/n\sqrt{n} \end{pmatrix} \leq 2^{2^{n-\frac{3}{2}\log n + \log\log(en)}}. \tag{4.10}$$

Замечания

1. Сравнивая эту оценку с лучшей нижней оценкой числа различных совершенных двоичных кодов, убеждаемся в большом разрыве между ними.

2. Тривиальная верхняя оценка имеет вид

$$2^{2^{n-\log n}}.$$

Упражнение 20. Доказать, используя приведенный выше вариант формулы Стирлинга, неравенство

$$\left(\begin{array}{c} n \\ (n-1)/2 \end{array} \right) \leq \frac{2^n}{\sqrt{n}}.$$

Упражнение 21. Доказать, используя формулу Стирлинга, что число $A_{\frac{n-1}{2}}$ из следствия 8 удовлетворяет неравенству

$$A_{\frac{n-1}{2}} \leq c\frac{2^n}{n\sqrt{n}},$$

где c — некоторая константа.

Упражнение 22. Доказать, используя формулу Стирлинга, неравенство (4.10).

Упражнение 23. Доказать, что базовое множество кода Хэмминга, состоящее из кодовых слов веса 3, может быть построено индуктивно из представления кода Хэмминга посредством конструкции Васильева.

Нерешенная проблема

Найти новую верхнюю оценку числа различных совершенных двоичных кодов длины $n \geq 15$.

Глава 5

Каскадные методы

5.1. Основная идея каскадного способа построения

Каскадный метод построения кодов впервые был предложен в 1966 г. Г. Д. Форни, затем, в начале 70-х гг., теория каскадных и обобщенных каскадных кодов была развита В. В. Зябловым, Э. Л. Блохом, В. А. Зиновьевым.

Рассмотрим основную идею каскадного способа построения кодов.

Пусть A является q-значным $(n, |A|, d)$-кодом, т. е. кодом длины n, мощности $|A|$, с кодовым расстоянием d. Пусть B — q'-значный $(N, |B|, d')$-код, где $|B| = q$. Обозначим кодовые слова кода B следующим образом: $B = \{\mathbf{b}(0), \ldots, \mathbf{b}(q-1)\}$. Для любого кодового слова $\mathbf{a} = (a_1, \ldots, a_n) \in A$ построим вектор $\mathbf{a}(B) = \big(\mathbf{b}(a_1)|\ldots|\mathbf{b}(a_n)\big)$, где символ | обозначает конкатенацию векторов. Множество $C = \{\mathbf{a}(B) : \mathbf{a} \in A\}$ является q'-значным кодом. Легко найти параметры этого кода: длина равна nN, мощность — $|C| = |A|$ и кодовое расстояние — $d(C) \geq dd'$. Коды A, B и C называются, соответственно, *внешним*, *внутренним* и *каскадным кодами*.

В этой главе будет приведено несколько каскадных методов построения кодов (иногда для простоты изложения рассматривается случай кодов с малыми кодовыми расстояниями), принадлежащих различным авторам. Сначала рассмотрим наиболее простые каскадные методы построения кодов, затем более сложные. В конце главы изложим обобщенную каскадную конструкцию В. А. Зиновьева для нелинейных кодов (обобщенная каскадная конструкция для линейных кодов была предложена ранее В. В. Зябловым и Э. Л. Блохом).

5.2. Коды Соловьевой (1981)

Для определения каскадной конструкции, предложенной Ф. И. Соловьевой в 1981 г., потребуются разбиения n-куба E^n на совершенные коды.

Разбиения E^n на совершенные коды. Рассмотрим произвольный совершенный код C с кодовым расстоянием 3 длины $n = 2^m - 1$ при $m \geq 2$. Используя свойство плотной упакованности совершенного кода в E^n, легко получить следующее тривиальное разбиение E^n на аналоги классов смежности по совершенному коду C:

$$E^n = C \cup (C + e_1) \cup \ldots \cup (C + e_n),$$

где e_i — вектор с единственной ненулевой координатой i. Приведем нетривиальную конструкцию широкого класса нетривиальных разбиений E^n на совершенные коды для любой допустимой длины кода $n > 7$, используя конструкцию Васильева. Обозначим этот класс через $\mathbf{P}^n$.

Теорема 21 (Соловьева Ф. И., 1981). *Существует класс $\mathbf{P}^n$ различных разбиений куба E^n на совершенные коды длины $n \geq 15$, где*

$$|\mathbf{P}^n| \geq 2^{2^{(n-1)/2}}.$$

Доказательство проведем индукцией по m, где $m = \log(n+1)$.

Для $m = 2$ существует лишь тривиальное разбиение, поскольку для $n = 3$ существует единственный совершенный код — это код Хэмминга H^3. К. Т. Фелпсом показано, что при $n = 7$ существует 11 неэквивалентных разбиений E^7 на различные коды Хэмминга длины 7.

Рассмотрим произвольное разбиение $E^{(n-1)/2}$, $m-1 = \log((n+1)/2)$, на совершенные коды длины $(n-1)/2$:

$$E^{(n-1)/2} = \bigcup_{i=0}^{(n-1)/2} C_i^{(n-1)/2}.$$

Перейдем к случаю m. Используя конструкцию Васильева, из каждого кода $C_i^{(n-1)/2}$, $i \in \{0, 1, \ldots, (n-1)/2\}$, построим следующие два совершенных кода длины n. Первый код имеет вид

$$C_i^n = \{(x+y, |x| + \lambda_i(y), x) : x \in E^{(n-1)/2}, y \in C_i^{(n-1)/2}\},$$

второй —

$$C^n_{i+(n+1)/2} = C_i^n + e_{(n+1)/2},$$

где, как и в конструкции Васильева, функция λ_i является произвольной функцией из кода $C_i^{(n-1)/2}$ в множество $\{0, 1\}$. Легко показать, что любые два построенных кода не пересекаются.

Число различных разбиений не меньше числа выборов различных функций λ_i для каждого $i = 0, 1, \ldots, (n-1)/2$. Следовательно,

$$|\mathbf{P}^n| \geq \left(2^{|C_i^{(n-1)/2}|}\right)^{\frac{n+1}{2}} = \left(2^{\frac{2^{(n+1)/2}}{n+1}}\right)^{\frac{n+1}{2}} = 2^{2^{(n-1)/2}},$$

что завершает доказательство. ▲

В дальнейшем в этой главе под совершенными кодами подразумеваются совершенные коды с расстоянием 3.

Теорема 22 (Соловьева Ф. И., 1981). *Пусть*

$$E^n = \bigcup_{i=0}^{n} C_i^n,\ E^n = \bigcup_{i=0}^{n} D_i^n$$

произвольные разбиения куба E^n на совершенные коды длины n и π — произвольная подстановка на множестве $n+1$ координат. Тогда множество

$$C = \{(x, y, |y|) : x \in C_i^n, y \in D_{\pi(i)}^n, i = 0, 1, \ldots, n\}$$

является совершенным двоичным кодом длины $2n+1$.

Доказательство. Легко видеть, что длина кода равна $2n+1 = 2^{m+1}-1$, где $n = 2^m - 1$. Мощность кода равна

$$|C| = (n+1)\cdot|C_i^n|\cdot|D_i^n| = (n+1)\cdot|C_i^n|^2 = (n+1)\cdot\frac{2^{2n}}{(n+1)^2} = \frac{2^{2n+1}}{(2n+1)+1}.$$

Проверим, что кодовое расстояние равно 3. Пусть $u = (x, y, |y|)$ и $v = (x', y', |y'|)$ — произвольные различные кодовые слова кода C. Возможны три случая.

1. Если $x = x', y \neq y', x \in C_i^n, i = 0, 1, \ldots, n$ то $y, y' \in D_{\pi(i)}^n$ и $d(y, y') \geq 3$, значит $d(u, v) \geq 3$.

2. Случай $x \neq x', y = y'$ аналогичен предыдущему.

3a. Пусть $x \neq x', y \neq y'$ и $x, x' \in C_i^n$. Тогда $d(x, x') \geq 3$ и, следовательно, $d(u, v) \geq 3$.

3b. Пусть $x \neq x', y \neq y'$ и $x \in C_i^n$, $x' \in C_j^n$, где $i, j \in \{0, 1, \ldots, n\}$ и $i \neq j$. Тогда $y \in D_{\pi(i)}^n$ и $y' \in D_{\pi(j)}^n$. Если $|y| = |y'|$, то $d(y, y') \geq 2$, $d(x, x') \geq 1$ и, значит, $d(u, v) \geq 3$. При $|y| \neq |y'|$ имеем $d(y, y') \geq 1$, $d((y, |y|), (y', |y'|)) \geq 2$, $d(x, x') \geq 1$ и снова $d(u, v) \geq 3$.

Теорема доказана. ▲

Конструкция легко обобщается с помощью общей проверки на четность на случай расширенных совершенных кодов. Иллюстрацию к теореме 22 для случая, когда C_i и $D_{\pi(i)}$ — расширенные коды, $n = 2^m$ см. на рис. 8.

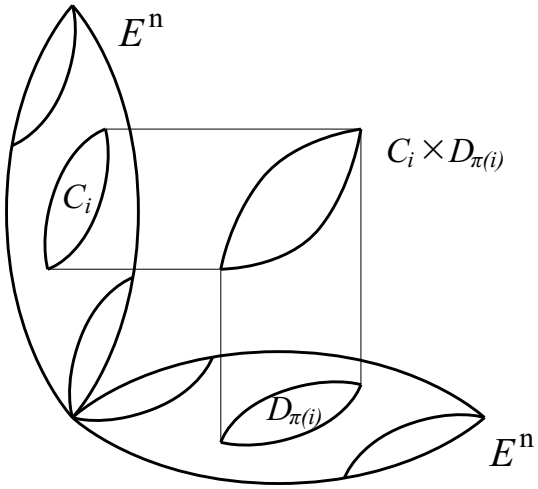

Рис. 8. Случай расширенных кодов

Замечания

1. Несложно показать, что эта конструкция является каскадной.

2. Конструкцию можно обобщить следующим образом: вместо двух разбиений пространства E^n на совершенные коды рассмотреть разбиения двух различных кодов C_1 и C_2 на непересекающиеся подкоды с параметрами некоторых кодов C_3 и C_4 соответственно (см., например, теорему 63 в разд. 9.2.). В случае разбиений кодов C_1 и C_2 на смежные классы по кодам C_3 и C_4 (т. е. тривиальных разбиений) эта конструкция называется *конструкцией* $X4$ (см. [1, гл. 18]).

3. Следует отметить, что, используя эту конструкцию, можно построить класс разбиений E^n на совершенные двоичные коды.

4. Класс совершенных кодов, описанный в этом разделе, не эквивалентен классу кодов Васильева. В 1984 г. К. Т. Фелпс обобщил эту конструкцию. В 2000 г. он доказал, что существует по крайней мере 963 и не более 15 408 неэквивалентных кодов Соловьевой длины 15. В 2004 г. В. А. Зиновьев и Д. В. Зиновьев доказали, что существует в точности 758 совершенных кодов длины 15 ранга 13. В 2006 г. С. А. Малюгин построил 55 совершенных кодов длины 15 ранга 15.

Упражнение 24. Доказать, что $C_i^n \cap C_j^n = \emptyset$ для любых $i, j \in \{1, 2, \ldots, (n-1)/2\}$, $i \neq j$ в теореме 21.

Упражнение 25. Построить класс разбиений куба E^n на совершенные двоичные коды, используя теорему 22.

Упражнение 26. Обобщить каскадную конструкцию теоремы 22 для расширенных совершенных кодов.

Упражнение 27. Как построить код Хэмминга, используя конструкцию теоремы 22?

5.3. Коды Романова

Рассмотрим применение каскадной конструкции для кодов, не являющихся совершенными, на примере кода длины 16, который имеет максимальную мощность среди всех кодов такой длины, исправляющих одну ошибку.

Хорошо известно, что существует разбиение множества D_3^9 всех двоичных слов длины 9 веса 3 на семь систем троек Штейнера порядка 9. Обозначим эти системы троек Штейнера через $S_i, i = 1, \ldots, 7$. Таким образом, имеем

$$D_3^9 = \bigcup_{i=1}^{7} S_i.$$

Рассмотрим также разбиение куба E^7 на классы смежности по коду Хэмминга H^7:

$$E^7 = \bigcup_{i=0}^{7} (H^7 + e_i).$$

Пусть S_i' — множество всех антиподальных слов к словам множества S_i, т. е.

$$S_i' = \{z + \mathbf{1}^9 | z \in S_i\}.$$

Теорема 23 (Романов А. М., 1983). *Множество* C^{16}*, определенное как*

$$C^{16} = \{(x, y) : x \in S_i \cup S_i',\ y \in H^7 + e_i,\ i = 1, \ldots, 7\} \cup \{(x, y) : x \in \{\mathbf{0}^7, \mathbf{1}^7\},\ y \in H^7\}$$

является кодом, исправляющим одну ошибку длины 16 *мощности* 2720*.*

Доказательство опустим, поскольку оно аналогично доказательству теоремы 22.

Применение конструкции Плоткина к этому коду позволяет получить хорошие коды длины n, где $2^m \leq n \leq 2^m + 2^{m-4}$.

Подставляя полный четновесовой код D длины 17 и расширенный код Романова C длины 17 в конструкцию Плоткина, получаем код длины 34 со следующими хорошими параметрами:

$$D : (17,\, 2^{16},\, 2),\ \ C : (17,\, \frac{85}{64} \cdot 2^{11},\, 4) \implies (34,\, \frac{85}{64} \cdot 2^{27},\, 4).$$

Укорачивая этот код дважды, получаем коды длин 33 и 32 со следующими параметрами:

$$(34,\, \frac{85}{64} \cdot 2^{27},\, 4) \implies (33,\, \frac{85}{64} \cdot 2^{27},\, 3) \implies (32,\, \frac{85}{64} \cdot 2^{26},\, 3).$$

Используя эти коды в качестве первого шага индукции, индукцией по $m = \log n$ можно доказать следующий факт.

Теорема 24 (Романов А. М., 1983). *Для любого целого числа* n*, удовлетворяющего* $2^m \leq n \leq 2^m + 2^{m-4} - 1$*, существует нелинейный* $(n, \lambda \cdot 2^{n-m-1}, 3)$ *код, где* $\lambda = 85/64$*.*

Следует отметить, что для кодов длины больше 16 существуют другие коды с хорошими параметрами, исправляющие одну ошибку. Например код Этциона длины 17, мощности 5312, с кодовым расстоянием 3 и код Хямяляйнена длины 18, мощности 10496, с кодовым расстоянием 3 (см. [19]). Используя конструкцию Плоткина и эти коды, можно аналогичным образом получить бесконечные классы кодов с хорошими параметрами.

Упражнение 28. Доказать теорему 23.

5.4. Коды Хямяляйнена

Для изложения конструкции Хямяляйнена нам потребуется q-значный код Хэмминга.

5.4.1. Код Хэмминга над $GF(q)$

Основная идея построения проверочной матрицы q-значного кода Хэмминга, где $q > 2$, такая же, как и для двоичного кода Хэмминга. В качестве столбцов проверочной матрицы возьмем все q-значные векторы длины m такие, что любые два из них линейно независимы и найдутся три линейно зависимых. В отличие от двоичного случая, мы не можем брать все ненулевые векторы длины m, поскольку некоторые из них могут быть линейно зависимыми. Например, векторы (11011) и (22022) линейно зависимы над полем Галуа $GF(3)$. В целях устранения этой ситуации возьмем, например, в качестве столбцов проверочной матрицы все ненулевые столбцы такие, что первая ненулевая координата в каждом из них равна 1. Количество ненулевых векторов длины m над $GF(q)$ равно $q^m - 1$. Нетрудно показать, что среди них имеем

$$1 + q + q^2 + \ldots + q^{m-1} = (q^m - 1)/(q - 1)$$

векторов с предписанным свойством. Следовательно, длина q-значного кода Хэмминга с m проверочными символами равна $n = (q^m - 1)/(q - 1)$, мощность кода равна q^{n-m} и по построению кодовое расстояние 3. Таким образом, мы построили код с параметрами

$$[n = (q^m - 1)/(q - 1), q^{n-m}, d = 3]_q.$$

Обозначим его через $\mathcal{H}_q^n$.

Пример 2. Построим код Хэмминга над $GF(3)$ с двумя проверочными символами. Рассмотрим проверочную матрицу в каноническом виде

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}.$$

Она задает троичный код Хэмминга $\mathcal{H}_3^4$ длины 4. Переходя от этой проверочной матрицы к порождающей матрице в каноническом виде

$$G = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 2 \\ 0 & 1 & 2 & 1 \end{pmatrix},$$

построим все кодовые слова. Они имеют вид

$$\alpha_1 x_1 + \alpha_2 x_2,$$

где x_1, x_2 строки матрицы G и $\alpha_1, \alpha_2 \in GF(3)$:

$$\text{информационные блоки} \Longrightarrow \text{кодовые слова}$$

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 2 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 2 & 0 \\ 2 & 1 \\ 2 & 2 \end{pmatrix} \Longrightarrow \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 2 & 1 & 2 \\ 1 & 0 & 2 & 2 \\ 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 1 & 1 \\ 2 & 1 & 0 & 2 \\ 2 & 2 & 2 & 0 \end{pmatrix}.$$

Упражнение 29. Показать, что произвольный q-значный код с параметрами кода Хэмминга является совершенным.

5.4.2. Конструкция Хямяляйнена

Основная идея конструкции Хямяляйнена состоит в следующем: сначала в пятизначном коде Хэмминга $\mathcal{H}_5^6$ с параметрами $[6, 5^4, 3]_5$ ищется с помощью метода включений и исключений подкод максимально возможной мощности над алфавитом из четырех элементов, затем к этому подкоду применяется каскадирование с помощью классов смежности по двоичному коду Хэмминга длины 3. Рассмотрим детально эту красивую комбинаторную конструкцию.

Пусть код Хэмминга $\mathcal{H}_5^6$ длины $n = 6$ над полем Галуа $GF(5)$, задан своей порождающей матрицей в каноническом виде:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 4 \\ 0 & 1 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Произвольное кодовое слово имеет вид

$$\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4,$$

где u_1, u_2, u_3, u_4 — строки из G и $\alpha_i \in \{0, 1, 2, 3, 4\}, i = 1, 2, 3, 4$. Мощность кода равна 5^4. Зафиксируем произвольный элемент k из множества $\{1, 2, 3, 4\}$. Удалим из кода $\mathcal{H}_5^6$ все кодовые слова, содержащие координату, равную k. Используя метод включений и исключений, вычислим число таких кодовых слов в коде Хэмминга $\mathcal{H}_5^6$:

$$5^4 + \sum_{i=1}^{4} (-1)^i \begin{pmatrix} 6 \\ i \end{pmatrix} 5^{4-i} - 1 = 164.$$

Имеется только один вектор с пятью координатами, равными k. Например, при $k = 4$, это вектор (424444). Полученный подкод фактически является сужением кода Хэмминга $\mathcal{H}_5^6$ над алфавитом из четырех элементов $\{0, 1, 2, 3\}$. Он имеет 164 кодовых слова и кодовое расстояние, равное 3.

Для получения кода длины 18 применим следующую каскадную конструкцию к полученному подкоду: вместо элемента 0 подставим слова двоичного кода Хэмминга $\mathcal{H}^3$ длины 3:

$$0 \to \{000, 111\};$$

вместо каждого элемента из множества $\{1, 2, 3\}$ возьмем слова класса смежности по коду $\mathcal{H}^3$ таким образом, что любым двум элементам будут отвечать различные классы смежности. В результате получим двоичный код с параметрами $(18, 10496, 3)$, т. е. длины 18, мощности

$$164 \cdot 2^6 = 10496,$$

с кодовым расстоянием 3.

Таким образом, мы доказали следующее утверждение.

Теорема 25 (Хямяляйнен Х., 1988). *Существует двоичный код с параметрами* $(18, 10496, 3)$.

Укорачивая одну координату в этом коде, получаем код, мощность которого меньше мощности известного кода Этциона длины 17 (описание кода Этциона можно найти в работе [12]).

Упражнение 30. Доказать, что подкод кода Хэмминга $\mathcal{H}_5^6$ над подалфавитом $\{1, 2, 3, 4\}$, который не содержит элемента 0, состоит из 160 кодовых слов. Именно по этой причине в конструкции Хямяляйнена существенно, что $k \neq 0$.

5.5. Каскадная конструкция Зиновьева (1988)

Рассмотрим каскадную конструкцию, предложенную В. А. Зиновьевым в 1988 г. Изложим ее для совершенных кодов (независимо этот метод построения в 1989 г. был предложен Ф. И. Соловьевой). Эта конструкция может быть рассмотрена как обобщение конструкции Хямяляйнена, изложенной в предыдущем разделе.

Пусть A — произвольный q-значный совершенный код с параметрами $(n, |A|, 3)$, $q = 2^k$, (например, код Хэмминга) над полем $GF(2^k), k > 1$ с двумя проверочными символами. Его длина равна $n = 2^k + 1$. Пусть объединение двоичных совершенных кодов $C_0, C_1, \ldots, C_r$ задает разбиение векторного пространства E^r, $r = 2^k - 1$.

Теорема 26 (Зиновьев В. А., 1988). *Множество* C^N, *определенное как*

$$C^N = \bigcup_{(x_1,x_2,\ldots,x_n)\in A} C_{x_1} \times C_{x_2} \times \ldots \times C_{x_n},$$

является двоичным совершенным кодом длины $N = nr = 2^{2k} - 1, k > 1$.

Доказательство. Длина кода, очевидно, равна

$$N = n(q-1) = (2^k + 1)(2^k - 1) = 2^{2k} - 1.$$

Мощность кода несложно вычислить:

$$|C^N| = |\mathcal{H}_q^n| \cdot |C_0|^n = 2^{k2^k-k}(2^{2^k-1-k})^{2^k+1} = 2^{k2^k-k} \cdot 2^{2^{2k}-1-k2^k-k} = 2^{N-\log(N+1)},$$

где $N = 2^{2k} - 1$.

Убедимся, что кодовое расстояние равно 3. Рассмотрим два произвольных кодовых слова

$$x = (x_1, x_2, \ldots, x_n),$$

$$y = (y_1, y_2, \ldots, y_n)$$

из A. Если $x \neq y$, то $d(x, y) \geq 3$ и, значит, найдутся по крайней мере три координаты i, j, k, в которых различаются кодовые слова x и y. Следовательно, существуют три пары кодов в разбиении E^n такие, что

$$d(C_{x_i}, C_{y_i}) \geq 1, \quad d(C_{x_j}, C_{y_j}) \geq 1, \quad d(C_{x_k}, C_{y_k}) \geq 1.$$

Отсюда следует

$$d(C_{x_1} \times C_{x_2} \times \ldots \times C_{x_n}, C_{y_1} \times C_{y_2} \times \ldots \times C_{y_n}) \geq 3.$$

Пусть $x = y$. Тогда имеем следующее множество кодовых векторов кода C^N:

$$C_{x_1} \times C_{x_2} \times \ldots \times C_{x_n}.$$

Учитывая, что каждое множество C_{x_i} является совершенным двоичным кодом, получаем, что расстояние между любыми двумя различными векторами этого множества равно по крайней мере 3. ▲

5.6. Каскадная конструкция Фелпса (1984)

Пусть $C_1^0, C_2^0, \ldots, C_r^0$ и $C_1^1, C_2^1, \ldots, C_r^1$ — произвольные разбиения полных четновесового кода и нечетно-весового кодов пространства E^r (множество всех векторов пространства E^r четного и нечетного веса соответственно) на расширенные совершенные двоичные коды длины r соответственно, пусть C^m — произвольный расширенный совершенный двоичный код длины m в E^m, в этом разделе пусть $r = 2^k$, $m = 2^p$. Для каждого вектора μ из C^m возьмем r-значный код C_μ с кодовым расстоянием 2, длины m, $|C_\mu| = r^{m-1}$ (C_μ является MDS-кодом). Напомним, что MDS-код C длины m, объема $|C|$, с кодовым расстоянием d над $GF(r)$ — это код, достигающий границы Синглтона, т. е. $m - \log_r |C| = d - 1$.

Теорема 27 (Фелпс К. Т., 1984). *Множество C^n, определенное как*

$$C^n = \{(c_1|c_2| \ldots |c_m) : c_i \in C_{j_i}^{\mu_i}, \mu = (\mu_1, \mu_2, \ldots, \mu_m) \in C^m,$$

$$j = (j_1, j_2, \ldots, j_m) \in C_\mu\}$$

является совершенным расширенным двоичным кодом длины $n = mr$.

Далее код C^m будем называть *базовым кодом.*

Доказательство. Для доказательства теоремы используем другое, эквивалентное данному выше определение кода C^n:

$$C^n = \bigcup_{\mu \in C^m} \bigcup_{j \in C_\mu} C_{j_1}^{\mu_1} \times \ldots \times C_{j_m}^{\mu_m}.$$

Очевидно, что длина кода равна $n = mr$.

Также несложно вычислить мощность кода:

$$|C^n| = |(C_{j_i}^{\mu_i})|^m \cdot |C_\mu| \cdot |C^m| = (2^{r - \log r - 1})^m \cdot r^{m-1} \cdot 2^{m - \log m - 1} = 2^{n - \log n - 1}.$$

Здесь $n = mr$.

Несложно заметить, что для двух различных кодовых слов v и v' кода C^n таких, что $\mu = \mu'$, $j = j'$ выполняется $d(v, v') \geq 4$.

Убедимся, что для кодового расстояния справедливо

$$d = d(C_{j_1}^{\mu_1} \times \ldots \times C_{j_m}^{\mu_m}, C_{j'_1}^{\mu'_1} \times \ldots \times C_{j'_m}^{\mu'_m}) \geq 4$$

для любых $\mu, \mu' \in C^m$ и $j, j' \in C_\mu$ таких, что пары (μ, j) и (μ', j') различны.

Возможны приведенные ниже случаи.

1. Предположим, $\mu = \mu'$, $j \neq j'$.

Тогда $d(j, j') \geq 2$ и найдутся координаты s, t такие, что $j_s \neq j_{s'}$, $j_t \neq j_{t'}$. Отсюда, учитывая, что $C_{j_s}^{\mu_s}$ и $C_{j'_s}^{\mu_s}$ одновременно являются четно- или нечетно-весовыми совершенными расширенными двоичными кодами (аналогично для кодов $C_{j_t}^{\mu_t}$ и $C_{j'_t}^{\mu_t}$), имеем $d(C_{j_s}^{\mu_s}, C_{j'_s}^{\mu_s}) \geq 2$ и $d(C_{j_t}^{\mu_t}, C_{j'_t}^{\mu_t}) \geq 2$. Следовательно,

$$d(C_{j_1}^{\mu_1} \times \ldots \times C_{j_s}^{\mu_s} \times C_{j_t}^{\mu_t} \times \ldots \times C_{j_m}^{\mu_m}\ ,\ C_{j'_1}^{\mu_1} \times \ldots \times C_{j'_s}^{\mu_s} \times C_{j'_t}^{\mu_t} \times \ldots \times C_{j'_m}^{\mu_m}) \geq 4.$$

2. Пусть $\mu \neq \mu'$.

Векторы μ и μ' принадлежат базовому коду C^m, т. е. $d(\mu, \mu') \geq 4$ и найдутся четыре координаты t, s, e и l, в которых различаются μ и μ'. Следовательно, имеются четыре пары совершенных кодов $C_{j_i}^{\mu_i}$ и $C_{j'_i}^{\mu'_i}$, $i \in \{t, s, e, l\}$ такие, что

$$d(C_{j_i}^{\mu_i}, C_{j'_i}^{\mu'_i}) \geq 1.$$

В итоге имеем $d \geq 4$. ▲

Замечания

1. Выкалывание произвольной координаты кода C^n приводит к совершенному двоичному коду длины $mr - 1$.

2. Для $m = 2$ код C^m является тривиальным "расширенным совершенным" кодом, состоящим из одного вектора (01). Код C_μ является r-значным кодом длины 2, с кодовым расстоянием 2, которому отвечает подстановка π на r элементах

$$C_\mu = \{(1, \pi(1)), (2, \pi(2)), \ldots, (r, \pi(r))\}.$$

Таким образом, теорема 27 является обобщением теоремы 22.

3. Число неэквивалентных совершенных расширенных кодов длины n полученных по теореме 27 будет не менее

$$2^{2^{\frac{n+1}{2}(1-\varepsilon_n)}},$$

где $\varepsilon_n \to 0$ при $n \to \infty$.

4. Обобщение конструкции Фелпса было дано Д. С. Кротовым в 2000 г.

5.7. Обобщенная каскадная конструкция

Пусть B является q_B-значным $\left(n_B, K_1, d_{B,1}\right)$ кодом. Предположим, что код B разбивается на q_1 подкодов:

$$B = \bigcup_{i=0}^{q_1-1} B_i,$$

где B_i является q_B-значным $(n_B, K_2, d_{B,2})$ кодом, $i = 0, 1, \ldots, q_1 - 1$.

Предположим далее, что B_i разбивается на q_2 подкодов: для $i = 0, 1, \ldots, q_1 - 1$ имеем

$$B_i = \bigcup_{j=0}^{q_2-1} B_{i,j},$$

где $B_{i,j}$ является q_B-значным $\left(n_B, K_3, d_{B,3}\right)$ кодом, $K_3 = q_3$.

Пусть произвольное кодовое слово $\mathbf{b} \in B$ имеет номер k в $B_{i,j}$, тогда тройка

$$(i, j, k) \in \{0, \ldots, q_1 - 1\} \times \{0, \ldots, q_2 - 1\} \times \{0, \ldots, q_3 - 1\}$$

однозначно характеризует вектор $\mathbf{b}$, это можно записать как $\mathbf{b} = \mathbf{b}(i, j, k)$.

Для каждого $\ell = 1, 2, 3$, рассмотрим q_ℓ-значный $\left(n_A, K_{A,\ell}, d_{A,\ell}\right)$ код A_ℓ и его произвольное кодовое слово $\mathbf{a}^\ell = (a_1^\ell, \ldots, a_{n_A}^\ell) \in A_\ell$. Для любого $s = 1, \ldots, n_A$ тройка (a_s^1, a_s^2, a_s^3) дает кодовое слово $\mathbf{b} = \mathbf{b}(a_s^1, a_s^2, a_s^3)$ из B.

Пусть

$$C = \{\left(\mathbf{b}(a_1^1, a_1^2, a_1^3)| \ldots |\mathbf{b}(a_{n_A}^1, a_{n_A}^2, a_{n_A}^3)\right) : \mathbf{a}^\ell \in A_\ell, 1 \le \ell \le 3\}. \tag{5.1}$$

Теорема 28 (Зиновьев В. А., 1975). *Множество C является q_B-значным кодом длины $n_C = n_A n_B$, мощности $K_{A,1}K_{A,2}K_{A,3}$, с кодовым расстоянием*

$$d_C \ge \min\{d_{A,1}d_{B,1}, d_{A,2}d_{B,2}, d_{A,3}d_{B,3}\}.$$

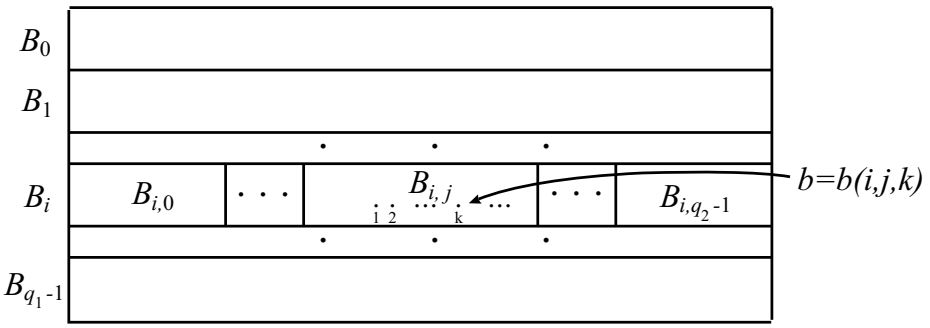

Рис. 9 Иллюстрация к теореме 28

Рассмотрим двоичный случай. Пусть $B = E^{n_B}$, $B = E_\bullet^{n_B} \cup (E^{n_B} \setminus E_\bullet^{n_B})$, где $E_\bullet^{n_B}$ — полный четно-весовой код B, $n_B = 2^m$. Рассмотрим произвольные разбиения кодов $E_\bullet^{n_B}$ и $E^n \setminus E_\bullet^{n_B}$ на 2^m расширенных совершенных кодов длины n_B.

Пусть A_1 — произвольный расширенный совершенный двоичный код $(n_A, 2^{n_A-1-u}, 4)$, $n_A = 2^u$. Пусть A_2 является n_B-значным $(n_A, n_B^{n_A-1}, 2)$ кодом (MDS кодом с расстоянием 2) и A_3 является q_3-значным $(n_A, q_3^{n_A}, 1)$ кодом, где $q_3 = 2^{n_B-1-m}$.

Используя конструкцию последней теоремы, формула (5.1) позволяет получить расширенный совершенный двоичный код C длины 2^{m+n}.

Теорема 29 (Зиновьев В. А., Лобстейн А., 1997). *Код C является расширенным совершенным двоичным кодом C длины 2^{m+n}.*

Замечания

1. Перечисление расширенных совершенных двоичных кодов длины 16, полученных обобщенной каскадной конструкцией, было получено В. А. Зиновьевым и Д. В. Зиновьевым в 2002 г., а именно ими было доказано, что существует 285 неэквивалентных таких кодов.

2. Следует отметить, что конструкция Фелпса может быть описана в терминах обобщенной каскадной конструкции Зиновьева.

Глава 6

Поля Галуа

6.1. Основные определения

В этой главе приведем необходимые определения и утверждения о полях Галуа (см. [1, 2, 8, 21, 24]), которые потребуются в дальнейшем для изложения теории циклических кодов.

Определение. Алгебраическая система $< F; +, \cdot >$ называется *полем*, если:

(i) $< F; + >$ является коммутативной группой по сложению (с единичным элементом 0),

(ii) $< F \setminus \{0\}; \cdot >$ является коммутативной группой по умножению (с единичным элементом 1),

(iii) Выполнены законы дистрибутивности: для любых элементов $a, b, c \in F$ справедливо

$$a(b + c) = ab + ac,$$

$$(b + c)a = ba + ca.$$

Порядком поля называется число его элементов. Поле F называется *конечным*, если оно имеет конечный порядок. В противном случае поле называется *бесконечным*.

Примерами бесконечных полей являются $< \mathbb{Q}; +, \cdot >$, $< \mathbb{R}; +, \cdot >$, $< \mathbb{C}; +, \cdot >$, где $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ обозначают множества рациональных, вещественных и комплексных чисел соответственно, а операции $+$ и $\cdot$ являются обычными операциями сложения и умножения. Примером конечного поля является кольцо вычетов $< \mathbb{Z}_p; +, \cdot >$ целых чисел по модулю простого числа p. Далее такое поле будем называть *простым* и обозначать через F_p.

Определение. *Характеристикой* поля F называется наименьшее положительное целое число p такое, что в поле F справедливо равенство

$$\underbrace{1 + \ldots + 1}_{p \text{ раз}} = 0.$$

Поскольку в поле нет делителей нуля, характеристика p всегда является простым числом. Если такого числа не существует, то говорят, что поле F имеет характеристику 0. Очевидно, любое конечное поле имеет характеристику отличную от нуля.

Определение. Подкольцо $< I; +, \cdot >$ произвольного кольца $< R; +, \cdot >$ называется *идеалом*, если для любых элементов $u \in I$, $v \in R$ выполняется $u \cdot v \in I$ и $v \cdot u \in I$.

Например, подкольцо $< p\mathbb{Z}; +, \cdot >$ является идеалом кольца целых чисел $< \mathbb{Z}; +, \cdot >$.

Пусть $< F; +, \cdot >$ — некоторое поле. Рассмотрим кольцо $F[x]$ всех многочленов от переменной x с коэффициентами из поля F. Пусть $s(x)$ — произвольный многочлен из $F[x]$. Тогда множество

$$(s(x)) = \{\, c(x) \cdot s(x) \,|\, c(x) \in F[x] \,\} \tag{6.1}$$

образует идеал кольца $F[x]$. Верно и обратное, любой идеал кольца $F[x]$ представим в виде совокупности произведений многочленов (6.1) для подходящего многочлена $s(x)$. Без ограничения общности можно считать, что $s(x)$ многочлен наименьшей степени в идеале $(s(x))$. Пользуясь алгоритмом Евклида для многочленов, можно показать, что в любом идеале такой многочлен $s(x)$ существует и единствен. Рассмотрим фактор-кольцо $F[x]/(s(x))$ кольца всех многочленов $F[x]$ по модулю идеала $(s(x))$. Элементами фактор-кольца $F[x]/(s(x))$ являются всевозможные многочлены степени меньшей, чем степень $s(x)$, а операции сложения и умножения в фактор-кольце производятся по модулю многочлена $s(x)$. Если степень многочлена $s(x)$ равна m и поле F конечно, то фактор-кольцо $F[x]/(s(x))$ содержит в точности $|F|^m$ элементов.

Определение. Многочлен $f(x)$ из кольца $F[x]$ называется *неприводимым над полем* F, если он нормированный (со старшим коэффициентом, равным 1) и не может быть представлен в виде произведения двух многочленов из $F[x]$ меньших степеней.

Справедлива следующая

Теорема 30. *Пусть $f(x)$ — многочлен степени m с коэффициентами из простого поля F_p и $(f(x))$ — идеал, порожденный многочленом $f(x)$ в кольце $F_p[x]$. Фактор-кольцо $F_p[x]/(f(x))$, состоящее из p^m элементов, является полем тогда и только тогда, когда многочлен $f(x)$ неприводим над F_p.*

Пример 3. Пусть $p = 2$, $F_2 = \{0, 1\}$. Рассмотрим многочлен $f(x) = 1 + x^3 + x^4$. Несложно проверить, что он неприводим над F_2. Действительно, так как элементы 0 и 1 не являются корнями, то $f(x)$ не имеет линейных многочленов x и $x + 1$ в качестве делителей. Легко проверить, что единственный неприводимый над F_2 многочлен второй степени x^2+x+1 также не делит $f(x)$. Следовательно, многочлен $f(x)$ неприводим и по теореме 30 фактор-кольцо $F_2[x]/(f(x))$ является конечным полем с 2^4 элементами. Все 16 его элементов представимы как многочлены степени меньшей 4 с операциями сложения над F_2 и умножения по модулю $f(x)$. Например,

$$(x^3 + x + 1)(x^2 + 1) = x^5 + x^3 + x^2 + x^3 + x + 1 = x^5 + x^2 + x + 1 =$$

$$x(x^4 + x) + x^2 + x + 1 = (x^3 + 1 + x) + x^2 + x + 1 = x^3 + x^2 \,(\mathrm{mod}\, f(x)).$$

Теорема 30 приводит к ряду вопросов:

1. Существует ли неприводимый над F_p многочлен степени m для произвольных простого числа p и положительного целого числа m?

2. Сколько существует неприводимых над F_p многочленов степени m?

3. Существуют ли другие конечные поля?

Ответы на эти вопросы будут приведены в следующем разделе.

6.2. Строение конечных полей

Теорема 31. *Для любого простого числа p и любого положительного целого числа m существует единственное с точностью до изоморфизма поле порядка p^m.*

Конечное поле порядка p^m называется *полем Галуа* и обозначается $GF(p^m)$ в честь первого исследователя таких полей Эвариста Галуа.

Порядком произвольного элемента β некоторого конечного поля называется наименьшее целое положительное число k такое, что $\beta^k = 1$. Непосредственно из определения следует, что в конечном поле $GF(p^m)$ для элемента β порядка k все элементы $1, \beta, \beta^2, \ldots, \beta^{k-1}$ различны. Поэтому порядок каждого элемента поля $GF(p^m)$ конечен и не превышает числа $p^m - 1$. Элемент α поля $GF(p^m)$ называется *примитивным*, если его порядок равен $p^m - 1$. Многочлен, корнем которого является примитивный элемент, называется *примитивным многочленом*. Заметим, что не всякий неприводимый многочлен является примитивным.

Пусть (a, b) означает наибольший общий делитель чисел a и b. Имеют место следующие утверждения.

Лемма 6. *Пусть элементы β и γ коммутативной группы имеют порядки k и l соответственно, причем $(k, l) = 1$. Тогда порядок элемента $\beta\gamma$ равен kl.*

Лемма 7. *Пусть порядок элемента β коммутативной группы равен k. Тогда порядок элемента β^l равен $\frac{k}{(k,l)}$.*

Справедлива

Теорема 32. *Ненулевые элементы поля $GF(p^m)$ образуют циклическую группу порядка $p^m - 1$ относительно умножения.*

Доказательство. Пусть n — максимальный порядок ненулевых элементов поля $GF(p^m)$ и α — элемент порядка n, т. е. $\alpha^n = 1$. Отсюда следует, что $n \leq p^m - 1$, так как все n степеней элемента α различны между собой и не равны 0.

Пусть β — произвольный элемент поля порядка k, т. е. $\beta^k = 1$. Покажем, что тогда $k \mid n$ (т. е. k делит n). Предположим противное: $k \nmid n$. Рассмотрим элемент $\beta^{(k,n)}$. По лемме 7 его порядок равен $\frac{k}{(k,n)}$, т. е.

$$(\beta^{(k,n)})^{\frac{k}{(k,n)}} = 1.$$

Числа $\frac{k}{(k,n)}$ и n взаимно просты, следовательно, по лемме 6 порядок элемента $\alpha \cdot \beta^{(k,n)}$ равен $\frac{nk}{(n,k)}$, т. е.

$$(\alpha \cdot \beta^{(k,n)})^{\frac{nk}{(n,k)}} = 1.$$

Но число $\frac{nk}{(n,k)}$ строго больше n, что противоречит выбору n. Отсюда заключаем, что $k \mid n$. Другими словами, любой ненулевой элемент поля является корнем многочлена $x^n - 1 = 0$. Поскольку степень этого многочлена равна n, то он имеет не больше, чем n различных корней в данном поле $GF(p^m)$, поэтому $p^m - 1 \leq n$.

Таким образом, показано, что в поле $GF(p^m)$ найдется элемент α порядка $n = p^m - 1$. Все ненулевые элементы $GF(p^m)$ являются различными степенями α. ▲

Эта группа, состоящая из ненулевых элементов, называется *мультипликативной группой поля* $GF(p^m)$. Каждый образующий элемент этой группы, очевидно, является примитивным элементом поля. Справедливо и обратное. Несложно показать, что поле $GF(p^m)$ содержит в точности $\varphi(p^m - 1)$ примитивных элементов, где φ обозначает функцию Эйлера. В качестве следствия получаем следующий факт.

Теорема 33 (Теорема Ферма). *Каждый элемент поля* $GF(p^m)$ *удовлетворяет уравнению*

$$x^{p^m} - x = 0,$$

т. е. в поле $GF(p^m)$ *справедливо разложение*

$$x^{p^m} - x = \prod_{\beta \in GF(p^m)} (x - \beta).$$

Следующая теорема говорит о том, что конечных полей порядка, отличного от p^m, не существует.

Теорема 34. *Пусть* F *— конечное поле порядка* q *характеристики* p*. Тогда для некоторого положительного целого числа* m *справедливо равенство* $q = p^m$*.*

Доказательство. Покажем, что поле F может быть построено как линейное m-мерное пространство над $GF(p)$ для некоторого положительного целого числа m. Поскольку характеристика поля F равна p, можно показать, что поле F содержит $GF(p)$ в качестве наименьшего подполя. Пусть m — мощность базиса поля F над $GF(p)$, т. е. произвольный элемент $u \in F$ может быть представлен в виде

$$u = a_1 v^1 + \ldots + a_m v^m \tag{6.2}$$

для некоторых $a_i \in GF(p)$ и элементы $v^i \in GF(q)$, $i = 1, \ldots, m$ линейно независимы над $GF(p)$. Тогда верно $q \leq p^m$. С другой стороны, так как элементы $v^1, \ldots, v^m$ образуют базис поля F, все элементы u вида (6.2) различны при различных $a_1, \ldots, a_m$ и принадлежат полю F. Следовательно, $q \geq p^m$. Таким образом, доказано $q = p^m$. ▲

Теорема 35. *Для любых элементов a, b произвольного поля характеристики p и любого положительного целого числа s справедливо равенство*

$$(a-b)^{p^s} = a^{p^s} - b^{p^s}.$$

В заключение приведем сводку основных результатов, относящихся к полям Галуа и необходимых нам в дальнейшем:

1. *Число элементов произвольного поля Галуа равно степени простого числа.*

2. *Для любого простого числа p и любого целого числа $m \geq 0$ существует единственное с точностью до изоморфизма поле Галуа $GF(p^m)$.*

3. *Наименьшим подполем поля $GF(p^m)$ является поле $GF(p)$.*

4. *Поле $GF(p^k)$ является подполем поля $GF(p^m)$ тогда и только тогда, когда k делит m.*

5. *Любое поле $GF(p^m)$ содержит хотя бы один примитивный элемент.*

6. *Над каждым полем Галуа $GF(p)$ существует хотя бы один примитивный многочлен любой положительной степени.*

6.3. Примеры конечных полей

Пример 4. Построим конечное поле $GF(2^4)$ с помощью теоремы 30. Используем для этого неприводимый над $GF(2)$ многочлен $f(x) = x^4 + x + 1$.

Согласно теореме 30, поле $GF(2^4)$ состоит из всех многочленов степени меньшей 4:

0	x	x^2	x^3
1	$x+1$	x^2+1	x^3+1
		x^2+x	x^3+x
		x^2+x+1	x^3+x+1
			x^3+x^2
			x^3+x^2+1
			x^3+x^2+x
			x^3+x^2+x+1

Определим на множестве элементов операции умножения и взятия обратного элемента. Особенно удобно производить эти операции с помощью представления всех ненулевых элементов поля $GF(2^4)$ в виде степеней некоторого примитивного элемента α. Выберем его. Нетрудно проверить, что в качестве α

можно взять x. Действительно, все его степени по модулю $f(x)$ различны между собой:

$$\begin{aligned}
x^1 &= x,\\
x^2 &= x^2,\\
x^3 &= x^3,\\
x^4 &= x+1,\\
x^5 &= x^2+x,\\
x^6 &= x^3+x^2,\\
x^7 &= x^3+x+1,\\
x^8 &= x^2+1,\\
x^9 &= x^3+x,\\
x^{10} &= x^2+x+1,\\
x^{11} &= x^3+x^2+x,\\
x^{12} &= x^3+x^2+x+1,\\
x^{13} &= x^3+x^2+1,\\
x^{14} &= x^3+1,\\
x^{15} &= 1,
\end{aligned}$$

и, следовательно, порядок x равен 15.

Представим поле с помощью таблицы. Здесь число i для элемента $\gamma = \alpha^i$ называется *логарифмом* γ (по основанию выбранного примитивного элемента α). Логарифм элемента 0 полагают обычно равным $-\infty$.

Логарифм	**Степень прим. эл-та**	**Многочлен**	**Вектор**
$-\infty$	0	0	(0000)
0	1	1	(1000)
1	α	x	(0100)
2	α^2	x^2	(0010)
3	α^3	x^3	(0001)
4	α^4	$x+1$	(1100)
5	α^5	x^2+x	(0110)
6	α^6	x^3+x^2	(0011)
7	α^7	x^3+x+1	(1101)
8	α^8	x^2+1	(1010)
9	α^9	x^3+x	(0101)
10	α^{10}	x^2+x+1	(1110)
11	α^{11}	x^3+x^2+x	(0111)
12	α^{12}	x^3+x^2+x+1	(1111)
13	α^{13}	x^3+x^2+1	(1011)
14	α^{14}	x^3+1	(1001)

Сложение элементов поля является обычным сложением по модулю 2 и не зависит от выбора примитивного элемента в поле. Например:

$$\alpha^7 + \alpha^{11} = (x^3+x+1) + (x^3+x^2+x) = x^2+1 = \alpha^8.$$

Умножение ненулевых элементов поля, представленных в виде степеней примитивного элемента, проводится путем сложения показателей степеней по модулю 15.

Например:

$$(x^3 + x^2) \cdot (x^3 + x^2 + 1) = \alpha^6 \cdot \alpha^{13} = \alpha^{19 \,(\mathrm{mod}\, 15)} = \alpha^4 = x + 1.$$

Операция умножения, в отличие от сложения, зависит от выбора многочлена $f(x)$. Тем не менее, согласно теореме 31, какие бы неприводимые многочлены одинаковой степени не использовались нами для построения поля, все построенные поля будут изоморфны между собой.

Нахождение обратного элемента покажем на примере. Найдем обратный элемент для многочлена $x^3 + x^2 + 1 = \alpha^{13}$, т. е. такой ненулевой элемент α^k, что

$$\alpha^{13} \cdot \alpha^k = 1.$$

Для этого запишем

$$(x^3 + x^2 + 1)^{-1} = \alpha^{-13} = \alpha^{15-13} = \alpha^2 = x^2.$$

Нетрудно проверить, что решение найдено верно:

$$(x^3 + x^2 + 1) \cdot (x^2) = x^5 + x^4 + x^2 = ((x^2 + x) + (x + 1) + x^2) \bmod (f(x)) = 1 \bmod (f(x)).$$

Возможность находить для заданного ненулевого многочлена обратный обеспечивается неприводимостью многочлена $f(x)$ (аналогично и в общем виде для любого поля над $GF(q)$).

Несложно убедиться, что в построенном поле элементы α^2, α^4 также примитивные, а элементы α^3, α^5 примитивными не являются. Например, степени элемента α^3 порождают не все поле, а только элементы

$$\alpha^3, \alpha^6, \alpha^9, \alpha^{12}, \alpha^{15} = 1.$$

Заметим, что многочлен $f(x)$ в нашем случае — примитивный, поскольку x, его корень по построению, является примитивным элементом α.

Пример 5. Построим конечное поле $GF(2^4)$ с помощью неприводимого над $GF(2)$ многочлена $f(x) = x^4 + x^3 + x^2 + x + 1$. Найдем некоторый примитивный элемент поля $GF(2^4)$. Например, элемент x не является примитивным элементом, так как его порядок равен 5, что меньше $2^4 - 1 = 15$. Действительно,

$$x^5 = x \cdot x^4 \equiv x \cdot (x^3 + x^2 + x + 1) \bmod(f(x)) \equiv 1 \bmod(f(x)).$$

В качестве примитивного элемента можно взять $x + 1$, несложно убедиться что его порядок равен 15.

Представим поле с помощью таблицы.

Логарифм	Степень прим. эл-та	Многочлен	Вектор
$-\infty$	0	0	(0000)
0	1	1	(1000)
1	α	$x+1$	(1100)
2	α^2	x^2+1	(1010)
3	α^3	x^3+x^2+x+1	(1111)
4	α^4	x^3+x^2+x	(0111)
5	α^5	x^3+x^2+1	(1011)
6	α^6	x^3	(0001)
7	α^7	x^2+x+1	(1110)
8	α^8	x^3+1	(1001)
9	α^9	x^2	(0010)
10	α^{10}	x^3+x^2	(0011)
11	α^{11}	x^3+x+1	(1101)
12	α^{12}	x	(0100)
13	α^{13}	x^2+x	(0110)
14	α^{14}	x^3+x	(0101)

Таким образом, мы убедились, что с помощью различных многочленов можно найти различные (но эквивалентные) представления поля Галуа $GF(2^4)$.

6.4. Число неприводимых многочленов

Функция Мёбиуса определяется следующим образом:

$$\mu(d) = \begin{cases} 1, & \text{если } d = 1; \\ (-1)^r, & \text{если } d \text{ — произведение } r \text{ различных простых чисел;} \\ 0, & \text{в остальных случаях.} \end{cases}$$

Рассмотрим простое поле $GF(p)$. Обозначим через $I_p(m)$ число неприводимых над $GF(p)$ многочленов степени m.

Теорема 36. *Справедливо*

$$I_p(m) = \frac{1}{m} \sum_{d,\, d|m} \mu(d) p^{\frac{m}{d}}.$$

Упражнение 31. Установить изоморфизм полей, приведенных в примерах 4 и 5.

Упражнение 32. Доказать леммы 6 и 7 и теорему 35.

Упражнение 33. Найти все неприводимые над $GF(2)$ многочлены степени, не превышающей 3.

Упражнение 34. Построить поле Галуа $GF(2^2)$, используя неприводимый многочлен x^2+x+1. Найти таблицы сложения и умножения элементов поля.

Упражнение 35. Построить два представления поля Галуа $GF(2^3)$, используя один неприводимый многочлен $x^3 + x + 1$ и разные примитивные элементы. Указать изоморфизм этих представлений.

Упражнение 36. Доказать, что многочлен $M(x) = x^5 + x^2 + 1$ неприводим над $GF(2)$.

Упражнение 37. Построить поля Галуа:

а) $GF(2^3)$, используя неприводимый многочлен $x^3 + x^2 + 1$. Показать изоморфизм между построенным полем и полем из упражнения 35;

б) $GF(3^2)$, используя неприводимый многочлен $x^2 + x + 2$;

в) $GF(3^3)$, используя неприводимый многочлен $x^3 + 2x + 1$.

Глава 7

Циклические коды

В этой главе рассмотрим введение в теорию циклических кодов. Важность циклических кодов обусловлена тем, что:

1. Класс циклических кодов содержит много кодов с конструктивно задаваемым расстоянием; с кодовым расстоянием, близким к наилучшему, в особенности для кодов длины ≤ 100.

2. Для циклических кодов существуют сравнительно простые алгебраические методы декодирования и кодирования, поэтому именно циклические коды чаще всего используются для передачи информации в каналах связи с шумами.

7.1. Определение и свойства

Определение. Линейный код длины n называется *циклическим*, если для любого кодового слова $(x_1, x_2, \ldots, x_n)$ слово $(x_2, \ldots, x_n, x_1)$ также является кодовым.

Циклические коды обладают несложными процедурами кодирования и декодирования, имеют хорошие алгебраические свойства, их группы автоморфизмов содержат циклические подстановки. Все совершенные линейные коды, коды Рида-Маллера и другие коды, построенные еще до того, как Прейндж ввел понятие циклического кода и описал общие свойства таких объектов, являются кодами, которые после перестановки координат и несущественной модификации оказываются циклическими (см. [5]). Многие важнейшие блоковые коды, открытые позже, также оказались либо циклическими, либо тесно связанными с ними.

Перейдем к алгебраическому описанию циклических кодов с помощью многочленов с коэффициентами из конечного поля.

Обозначим через F простое конечное поле $GF(p)$, где p — простое число. Напомним, что $F[x]$ — кольцо всех многочленов от переменной x с коэффициентами из поля F. Оно ассоциативно, коммутативно и содержит единицу. В кольце $F[x]$ рассмотрим фактор-множество $F[x]/(x^n-1)$, состоящее из классов вычетов кольца $F[x]$ по модулю многочлена x^n-1. Множество $F[x]/(x^n-1)$ замкнуто относительно операций сложения $+$ и умножения $\cdot$ и, следовательно, является кольцом. Заметим, что множество $F[x]/(x^n-1)$ не является полем, так как многочлен x^n-1 приводим.

Все многочлены степени не больше $n-1$ попадают в различные классы вычетов и их можно выбрать в качестве представителей этих классов. Кольцо классов вычетов $F[x]/(x^n-1)$ изоморфно n-мерному векторному пространству над F:

$$c(x) = c_0 + c_1x + c_2x^2 + \ldots + c_{n-1}x^{n-1} \longleftrightarrow c = (c_0, c_1, c_2, \ldots, c_{n-1}).$$

В дальнейшем мы не будем различать векторы и многочлены степени меньше n, но из контекста всегда будет понятно, речь идет о многочленах или о векторах. Пусть дан многочлен

$$c(x) = \sum_{i=0}^{n-1} c_i x^i = c_0 + c_1x + \ldots + c_{n-1}x^{n-1}.$$

Обозначим через

$$\overleftarrow{c(x)} = \sum_{i=0}^{n-1} c_{n-i} x^i = c_{n-1} + c_{n-2}x + \ldots + c_0x^{n-1}$$

инверсию многочлена $c(x)$, записав его коэффициенты при степенях x в обратном порядке.

Определение. *Идеалом* I кольца $F[x]/(x^n-1)$ называется такое его линейное подпространство, что для любых многочленов $r(x) \in F[x]/(x^n-1)$ и $c(x) \in I$ многочлен $r(x) \cdot c(x)$ принадлежит I.

Теорема 37. *Подпространство кольца $F[x]/(x^n-1)$ является циклическим кодом тогда и только тогда, когда оно образует идеал.*

Доказательство. Существенным моментом в доказательстве этой теоремы является то, что в кольце $F[x]/(x^n-1)$ умножение многочлена на x соответствует циклическому сдвигу вектора в пространстве F^n. Действительно,

$$x \cdot c(x) = x \cdot (c_0 + c_1x + c_2x^2 + \ldots + c_{n-1}x^{n-1}) =$$

$$= c_{n-1} + c_0x + c_1x^2 + \ldots + c_{n-2}x^{n-1}.$$

Таким образом, вектор

$$(c_0, c_1, \ldots, c_{n-1})$$

переходит в вектор

$$(c_{n-1}, c_0, c_1, \ldots, c_{n-2}),$$

т. е. получаем циклический сдвиг в F^n.

Достаточность. Пусть C — циклический код. Тогда для любого кодового вектора c, его циклический сдвиг принадлежит C. Другими словами, для любого кодового многочлена $c(x)$ произведение $x \cdot c(x)$ принадлежит коду. Отсюда следует, в силу линейности циклического кода, что $f(x) \cdot c(x)$ является кодовым многочленом, где $f(x)$ — произвольный многочлен из $F[x]/(x^n-1)$. Следовательно, C — идеал в кольце $F[x]/(x^n-1)$.

Необходимость. Пусть подпространство D кольца $F[x]/(x^n-1)$ образует идеал. Рассмотрим многочлен $c(x) \in D$. По определению идеала, многочлен $f(x) \cdot c(x)$ принадлежит D для любого многочлена $f(x) \in F[x]/(x^n-1)$. Тогда $x \cdot c(x)$ принадлежит

D. Кроме того, сумма любых двух элементов из D также лежит в D и, следовательно, D — циклический код. ▲

Иногда пользуются следующим эквивалентным определением циклического кода.

Определение. *Циклическим кодом* длины n называется идеал кольца $F[x]/(x^n-1)$.

7.2. Порождающий многочлен

Выберем в циклическом коде C ненулевой многочлен наименьшей степени, обозначим его степень через r. Умножим многочлен на подходящий элемент поля F, чтобы он стал *нормированным* (или *приведенным*), т. е. чтобы коэффициент при старшей степени многочлена равнялся 1. В силу линейности кода C, полученный многочлен также принадлежит C. Обозначим его через $g(x)$.

Утверждение 12. *Циклический код содержит единственный ненулевой нормированный многочлен наименьшей степени.*

Доказательство. Пусть существуют два нормированных многочлена $f(x)$ и $g(x)$ наименьшей степени r. Тогда многочлен $f(x)-g(x)$, принадлежащий коду, имеет степень меньше r, что приводит к противоречию. ▲

Определение. Нормированный ненулевой многочлен наименьшей степени, принадлежащий циклическому коду, называется *порождающим* многочленом кода и обозначается через $g(x)$.

Теорема 38. *Циклический код состоит из всех многочленов вида*

$$f(x)\cdot g(x),$$

где $g(x)$ — порождающий многочлен кода степени r, степень $f(x)$ меньше $(n-r)$.

Доказательство. Согласно тому что циклический код образует идеал в кольце $F[x]/(x^n-1)$ (см. теорему 37), произведение любого многочлена $f(x)$ из $F[x]/(x^n-1)$ на $g(x)$ принадлежит коду. Тогда произведения $g(x)$ на все многочлены степени, меньшей чем $n-r$, принадлежат C.

Покажем, что любой кодовый многочлен представим в виде такого произведения. Пусть $c(x)$ — кодовый. Разделим его в кольце $F[x]/(x^n-1)$ с остатком на многочлен $g(x)$. Имеем

$$c(x)=q(x)g(x)+s(x),$$

где степень $s(x)$ меньше степени $g(x)$, а степень $q(x)$ меньше $n-r$. Многочлен

$$s(x)=c(x)-q(x)g(x)$$

является кодовым, так как слагаемые в правой части принадлежат коду C и он линеен. Но степень многочлена $s(x)$ меньше степени $g(x)$ — наименьшей степени ненулевого многочлена в C. Отсюда $s(x)=0$ и $c(x)=q(x)g(x)$, т. е. $c(x)$ имеет в кольце $F[x]/(x^n-1)$ искомое представление. ▲

Теорема 39. *Циклический код длины n с порождающим многочленом $g(x)$ существует тогда и только тогда, когда $g(x)$ делит x^n-1.*

Доказательство. Необходимость. Пусть дан циклический код C длины n с порождающим многочленом $g(x)$. Рассмотрим в кольце многочленов $F[x]$ деление многочлена $x^n - 1$ на $g(x)$ с остатком:

$$x^n - 1 = q(x)g(x) + r(x),$$

где степень $r(x)$ меньше степени $g(x)$. Переходя в кольцо $F[x]/(x^n - 1)$, получаем

$$0 = q(x)g(x) + r(x).$$

Отсюда $-r(x) = q(x)g(x)$. Из теоремы 37 следует, что многочлен $q(x)g(x)$ принадлежит коду и, следовательно, многочлен $-r(x)$ кодовый, причем $\deg r(x) < \deg g(x)$, что противоречит утверждению 12. Таким образом, $r(x) = 0$.

Достаточность. Пусть многочлен $g(x)$ делит $x^n - 1$. Покажем, что тогда существует циклический код длины n с порождающим многочленом $g(x)$. Поскольку степень $g(x)$ равна r, то размерность кода должна быть равна $n - r$.

Рассмотрим векторы, отвечающие многочленам

$$g(x), xg(x), x^2g(x), \ldots, x^{n-r-1}g(x),$$

где $r = \deg g(x)$. Будучи циклическими сдвигами вектора $g(x)$, они должны принадлежать циклическому коду с порождающим многочленом $g(x)$. Так как эти циклические сдвиги линейно независимы, то матрица

$$G = \begin{array}{c} \begin{matrix} x^0 & x^1 & \ldots & x^r & x^{r+1} & \ldots & \ldots & x^{n-1} \end{matrix} \\ \begin{pmatrix} g_0 & g_1 & \ldots & g_r & 0 & \ldots & \ldots & 0 \\ 0 & g_0 & g_1 & \ldots & g_r & 0 & \ldots & 0 \\ & & . & . & . & . & & \\ 0 & \ldots & 0 & g_0 & g_1 & \ldots & g_r & 0 \\ 0 & \ldots & \ldots & 0 & g_0 & g_1 & \ldots & g_r \end{pmatrix} \end{array} \sim \begin{pmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ \ldots \\ x^{n-r-1}g(x) \end{pmatrix},$$

где

$$g(x) = g_0 + g_1x + \ldots + g_rx^r,$$

имеет ранг $n - r$ (т. е. имеем искомое количество кодовых слов в базисе линейного кода). Следует отметить, что в матрицу G не может быть включен многочлен $x^{n-r+i}g(x)$, $i \geq 0$, поскольку по модулю многочлена $x^n - 1$ он равен линейной комбинации уже выбранных в G многочленов.

Покажем, что линейный код с порождающей матрицей G является идеалом, порожденным многочленом $g(x)$, в кольце $F[x]/(x^n - 1)$ и, кроме того, многочлен $g(x)$ имеет наименьшую ненулевую степень в этом идеале. Для этого представим вектор, отвечающий произвольному многочлену

$$f(x) = u(x)g(x)$$

из $F[x]/(x^n - 1)$ в виде линейной комбинации строк матрицы G, по теореме 38 этот многочлен кодовый. Действительно, имеем

$$f(x) = (u_0 + u_1x + \ldots + u_{n-r-1}x^{n-r-1})g(x) =$$

$$= u_0g(x) + u_1xg(x) + \ldots + u_{n-r-1}x^{n-r-1}g(x).$$

Остается показать, что не существует в этом коде ненулевого многочлена со степенью, меньшей степени многочлена $g(x)$. Если найдется такой кодовый многочлен $r(x)$, то он представим линейными комбинациями строк матрицы G, т. е. имеет вид $r(x) = q(x)g(x)$ для некоторого многочлена $q(x)$. Отсюда, переходя в кольцо $F[x]$, получаем

$$q(x)g(x) - r(x) = x^n - 1,$$

где по условию теоремы $x^n - 1$ делится на $g(x)$. Это возможно только в случае $r(x) = 0$. Таким образом, циклический код с порождающим многочленом $g(x)$ построен. ▲

Следствие 9. *Порождающая матрица циклического кода длины n с порождающим многочленом $g(x) = g_0 + g_1x + \ldots + g_rx^r$ имеет вид*

$$G = \begin{pmatrix} g_0 & g_1 & \ldots & g_r & 0 & \ldots & \ldots & 0 \\ 0 & g_0 & g_1 & \ldots & g_r & 0 & \ldots & 0 \\ & & . & . & . & . & & \\ 0 & \ldots & 0 & g_0 & g_1 & \ldots & g_r & 0 \\ 0 & \ldots & \ldots & 0 & g_0 & g_1 & \ldots & g_r \end{pmatrix}. \tag{7.1}$$

Матрица G имеет n столбцов и $n - r$ строк.

7.3. Кодирование циклических кодов

Определение. Код длины n размерности k называется *систематическим*, если после вычеркивания некоторых $(n - k)$ столбцов из его кодовой матрицы остаются в точности все различные векторы длины k.

Утверждение 13. *Любой циклический код эквивалентен систематическому коду.*

Для циклического кода существует простой способ нахождения порождающей матрицы в каноническом (или приведенно-ступенчатом) виде. При этом получается порождающая матрица того же кода, а не эквивалентного. Это существенно, поскольку перестановка координат может нарушить свойство цикличности.

Рассмотрим несколько общих принципов кодирования циклических кодов.

Первый систематический кодер. Пусть циклический код C длины n имеет порождающий многочлен $g(x)$ степени r. Тогда размерность k кода C равна $n-r$. Для построения порождающей матрицы G осуществим деление с остатком многочленов

$$x^{n-1}, x^{n-2}, \ldots, x^{n-k}$$

на $g(x)$. Имеем

$$x^{n-1} = g(x) \cdot a_1(x) + r_1(x),$$
$$x^{n-2} = g(x) \cdot a_2(x) + r_2(x),$$
$$\ldots$$
$$x^{n-k} = g(x) \cdot a_k(x) + r_k(x).$$

Поскольку каждый многочлен $x^{n-i} - r_i(x)$ для $i = 1, 2, \ldots, k$ делится на $g(x)$, то он является кодовым по теореме 38. Степень остатка $r_i(x)$ не превышает $r-1$. Пусть многочлен $-r_i(x)$ имеет вид

$$-r_i(x) = \lambda_{i,r-1}x^{r-1} + \lambda_{i,r-2}x^{r-2} + \ldots + \lambda_{i,0}.$$

Порождающая матрица

$$G = \begin{pmatrix} 1 & 0 & \ldots & 0 & \lambda_{1,r-1} & \lambda_{1,r-2} & \ldots & \lambda_{1,0} \\ 0 & 1 & \ldots & 0 & \lambda_{2,r-1} & \lambda_{2,r-2} & \ldots & \lambda_{2,0} \\ & & . & . & . & . & & \\ 0 & 0 & \ldots & 1 & \lambda_{n-r,r-1} & \lambda_{n-r,r-2} & \ldots & \lambda_{n-r,0} \end{pmatrix} \sim \begin{pmatrix} \overleftarrow{x^{n-1} - r_1 x} \\ \overleftarrow{x^{n-2} - r_2 x} \\ \ldots \\ \overleftarrow{x^{n-k} - r_k x} \end{pmatrix},$$

строки которой отвечают многочленам $\overleftarrow{x^{n-i} - r_i(x)}$ для $i = 1, 2, \ldots, k$, имеет приведенно-ступенчатый вид. Таким образом, для данного циклического кода найдено систематическое представление.

Второй систематический кодер. Можно рассмотреть систематическое кодирование для циклического кода длины n с порождающим многочленом $g(x)$ степени r без использования его порождающей матрицы. Пусть задана информационная последовательность

$$\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1}),$$

$k = n - r$. Рассмотрим многочлен

$$f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{k-1}x^{k-1}.$$

Домножая его на x^r, имеем многочлен

$$x^r f(x) = \alpha_0 x^r + \alpha_1 x^{r+1} + \ldots + \alpha_{k-1}x^{n-1}.$$

Разделив многочлен $x^r f(x)$ на $g(x)$ с остатком, получим

$$x^r f(x) = g(x)q(x) + r(x),$$

где степень $r(x)$ меньше r. Так как многочлен $x^r f(x) - r(x)$ делится на $g(x)$, то по теореме 38 он принадлежит коду. В силу того что $r(x)$ не меняет последние k компонент вектора, отвечающего многочлену $x^r f(x) - r(x)$, эти компоненты совпадают с компонентами информационной последовательности α. Таким образом, если многочлен $-r(x)$ имеет вид

$$-r(x) = \lambda_0 + \lambda_1 x + \ldots + \lambda_{r-1}x^{r-1},$$

то информационному блоку

$$\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1})$$

отвечает кодовое слово

$$x = (\lambda_0, \lambda_1, \ldots, \lambda_{r-1}, \alpha_0, \alpha_1, \ldots, \alpha_{k-1}).$$

Несистематический кодер. Рассмотрим простое несистематическое кодирование для циклического кода длины n с порождающим многочленом $g(x)$ степени r. Пусть информационному блоку

$$\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{k-1}),$$

$k = n - r$ отвечает многочлен

$$f(x) = \alpha_0 + \alpha_1 x + \ldots \alpha_{k-1} x^{k-1}.$$

Сопоставим ему кодовый многочлен $c(x)$ по правилу

$$c(x) = f(x) \cdot g(x).$$

Упражнение 38. Найти два систематических кодера для кода длины 7 с порождающим многочленом $g(x) = x^3 + x + 1$. С помощью второго систематического кодера закодировать вектор (1101).

7.4. Проверочный многочлен

По теореме 39 порождающий многочлен $g(x)$ делит многочлен $x^n - 1$.

Определение. Многочлен $h(x)$ такой, что

$$g(x) \cdot h(x) = x^n - 1$$

называется *проверочным многочленом* циклического кода. Его степень k равна $n - r$.

Утверждение 14. *Для произвольного кодового слова $c(x)$ циклического кода с проверочным многочленом $h(x)$ справедливо $c(x) \cdot h(x) = 0$ по модулю многочлена $x^n - 1$.*

Доказательство. Согласно теореме 38, любой кодовый многочлен $c(x)$ циклического кода имеет вид $q(x) \cdot g(x)$ и, следовательно, в фактор-кольце $F[x]/(x^n - 1)$ справедливо равенство

$$c(x) \cdot h(x) = q(x) \cdot g(x) \cdot h(x) = q(x) \cdot (x^n - 1) = 0.$$

▲

Теорема 40. *Проверочная матрица циклического кода длины n с проверочным многочленом $h(x) = h_0 + h_1 x + \ldots + h_k x^k$ имеет вид*

$$H = \begin{pmatrix} 0 & \ldots & \ldots & 0 & h_k & \ldots & h_1 & h_0 \\ 0 & \ldots & 0 & h_k & \ldots & h_1 & h_0 & 0 \\ & & . & . & . & . & & \\ 0 & h_k & \ldots & h_1 & h_0 & 0 & \ldots & 0 \\ h_k & \ldots & h_1 & h_0 & 0 & \ldots & \ldots & 0 \end{pmatrix}. \tag{7.2}$$

Доказательство. Согласно утверждению 14, для любого кодового слова $c(x)$ циклического кода выполняется

$$c(x) \cdot h(x) = 0.$$

Тогда

$$c(x) \cdot h(x) = \Big(\sum_{i=0}^{n-1} c_i x^i\Big) \cdot \Big(\sum_{j=0}^{n-r} h_j x^j\Big) = 0,$$

где $c(x) = \sum_{i=0}^{n-1} c_i x^i$. Коэффициент при x^j, $j = 0, 1, \ldots, n-1$, в этом произведении равен

$$\sum_{i=0}^{n-1} c_i h_{j-i} = 0, \tag{7.3}$$

где индексы берутся по модулю n. Эти равенства задают проверочные уравнения, которым должен удовлетворять код. Рассмотрим матрицу

$$H = \begin{pmatrix} \overset{x^0}{0} & \overset{\ldots}{\ldots} & \overset{\ldots}{\ldots} & \overset{x^{r-2}}{0} & \overset{x^{r-1}}{h_k} & \overset{\ldots}{\ldots} & \overset{x^{n-2}}{h_1} & \overset{x^{n-1}}{h_0} \\ 0 & \ldots & 0 & h_k & \ldots & h_1 & h_0 & 0 \\ & & \cdot & \cdot & \cdot & \cdot & & \\ 0 & h_k & \ldots & h_1 & h_0 & 0 & \ldots & 0 \\ h_k & \ldots & h_1 & h_0 & 0 & \ldots & \ldots & 0 \end{pmatrix} \sim \begin{pmatrix} \overleftarrow{h(x)} \\ \overleftarrow{xh(x)} \\ \overleftarrow{x^2 h(x)} \\ \ldots \\ \overleftarrow{x^{n-k-1} h(x)} \end{pmatrix}.$$

Из уравнений (7.3) следует, что если вектор

$$c = (c_0, c_1, c_2, \ldots, c_{n-1})$$

кодовый, то

$$H \cdot c^T = \mathbf{0}. \tag{7.4}$$

Так как $\deg h(x) = k = n - \deg g(x)$ есть размерность кода, и строки H линейно независимы, то условие (7.4) является также достаточным для того, чтобы вектор c принадлежал коду. Таким образом, H — проверочная матрица циклического кода с проверочным многочленом $h(x)$. ▲

Пример 6. Для двоичного циклического $[n, n-1, 2]$-кода с проверкой на четность справедливо

$$g(x) = x + 1,$$

$$h(x) = \frac{x^n - 1}{x + 1} = x^{n-1} + x^{n-2} + \ldots + x + 1.$$

Тогда матрицы

$$G = \begin{pmatrix} 1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & \ldots & 0 & 0 & 0 \\ & & \cdot & \cdot & \cdot & \cdot & & \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 1 & \ldots & 1 \end{pmatrix},$$

являются соответственно порождающей и проверочной матрицами кода.

7.5. Декодирование циклического кода

Пусть при передаче кодового вектора

$$a = (a_0, a_1, \ldots, a_{n-1})$$

произошли ошибки и на приемном конце получен вектор

$$c = (c_0, c_1, \ldots, c_{n-1}) = a + \varepsilon,$$

где

$$\varepsilon = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1})$$

— вектор ошибок. На языке многочленов имеем

$$c(x) = a(x) + \varepsilon(x).$$

Так как $a(x)$, согласно теореме 38, делится на $g(x)$, многочлены $c(x)$ и $\varepsilon(x)$ имеют при делении на $g(x)$ одинаковые остатки. Другими словами, вектор ошибок находится среди тех наборов, которым соответствуют многочлены, дающие при делении на $g(x)$ тот же остаток, что и многочлен $c(x)$, соответствующий принятому вектору. Обратно, любой вектор, обладающий указанным свойством, может оказаться вектором ошибок. Из стратегии декодирования по принципу максимума правдоподобия в качестве вектора ошибки ε принимают вектор наименьшего веса (как наиболее вероятный вектор ошибок). Отыскание вектора ε можно осуществить, например, перебором всевозможных векторов в порядке возрастания их весов вплоть до вектора, которому отвечает многочлен с нужным остатком. В зависимости от конкретной задачи на практике, как правило, проводят существенно более простые процедуры декодирования.

Существуют также алгоритмы декодирования, использующие циклический сдвиг и позволяющие за счет этого сократить объем таблицы синдромов.

Упражнение 39. Пусть для передачи информации использовался циклический код длины 7 с порождающим многочленом $g(x) = x^3 + x + 1$. Декодировать слово $y = (1011110)$.

7.6. Минимальный многочлен и его свойства

В этом разделе рассмотрим минимальные многочлены и их свойства, эти многочлены потребуются для построения циклических кодов, оценки их кодового расстояния (см. далее гл. 8).

Определение. *Минимальным многочленом* элемента $\beta \in GF(p^m)$ называется нормированный многочлен $M(x)$ с коэффициентами из $GF(p)$ наименьшей степени такой, что $M(\beta) = 0$.

Пусть $M(x)$ — минимальный многочлен над $GF(p)$ элемента β из $GF(p^m)$. Тогда имеют место следующие его свойства:

1. *Многочлен $M(x)$ неприводим.*

2. *Если $f(x)$ — такой многочлен, что $f(\beta) = 0$, то $M(x)$ делит $f(x)$.*

3. *Многочлен $M(x)$ делит $x^{p^m} - x$.*

4. *Степень многочлена $M(x)$ не превосходит m.*

5. *Степень многочлена $M(x)$ делит m.*

6. *Степень минимального многочлена $M(x)$ примитивного элемента равна m.*

7. *Если многочлен $M(x)$ является минимальным для элемента β, то он минимален и для элемента β^p.*

Множество целых чисел по модулю $p^m - 1$ следующим образом распадается на подмножества, называемые *циклотомическими классами по модулю* $p^m - 1$: класс, содержащий число s, имеет вид

$$C_s = \{s, ps, p^2s, p^3s, \ldots, p^{m_s-1}s\},$$

где m_s — наименьшее положительное целое число такое, что

$$p^{m_s} \cdot s \equiv s (\mathrm{mod}(p^m - 1)).$$

Пусть $M^{(i)}(x)$ — минимальный многочлен элемента α^i из $GF(p^m)$, где α — примитивный элемент поля. Из свойства 7 следует, что все ненулевые элементы поля $GF(p^m)$ вида α^k, показатели степеней которых лежат в одном циклотомическом классе, имеют один и тот же минимальный многочлен. Иначе говоря, над $GF(p)$ выполнено равенство

$$M^{(s)}(x) = M^{(ps)}(x) = M^{(p^2s)}(x) = M^{(p^3s)}(x) = \ldots = M^{(p^{m_s-1}s)}(x).$$

8. *Если число i принадлежит циклотомическому классу C_s, то в поле $GF(p^m)$ справедливо разложение*

$$M^{(i)}(x) = \prod_{j \in C_s} (x - \alpha^j).$$

Из теоремы Ферма (см. теорему 33) и свойства 8 следует

Теорема 41. *Над $GF(p)$ справедливо равенство*

$$x^{p^m-1} - 1 = \prod_s M^{(s)}(x),$$

где s пробегает все множество представителей циклотомических классов по модулю $p^m - 1$.

Теорема 42. *Многочлен $x^{p^m} - x$ равен произведению всех нормированных неприводимых над $GF(p)$ многочленов, степени которых делят m.*

Эти теоремы позволяют находить неприводимые и минимальные многочлены. Вернемся к примеру 4 поля $GF(2^4)$ из параграфа 6.3. Сначала выпишем все циклотомические классы по модулю 15:

$$C_0 = \{0\},$$
$$C_1 = \{1, 2, 4, 8\},$$
$$C_3 = \{3, 6, 12, 9\},$$
$$C_5 = \{5, 10\},$$
$$C_7 = \{7, 14, 13, 11\}.$$

Затем рассмотрим разложение многочлена $x^{16} - x$ на неприводимые многочлены:

$$x^{16} + x = x(x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)(x^4+x^3+x^2+x+1).$$

Задавая поле $GF(2^4)$ с помощью неприводимого многочлена x^4+x+1 и примитивного элемента $\alpha = x$, получаем

Элемент	Минимальный многочлен элемента
0	x
1	$M^{(0)}(x) = x + 1$
$\alpha, \alpha^2, \alpha^4, \alpha^8$	$M^{(1)}(x) = M^{(2)}(x) = M^{(4)}(x) = M^{(8)}(x) = x^4 + x + 1$
$\alpha^3, \alpha^6, \alpha^{12}, \alpha^9$	$M^{(3)}(x) = M^{(6)}(x) = M^{(12)}(x) = M^{(9)}(x) = x^4 + x^3 + x^2 + x + 1$
α^5, α^{10}	$M^{(5)}(x) = M^{(10)}(x) = x^2 + x + 1$
$\alpha^7, \alpha^{14}, \alpha^{13}, \alpha^{11}$	$M^{(7)}(x) = M^{(14)}(x) = M^{(13)}(x) = M^{(11)}(x) = x^4 + x^3 + 1$

По свойству 6 полей Галуа (см. разд.6.2.) многочлен, задающий поле, всегда может быть выбран примитивным. Для этого нужно выбрать минимальный многочлен примитивного элемента. Однако задача определения какой из неприводимых многочленов является примитивным, весьма трудна.

Упражнение 40. Доказать свойства 1–7 минимальных многочленов.

7.7. Число циклических кодов

Свяжем изложенную ранее теорию полей Галуа с циклическими кодами. В разд. 7.2. было показано, что циклический код длины n над полем $GF(p)$ существует для каждого многочлена $g(x)$, делящего многочлен $x^n - 1$, где $n = p^m - 1$ для некоторого $m > 0$. Согласно теореме 41, имеем

$$x^n - 1 = M^{(1)}(x)M^{(2)}(x)\dots M^{(\ell)}(x),$$

где ℓ — число циклотомических классов по модулю n, на которые разбиваются числа от 0 до $n - 1$. Произведение произвольного подмножества следующего множества многочленов

$$\{M^{(1)}(x), M^{(2)}(x), \dots, M^{(\ell)}(x)\}$$

дает порождающий многочлен $g(x)$ для некоторого циклического кода. Тогда, за исключением тривиальных случаев $g(x) = 1$ и $g(x) = x^n - 1$, получаем, что число нетривиальных циклических кодов длины n не превышает числа $2^\ell - 2$.

Какие из этих циклических кодов имеют наибольшее расстояние? Ответ на этот вопрос не прост.

Еще раз вернемся к примеру 4 из разд. 6.3. В разложении многочлена $x^{15}-1$ имеем пять сомножителей (см. разд. 7.6.). Следовательно, получаем $2^5 - 2 = 30$ нетривиальных циклических кодов длины 15. Рассмотрим, например, код с порождающим многочленом

$$g(x) = (x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1) = x^8 + x^4 + x^2 + x + 1.$$

Так как степень $g(x)$ равна 8, то размерность кода $k = n - 8 = 7$. Однако определить кодовое расстояние кода из простых соображений не удается.

Упражнение 41. Построить циклический код длины 7 с порождающим многочленом $g(x) = (x^3+x+1)$, найти его проверочный многочлен, определить параметры кода.

Упражнение 42. Что собой представляет циклический код длины 15 с порождающим многочленом $g(x)$? Найти его проверочный многочлен, определить параметры кода:

а) $g(x) = (x^4 + x + 1)$;

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

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

в случаях *б* и *в* построить коды.

Глава 8

Коды БЧХ

8.1. Нули кода

Пусть α — примитивный элемент поля $GF(p^m)$. Согласно теореме 41, над полем Галуа $GF(p)$ справедливо разложение

$$x^{p^m-1} - 1 = \prod_s M^{(s)}(x),$$

где s пробегает все множество представителей циклотомических классов по модулю $p^m - 1$. Пусть $g(x)$ — порождающий многочлен степени r некоторого циклического кода длины $n = p^m - 1$. Тогда, в силу теоремы 39 и приведенного разложения многочлена $x^{p^m-1} - 1$ на множители, имеем

$$g(x) = M^{(i_1)}(x) \cdot \ldots \cdot M^{(i_t)}(x)$$

для некоторых представителей циклотомических классов $i_1, \ldots, i_t$.

Определение. Корни порождающего многочлена $g(x)$ называются *нулями кода.*

Теорема 43. О нулях кода. *Пусть попарно различные элементы $\beta_1, \ldots, \beta_r$ из $GF(p^m)$ являются корнями порождающего многочлена $g(x)$ степени r циклического кода. Многочлен $f(x)$ с коэффициентами из $GF(p)$ принадлежит этому коду тогда и только тогда, когда*

$$f(\beta_1) = \ldots = f(\beta_r) = 0.$$

Доказательство.

Необходимость. Пусть $f(x)$ — кодовый многочлен. Тогда по теореме 38 имеем $f(x) = g(x)q(x)$, где $\deg q(x) < n-r$. Подставляя β_i вместо x для $i = 1, \ldots, r$, получаем

$$f(\beta_i) = g(\beta_i)q(\beta_i),$$

где $g(\beta_i) = 0$. Отсюда следует, что β_i — корень $f(x)$.

Достаточность. Пусть β_i — корень $f(x)$ для всех $i = 1, \ldots, r$ и справедливо

$$f(x) = g(x)q(x) + r(x),$$

где $\deg r(x) < r$. Подставляя β_i, получаем

$$f(\beta_i) = g(\beta_i)q(\beta_i) + r(\beta_i),$$

где $f(\beta_i) = 0$ и $g(\beta_i) = 0$. Отсюда следует $r(\beta_i) = 0$ для всех $i = 1, \ldots, r$. Поскольку $\deg r(x) < r$ и все β_i различны, заключаем, что $r(x) = 0$. Из теоремы 38 следует, что $f(x)$ — кодовый. ▲

8.2. Циклическое представление кода Хэмминга

Теорема 44. *Двоичный код Хэмминга является циклическим кодом с порождающим многочленом* $g(x) = M^{(1)}(x)$.

Доказательство. Проверочная матрица двоичного кода Хэмминга $\mathcal{H}^n$ длины $n = 2^m - 1$, по определению, состоит из всех ненулевых векторов-столбцов длины m. Пусть α — примитивный элемент поля Галуа $GF(2^m)$. Тогда α является порождающим элементом мультипликативной группы поля $GF(2^m)$ и все элементы

$$1, \alpha, \alpha^2, \ldots, \alpha^{2^m-2}$$

различны и могут быть представлены как ненулевые двоичные m-векторы. Таким образом, проверочную матрицу H двоичного кода Хэмминга с параметрами

$$[n = 2^m - 1, k = n - m, d = 3]$$

можно представить в виде

$$H = \begin{pmatrix} 1 & \alpha^1 & \alpha^2 & \ldots & \alpha^{2^m-2} \end{pmatrix},$$

где каждый элемент α^i должен быть заменен на соответствующий ему двоичный вектор-столбец длины m.

По определению, вектор $c = (c_0, c_1, \ldots, c_{n-1})$ принадлежит коду $\mathcal{H}^n$ тогда и только тогда, когда выполнено

$$H \cdot c^T = \mathbf{0}$$

т. е.

$$\sum_{i=0}^{n-1} c_i \cdot \alpha^i = 0.$$

Другими словами, элемент α является корнем многочлена $c(x)$:

$$c(\alpha) = 0.$$

Последнее равенство, согласно свойству 2 минимальных многочленов (см. разд. 7.6.), имеет место в том и только том случае, когда минимальный многочлен $M^{(1)}(x)$ элемента α делит многочлен $c(x)$. Таким образом, код $\mathcal{H}^n$ состоит из всех многочленов, кратных многочлену $M^{(1)}(x)$. ▲

Пример 7. Для $n = 7$ код Хэмминга $\mathcal{H}^7$ имеет проверочную матрицу

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

где α — корень примитивного минимального многочлена $M^{(1)}(x) = 1 + x + x^3$.

8.3. Определитель Вандермонда

Для доказательства теоремы 45 о границе Боуза (см. следующий разд.) нам потребуется матрица Вандермонда.

Определение. *Матрицей Вандермонда* называется матрица

$$A = \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-1} \\ & \cdot & \cdot & \cdot & \\ 1 & a_n & a_n^2 & \dots & a_n^{n-1} \end{pmatrix},$$

где $a_1, a_2, \dots, a_n$ — элементы некоторого конечного поля.

Лемма 8 Вандермонда. *Если все a_i, $i = 1, \dots, n$ различны, то* $\det A \neq 0$.

Доказательство. Индукцией по n докажем формулу

$$\det A = \prod_{j=1}^{n-1} \prod_{i=j+1}^{n} (a_i - a_j). \tag{8.1}$$

Для $n = 2$ имеем $\det A = \det \begin{pmatrix} 1 & a_1 \\ 1 & a_2 \end{pmatrix} = a_2 - a_1$. Формула (8.1) справедлива.

Предположим, что (8.1) выполнена для $n - 1$. Докажем ее для n. Запишем

$$\det A = \det \begin{pmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{n-2} & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{n-2} & a_2^{n-1} \\ & & \cdot & \cdot & \cdot & \\ 1 & a_n & a_n^2 & \dots & a_n^{n-2} & a_n^{n-1} \end{pmatrix}.$$

Домножим каждый i-й столбец матрицы (кроме последнего) на $-a_1$ и прибавим его к $(i+1)$-му столбцу. От этого, как известно, определитель матрицы не изменится. Получим

$$\det A = \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & (a_2 - a_1) & (a_2 - a_1)a_2 & \dots & (a_2 - a_1)a_2^{n-3} & (a_2 - a_1)a_2^{n-2} \\ & & \dots & \dots & \dots & \\ 1 & (a_n - a_1) & (a_n - a_1)a_n & \dots & (a_n - a_1)a_n^{n-3} & (a_n - a_1)a_n^{n-2} \end{pmatrix}.$$

Раскладывая определитель по первой строке, имеем

$$\det A = \det \begin{pmatrix} (a_2 - a_1) & (a_2 - a_1)a_2 & \dots & (a_2 - a_1)a_2^{n-3} & (a_2 - a_1)a_2^{n-2} \\ & \dots & \dots & \dots & \\ (a_n - a_1) & (a_n - a_1)a_n & \dots & (a_n - a_1)a_n^{n-3} & (a_n - a_1)a_n^{n-2} \end{pmatrix}.$$

Вынесем из каждой j-й строки множитель $(a_j - a_1)$, тогда

$$\det A = (a_2 - a_1)\dots(a_n - a_1) \cdot \det \begin{pmatrix} 1 & a_2 & \dots & a_2^{n-3} & a_2^{n-2} \\ & \dots & \dots & \dots & \\ 1 & a_n & \dots & a_n^{n-3} & a_n^{n-2} \end{pmatrix}.$$

Отсюда, по предположению индукции, заключаем

$$\det A = (a_2 - a_1)\dots(a_n - a_1) \cdot \prod_{j=2}^{n-1} \prod_{i=j+1}^{n} (a_i - a_j) = \prod_{j=1}^{n-1} \prod_{i=j+1}^{n} (a_i - a_j).$$

Таким образом, формула (8.1) доказана. Очевидно, что $\det A$ отличен от нуля тогда и только тогда, когда все a_i различны. ▲

8.4. Граница БЧХ

Следующая теорема называется границей Боуза-Чоудхури-Хоквингема (кратко границей БЧХ) или теоремой о конструктивном расстоянии циклического кода. Эта теорема позволяет оценивать кодовое расстояние снизу.

Теорема 45. Граница БЧХ. *Пусть $g(x)$ — порождающий многочлен циклического кода C длины n такой, что существуют целые числа $b \geq 0$ и $\delta > 1$ такие, что выполняется*

$$g(\alpha^b) = g(\alpha^{b+1}) = \dots = g(\alpha^{b+\delta-2}) = 0,$$

т. е. $\delta - 1$ подряд идущих степеней примитивного элемента α поля $GF(p^m)$ являются корнями $g(x)$. Тогда кодовое расстояние d не меньше δ.

Замечание. Число δ называется *конструктивным расстоянием* кода.

Доказательство. Рассмотрим произвольный кодовый вектор $c = (c_0, c_1, \dots, c_{n-1})$ и соответствующий ему многочлен $c(x)$. Поскольку элементы $\alpha^b, \alpha^{b+1}, \dots, \alpha^{b+\delta-2}$ являются нулями кода C, то по теореме 43 о нулях кода имеем

$$c(\alpha^b) = c(\alpha^{b+1}) = \dots = c(\alpha^{b+\delta-2}) = 0.$$

Отсюда получаем систему уравнений

$$\begin{cases} c_0 + c_1\alpha^b + c_2\alpha^{2b} + \dots + c_{n-1}\alpha^{(n-1)b} = 0 \\ c_0 + c_1\alpha^{b+1} + c_2\alpha^{2(b+1)} + \dots + c_{n-1}\alpha^{(n-1)(b+1)} = 0 \\ \quad \dots \\ c_0 + c_1\alpha^{b+\delta-2} + c_2\alpha^{2(b+\delta-2)} + \dots + c_{n-1}\alpha^{(n-1)(b+\delta-2)} = 0 \end{cases}$$

Иначе говоря, выполняется равенство

$$H \cdot c^T = \mathbf{0}, \tag{8.2}$$

где матрица

$$H = \begin{pmatrix} 1 & \alpha^b & \alpha^{2b} & \dots & \alpha^{(n-1)b} \\ 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \dots & \alpha^{(n-1)(b+1)} \\ & \cdot & \cdot & \cdot & \\ 1 & \alpha^{b+\delta-2} & \alpha^{2(b+\delta-2)} & \dots & \alpha^{(n-1)(b+\delta-2)} \end{pmatrix}$$

не обязательно является полной проверочной матрицей кода C, поскольку порождающий многочлен $g(x)$ возможно имеет и другие корни, отличные от $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2}$. Кроме того, строки этой матрицы могут оказаться линейно зависимыми. Другими словами, код C', заданный проверочной матрицей H, будет содержать в качестве подкода код C, но возможно не будет совпадать с ним. Достаточно оценить кодовое расстояние кода C', так как оно, очевидно, не будет больше кодового расстояния кода C. Поэтому если мы покажем, что любые $\delta - 1$ или менее столбцов матрицы H линейно независимы, то кодовое расстояние кода C' (а следовательно, и C) будет по меньшей мере δ согласно теореме 3 из разд. 1.2.

Предположим, что кодовый вектор c имеет вес w меньший δ, т. е. найдутся w линейно зависимых столбцов H и $w < \delta$. Пусть ненулевые координаты вектора c имеют номера $a_1, a_2, \ldots, a_w$. Построим квадратную матрицу H' из матрицы H следующим образом: выберем w первых строк из матрицы H и вычеркнем все столбцы с номерами, отличными от $a_1, a_2, \ldots, a_w$. Таким образом, матрица H' имеет вид

$$H' = \begin{pmatrix} \alpha^{a_1 b} & \alpha^{a_2 b} & \ldots & \alpha^{a_w b} \\ \alpha^{a_1 (b+1)} & \alpha^{a_2 (b+1)} & \ldots & \alpha^{a_w (b+1)} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha^{a_1 (b+w-1)} & \alpha^{a_2 (b+w-1)} & \ldots & \alpha^{a_w (b+w-1)} \end{pmatrix}.$$

Из равенства (8.2) следует, что

$$H' \cdot \begin{pmatrix} c_{a_1} \\ c_{a_2} \\ \vdots \\ c_{a_w} \end{pmatrix} = \mathbf{0}.$$

Последнее равенство возможно, только если матрица H' вырождена, т. е. $\det H' = 0$. Но с другой стороны, имеем

$$\det H' = \alpha^{a_1 b + a_2 b + \ldots + a_w b} \cdot \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha^{a_1} & \alpha^{a_2} & \ldots & \alpha^{a_w} \\ \cdot & \cdot & \cdot & \cdot \\ \alpha^{a_1 (w-1)} & \alpha^{a_2 (w-1)} & \ldots & \alpha^{a_w (w-1)} \end{pmatrix} \neq 0,$$

так как определитель в последнем равенстве есть определитель Вандермонда со всеми различными элементами, и согласно лемме 8 он не равен нулю.

Из полученного противоречия следует, что в линейном коде C не существует кодового слова веса меньше δ. Следовательно, кодовое расстояние C не меньше δ. ▲

8.5. Коды БЧХ

Определение. *Кодом БЧХ над полем $GF(p)$* длины $n = p^m - 1$ с конструктивным расстоянием $\delta > 1$ называется циклический код с порождающим многочленом наименьшей степени, нулями которого являются элементы

$$\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2},$$

где α — примитивный элемент поля $GF(p^m)$ и b — некоторое неотрицательное целое число.

Замечание. Для кодов БЧХ в литературе имеет место более общее определение. Нами рассматриваются так называемые *примитивные коды БЧХ*, т. е. коды, длина которых равна $n = p^m - 1$. Далее термин *примитивные* будет опущен. Коды БЧХ при $b = 1$ иногда называют *кодами БЧХ в узком смысле.*

Справедливо следующее (эквивалентное) определение кодов БЧХ.

Определение. *Кодом БЧХ над полем* $GF(p)$ длины $n = p^m - 1$ с конструктивным расстоянием $\delta > 1$ называется циклический код с порождающим многочленом

$$g(x) = \text{НОК}\{M^{(b)}(x), M^{(b+1)}(x), \ldots, M^{(b+\delta-2)}(x)\},$$

где b — некоторое неотрицательное целое число.

Теорема 46 *Код БЧХ над* $GF(p)$ *длины* $n = p^m - 1$ *с порождающим многочленом*

$$g(x) = \text{НОК}\{M^{(b)}(x), M^{(b+1)}(x), \ldots, M^{(b+\delta-2)}(x)\}$$

для некоторого целого $b \geq 0$ *имеет параметры*

$$[n = p^m - 1,\ k \geq n - (\delta - 1)m,\ d \geq \delta].$$

Доказательство. Согласно теореме 45 о границе БЧХ, кодовое расстояние кода не меньше δ. Из вида порождающего многочлена кода следует, что проверочная матрица кода БЧХ равна

$$H = \begin{pmatrix} 1 & \alpha^{b} & \alpha^{2b} & \ldots & \alpha^{(n-1)b} \\ 1 & \alpha^{b+1} & \alpha^{2(b+1)} & \ldots & \alpha^{(n-1)(b+1)} \\ & . & . & . & \\ 1 & \alpha^{b+\delta-2} & \alpha^{2(b+\delta-2)} & \ldots & \alpha^{(n-1)(b+\delta-2)} \end{pmatrix},$$

где каждый элемент должен быть заменен на соответствующий столбец из m элементов поля $GF(p)$. Строки этой матрицы задают проверочные соотношения кода. Общее число строк равно $(\delta - 1)m$, но они могут оказаться линейно зависимыми, поэтому для числа проверок выполняется

$$r \leq (\delta - 1)m$$

и, следовательно, для размерности кода имеем

$$k \geq n - (\delta - 1)m,$$

что завершает доказательство теоремы. ▲

8.6. Двоичные коды БЧХ

Отдельно рассмотрим случай двоичных кодов БЧХ в узком смысле, т. е. $b = 1$.

Теорема 47. *Двоичный код БЧХ длины $n = 2^m - 1$ с порождающим многочленом*

$$g(x) = \text{НОК}\{M^{(1)}(x), M^{(2)}(x), \ldots, M^{(2t-1)}(x)\}$$

имеет параметры

$$[n = 2^m - 1,\ k \geq n - tm,\ d \geq 2t + 1].$$

Доказательство. При $p = 2$ в силу свойства 7 минимальных многочленов (см. разд. 7.6.) имеем

$$M^{(2i)}(x) = M^{(i)}(x).$$

Отсюда следует, что степень $g(x)$ может быть понижена. А именно, определяя $g(x)$, можно не рассматривать минимальные многочлены для степеней примитивного элемента с четными показателями.

Конструктивное расстояние δ данного в теореме кода БЧХ равно $2t$. В силу отмеченного свойства минимальных многочленов, коды с конструктивными расстояниями $2t$ и $2t + 1$ совпадают, для каждого из них порождающий многочлен имеет вид

$$g(x) = \text{НОК}\{M^{(1)}(x), M^{(3)}(x), \ldots, M^{(2t-1)}(x)\}.$$

Таким образом, $\deg g(x) \leq tm$. Для размерности кода выполняется соответственно

$$k \geq n - tm.$$

Проверочная матрица кода равна

$$H = \begin{pmatrix} 1 & \alpha & \alpha^2 & \ldots & \alpha^{n-1} \\ 1 & \alpha^3 & \alpha^6 & \ldots & \alpha^{(n-1)3} \\ & \cdot & \cdot & \cdot & \\ 1 & \alpha^{2t-1} & \alpha^{2(2t-1)} & \ldots & \alpha^{(n-1)(2t-1)} \end{pmatrix},$$

где каждый элемент должен быть заменен соответствующим двоичным вектором-столбцом длины m. Второй столбец содержит степени элемента α:

$$\alpha, \alpha^3, \ldots, \alpha^{2t-1},$$

показатели которых лежат в различных циклотомических классах по модулю $2^m - 1$. ▲

В разд. 8.2. показано, что двоичный код Хэмминга имеет циклическое представление. Этот факт также непосредственно следует из доказанной теоремы.

Следствие 10. *Код БЧХ, исправляющий одну ошибку, имеет параметры*

$$[n = 2^m - 1,\ k = n - m,\ d = 3],$$

порождающий многочлен $g(x) = M^{(1)}(x)$ и является кодом Хэмминга.

Следствие 11. *Код БЧХ, исправляющий две ошибки, имеет параметры*

$$[n = 2^m - 1,\ k = n - 2m,\ d \geq 5],$$

где $m \geq 3$ нечетно и порождающий многочлен $g(x) = M^{(1)}(x) \cdot M^{(3)}(x)$.

Доказательство. Кодовое расстояние непосредственно следует из теоремы 47. Докажем, что

$$g(x) = M^{(1)}(x) \cdot M^{(3)}(x)$$

и $k = n - 2m$. По теореме 47 имеем

$$g(x) = \text{НОК}\{M^{(1)}(x), M^{(2)}(x), M^{(3)}(x), M^{(4)}(x)\} = \text{НОК}\{M^{(1)}(x), M^{(3)}(x)\}.$$

Далее нам потребуются некоторые свойства минимальных многочленов из разд. 7.6. Поскольку α — примитивный элемент поля $GF(2^m)$, то по свойству 6 минимальных многочленов степень $M^{(1)}(x)$ равна m. Так как m нечетно, имеем $(3, 2^m - 1) = 1$. По лемме 7 из разд. 6.2. получаем, что порядок элемента α^3 равен $2^m - 1$. Таким образом, многочлен $M^{(3)}(x)$ примитивен и по свойству 6 его степень также равна m. Минимальные многочлены $M^{(1)}(x)$ и $M^{(3)}(x)$ неприводимы согласно свойству 1 минимальных многочленов, поэтому

$$g(x) = M^{(1)}(x) \cdot M^{(3)}(x),$$

и размерность кода k в точности равна $n - 2m$. ▲

Коды БЧХ (более точно: примитивные коды БЧХ) асимптотически плохие, а именно справедливо следующее утверждение.

Теорема 48 *Для любой бесконечной последовательности $[n, k, d]$-кодов БЧХ над $GF(q)$ скорость кода k/n и отношение d/n стремятся к нулю с ростом n.*

8.7. Коды Рида-Соломона

8.7.1. Определение и свойства

Коды Рида-Соломона — это коды БЧХ над $GF(q)$, $q = p^m$, длина n которых равна $q - 1$, $q \neq 2$, а порождающий многочлен имеет вид

$$g(x) = (x - \alpha^b) \cdot (x - \alpha^{b+1}) \cdot \ldots \cdot (x - \alpha^{b+\delta-2}) \tag{8.3}$$

для некоторых $b \geq 0$, $\delta > 1$. Иногда удобно рассматривать b равным единице. Такие коды представляют собой значительный практический и теоретический интерес и обладают целым рядом хороших свойств. В следующей теореме, например, покажем, что для кода Рида-Соломона можно точно вычислить как кодовое расстояние, так и мощность.

Теорема 49. *Код Рида-Соломона длины n имеет мощность $q^{n-\delta+1}$ и кодовое расстояние $d = n - k + 1 = \delta$.*

Доказательство. Поскольку порождающий многочлен кода Рида-Соломона длины $q-1$ имеет вид (8.3) для некоторых $b \geq 0$, $\delta > 1$, код Рида-Соломона — это БЧХ-код с конструктивным расстоянием δ длины $n = q-1$, $q \neq 2$. Размерность этого циклического кода равна $k = n - \deg g(x) = n - \delta + 1$. Согласно границе Синглтона (см. теорему 4 разд. 1.2.), если B — (n, k, d)-код, то $n - k \geq d - 1$. Иными словами, $d \leq n - k + 1$. Отсюда для кода Рида-Соломона имеем $d = n - k + 1$, а следовательно, этот код является MDS-кодом. ▲

Достоинства кодов Рида-Соломона:

1. Коды Рида-Соломона удобно использовать, когда требуется код, длина которого меньше, чем размер поля, так как, являясь MDS-кодами, они имеют наибольшее возможное минимальное расстояние.

2. Они используются для получения двоичных кодов с очень большими минимальными расстояниями.

3. Коды Рида-Соломона используются при построении каскадных кодов с хорошими параметрами.

4. Они используются при построении кодов, исправляющих пакеты ошибок (см. подробнее разд. 8.7.2.).

Пример 8. Рассмотрим код Рида-Соломона над $GF(5)$ длины 4 с конструктивным расстоянием 3. В качестве примитивного элемента поля $GF(5)$ возьмем, например, $\alpha = 2$, тогда $g(x) = (x - \alpha)(x - \alpha^2) = (x - 2)(x - 4) = x^2 + 4x + 3$. Код Рида-Соломона имеет $5^2 = 25$ кодовых слов длины 4, среди них например,

$$(3410), (2140), (0341), (3201) \ldots .$$

Упражнение 43. Найти все 25 кодовых слов этого кода.

Добавление к коду общей проверки на четность не всегда увеличивает его минимальное расстояние, если коды недвоичные и есть слова нечетного веса. Например, добавление общей проверки на четность к коду над $GF(3)$ с кодовой матрицей

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

не увеличивает кодовое расстояние. Однако для кодов Рида-Соломона с порождающим многочленом $g(x) = (x - \alpha) \cdot (x - \alpha^2) \cdot \ldots \cdot (x - \alpha^{\delta-1})$ кодовое расстояние всегда увеличивается на 1.

Теорема 50. *Пусть P — $[n = p^m - 1, k, d]$-код Рида-Соломона с порождающим многочленом*

$$g(x) = (x - \alpha) \cdot (x - \alpha^2) \cdot \ldots \cdot (x - \alpha^{\delta-1}).$$

Тогда расширение каждого кодового слова $c = (c_0, c_1, \ldots, c_{n-1})$ посредством добавления общей проверки на четность над $GF(p)$

$$c_n = -\sum_{i=0}^{n-1} c_i$$

приводит к коду с параметрами $[n + 1, k, d + 1]$.

Доказательство. Пусть $w(c) = d$. Кодовому слову c соответствует многочлен $c(x)$, и минимальный вес c увеличивается до $d+1$, если

$$\sum_{i=0}^{n-1} c_i = -c_n \neq 0.$$

Но $c(x) = c_0 + c_1 x + \ldots + c_{n-1}x^{n-1}$ и, следовательно,

$$c(1) = \sum_{i=0}^{n-1} c_i = -c_n.$$

Покажем, что $c(1) \neq 0$. По теореме 38 (см. разд. 7.2.) имеем $c(x) = a(x)g(x)$, где $g(x)$ — порождающий многочлен кода. По определению $g(x)$, поскольку α^0 не является его корнем заключаем, что $g(1) \neq 0$. Для $a(x)$ имеем $a(1) \neq 0$, так как в противном случае многочлен $c(x)$ делился бы на многочлен $(x-1)$ и, согласно границе БЧХ (см. теорему 45), кодовое слово c имело бы вес не менее чем $d+1$. Следовательно, $c(1) = a(1)g(1) \neq 0$, то есть кодовое расстояние полученного кода увеличивается на 1. ▲

Упражнение 44. Построить код Рида-Соломона с параметрами $[3, 2, 2]$ над полем Галуа

$$GF(4) = \{0, 1, \alpha, \alpha^2\},$$

где $\alpha^2 + \alpha + 1 = 0$ и рассмотреть расширенный код с параметрами $[4, 2, 3]$.

8.7.2. Использование кодов Рида-Соломона для получения двоичных кодов

Элементы поля $GF(p^m)$ могут быть представлены m-векторами над $GF(p)$ (вспомним пример поля $GF(2^4)$, где элементы поля были представлены двоичными векторами длины 4). Рассмотрим q-значный код Рида-Соломона с параметрами

$$[n = q-1, k = n-\delta+1, d = \delta]_q,$$

где $q = p^m$. Произведя замену на p-значные векторы, получим p-значный код с параметрами

$$[n' = nm, k' = km, d' \geq d].$$

Если $q = 2^m$, то полученные двоичные коды имеют часто большое минимальное расстояние. Далее покажем это.

Пусть $v_1, \ldots, v_m$ — базис векторного пространства $GF(2^m)$ над $GF(2)$, т. е. базис в m-мерном единичном кубе E^m. Тогда любой элемент β, принадлежащий $GF(2^m)$, представим в виде

$$\beta = \sum_{i=1}^{m} b_i v_i = b_1 v_1 + \ldots + b_m v_m,$$

где b_i принадлежит $GF(2)$. Рассмотрим отображение $\beta \to (b_1, \ldots, b_m)$. Это отображение переводит линейные коды над $GF(2^m)$ в линейные над $GF(2)$, т. е. сохраняет линейность, но необязательно при этом сохраняет цикличность!

Построим из кода Рида-Соломона двоичный код с большим кодовым расстоянием. Пусть вектор $c = (c_0, \ldots, c_{n-1})$ принадлежит $[n, k, d]$-коду Рида-Соломона над $GF(2^m)$. Заменим элементы c_i соответствующими двоичными m-векторами и к каждому такому m-вектору добавим общую проверку на четность. Получим двоичный код с параметрами

$$[n' = (m+1)(2^m - 1), k' = mk, d' \geq 2d = 2(2^m - k)].$$

Аналогичное преобразование можно применить к расширенному коду Рида-Соломона, что приводит к двоичному коду с параметрами

$$[(m+1)2^m, mk, d' \geq 2(2^m - k + 1)].$$

Полученный код является каскадным кодом.

Каскадные коды имеют широкое применение на практике, например, используются при записи информации на компакт дисках. Повреждения в компакт дисках могут вызвать длинные последовательности ошибок. Ошибки в цифровой записи и воспроизведении бывают двух типов:

1) случайные (в несколько бит, обычно разбросаны по диску, их можно легко исправить);

2) ошибки типа "потеря пакета".

Определение. *Пакетом* длины s называется вектор, все ненулевые элементы которого расположены среди s подряд идущих компонент, из которых первая и последняя являются ненулевыми.

Иначе говоря, ошибка составляет большое число последовательно расположенных бит. Например, в компакт дисках такая ошибка может быть вызвана физическим повреждением диска или крупным дефектом ленты. Эффект, вызванный таким дефектом, может быть существенно уменьшен, если биты, составляющие посылку, располагать не последовательно, а дискретно через некоторые интервалы (кадры). Тогда потерю сравнительно большого пакета ошибок можно рассматривать как случайную. Такой метод передачи данных использует код Рида-Соломона, а точнее двоичный код, полученный из кода Рида-Соломона. Иногда рассматриваются более сложные преобразования кода Рида-Соломона в двоичный с помощью процедуры перемежения, с применением двух кодеров, например, при использовании функции четности для любого двоичного набора длины m либо каскадирования. Однако двоичный код, полученный из кода Рида-Соломона, все же довольно плохо исправляет случайные ошибки. В этом случае может быть использован код Юстесена.

8.8. Коды Юстесена

Рассмотрим код Рида-Соломона над полем $GF(2^m)$ с параметрами

$$[n = 2^m - 1, k, d = n - k + 1]_{2^m}.$$

Пусть α — примитивный элемент поля $GF(2^m)$. Рассмотрим произвольный вектор

$$(a_0, a_1, a_2, \ldots, a_{n-1})$$

кода Рида-Соломона и составим с помощью него вектор длины $2n$ над $GF(2^m)$ следующего вида

$$a = (a_0, a_0, a_1, \alpha a_1, a_2, \alpha^2 a_2, \ \ldots, \ a_{n-1}, \alpha^{n-1} a_{n-1}).$$

Заменяя в этом векторе каждый элемент a_i из $GF(2^m)$ отвечающим ему двоичным вектором длины m, получаем двоичный вектор длины $2mn$. Полученное множество двоичных слов образует *код Юстесена* с параметрами

$$[N = 2mn, K = mk, D],$$

который по построению является линейным кодом. Кодовое расстояние D определяется из следующего предложения.

Утверждение 15. *Минимальное расстояние D кода Юстесена удовлетворяет неравенству*

$$D \geq \sum_{i=1}^{w} i \cdot \left(\begin{array}{c} 2m \\ i \end{array} \right),$$

где для кодового расстояния d исходного кода Рида-Соломона справедливо:

$$\sum_{i=1}^{w} \left(\begin{array}{c} 2m \\ i \end{array} \right) \leq d \leq \sum_{i=1}^{w+1} \left(\begin{array}{c} 2m \\ i \end{array} \right).$$

Доказательство. Любое ненулевое кодовое слово кода Юстесена содержит по крайней мере d различных двоичных векторов длины $2m$, где $u \neq 0$ и $v \neq 0$. Очевидно, что количество различных векторов длины $2m$ малого веса невелико. Следовательно, вес любого ненулевого кодового слова кода Юстесена должен быть большим. Минимальный вес ненулевого слова кода Юстесена не меньше суммы весов d ненулевых двоичных векторов длины $2m$ минимально возможного веса, а именно:

$$D \geq \sum_{i=1}^{w} i \cdot \left(\begin{array}{c} 2m \\ i \end{array} \right) + (w+1) \left(d - \sum_{i=1}^{w} \left(\begin{array}{c} 2m \\ i \end{array} \right) \right),$$

где в первом слагаемом под знаком суммы стоит количество единиц в наборе длины $2m$, а во втором — количество оставшихся столбцов в двоичном аналоге проверочной матрицы. Причем вес w определяется из соотношения

$$\sum_{i=1}^{w} \left(\begin{array}{c} 2m \\ i \end{array} \right) \leq d \leq \sum_{i=1}^{w+1} \left(\begin{array}{c} 2m \\ i \end{array} \right).$$

▲

Нетрудно видеть, что коды Юстесена являются каскадными кодами. Они также хороши тем, что являются асимптотически хорошими кодами.

Глава 9

Другие коды

9.1. Матрицы Адамара, коды Адамара

9.1.1. Матрицы Адамара

Определение. *Матрицей Адамара* порядка n называется $n \times n$ матрица H, элементами которой являются $+1$ и -1 такая, что

$$H \cdot H^T = nE_n,$$

где E_n — единичная матрица размера $n \times n$.

Это равенство эквивалентно тому, что любые две строки матрицы ортогональны, т. е. их скалярное произведение в поле действительных чисел равно 0, а скалярное произведение любой строки на саму себя равно n.

Матрица H носит название *матрицы Адамара*, поскольку ее детерминант *достигает* границы, принадлежащей Адамару. Справедлива

Теорема 51 (Адамар Ж., 1897). *Если $A = (a_{ij})$ — произвольная вещественная $(n \times n)$ матрица с элементами $-1 \leq a_{ij} \leq 1$, то $|\det A| \leq n^{n/2}$ (т. е. $\det^2 A \leq n^n$).*

Для матрицы Адамара справедливо $\det H \cdot H^T = (\det H)^2 = n^n$. Имеем

$$H^{-1} \cdot H \cdot H^T = nH^{-1} \quad \Rightarrow \quad H^T = nH^{-1} \quad \text{и} \quad H^T \cdot H = nH^{-1} \cdot H = nE_n,$$

т. е. $H^T \cdot H = nE_n$. Отсюда следует также, что столбцы матрицы H обладают теми же свойствами, что и строки матрицы H.

Очевидно, что перестановка строк или столбцов матрицы H, а также умножение строк или столбцов матрицы H на -1, переводят H в другую матрицу Адамара H'. Такие матрицы Адамара будем называть эквивалентными. Это равносильно тому, что

$$H' = P \cdot H \cdot Q,$$

где P и Q — мономиальные матрицы перестановки длины n с элементами $+1$ и -1. Напомним, что матрица называется *мономиальной*, если в каждой строке и каждом

столбце имеется точно один ненулевой элемент. В нашем случае для P и Q эти ненулевые элементы равны $+1$ или -1. Матрица P осуществляет перестановку и меняет знаки у строк H, а матрица Q — у столбцов.

Для данной матрицы Адамара H всегда можно найти эквивалентную ей матрицу Адамара, первая строка и первый столбец которой целиком состоят из $+1$. Такая матрица Адамара называется *нормализованной*.

Пример 9. При $n = 1$ имеем $H(1) = (1)$,

$$\text{при } n = 2, \qquad H(2) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

$$\text{при } n = 4, \qquad H(4) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Перестановка строк, кроме первой, и столбцов, кроме первого, не нарушают нормализованности матрицы Адамара. Но, вообще говоря, могут существовать эквивалентные нормализованные матрицы, которые *не получаются* одна из другой простой перестановкой строк и столбцов.

Упражнение 45. Показать, что матрицы Адамара порядков 2 и 4 существуют и единственны с точностью до эквивалентности.

Матрицы Адамара порядков 8 и 12 также единственны с точностью до эквивалентности. Существует пять неэквивалентных матриц Адамара порядка 16 и три неэквивалентных матрицы порядка 20. Для $n = 24$ число неэквивалентных матриц Адамара равно 60, для $n = 28$ таких матриц 487.

Теорема 52. *Если существует матрица Адамара порядка $n > 2$, то n кратно 4.*

Доказательство. Без ограничения общности предположим, что H представлена в нормализованном виде и пусть $n > 2$. Найдется матрица, эквивалентная матрице H, первые три строки которой имеют вид

$$\begin{array}{cccc|cccc|cccc|cccc} 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 & -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 & 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 \\ \underbrace{\qquad\qquad}_{i} & & & & \underbrace{\qquad\qquad}_{j} & & & & \underbrace{\qquad\qquad}_{k} & & & & \underbrace{\qquad\qquad}_{l} & & & \end{array}$$

Тогда из ортогональности строк имеем следующую систему уравнений

$$\begin{cases} i + j + k + l = n, \\ i + j - k - l = 0 \quad \text{(умножая 1-ю и 2-ю строки)}, \\ i - j - k + l = 0 \quad \text{(умножая 2-ю и 3-ю строки)}, \\ i - j + k - l = 0 \quad \text{(умножая 1-ю и 3-ю строки)}. \end{cases}$$

Решая эту систему, получаем $i = j = k = l = n/4$, откуда следует, что 4 делит n. ▲

Гипотеза. Матрицы Адамара существуют для всех $n \equiv 0 (\mathrm{mod}\ 4)$.

Существует большое количество методов построения матриц Адамара. Наименьший порядок, для которого матрица Адамара не построена, равен $n = 428 = 4 \cdot 107$ (2002 г.).

Приведем неполный спектр значений n, для которых построены матрицы Адамара:

1) $n = 2^r$;

2) $n = p^r + 1 \equiv 0 (\mathrm{mod}\ 4)$ для простого p;

3) $n = h(p^r + 1)$, где $h \geq 2$ — порядок матрицы Адамара;

4) $n = h(h-1)$, где h — произведение чисел вида 1 и 2;

5) $n = h(h+3)$, где h и $h+4$ — произведения чисел вида 1 и 2;

6) $n = h_1 h_2 (p^r + 1)p^r$, где $h_1, h_2 > 1$ — порядки матриц Адамара.

Рассмотрим два особенно важных в теории кодирования метода построения матриц Адамара.

9.1.2. Матрица Сильвестра

Прямым, или *кронекеровым*, произведением двух матриц

$$A = (a_{ij}) \text{ порядка } m \times m \text{ и } B = (b_{ij}) \text{ порядка } n \times n$$

называется матрица $A \times B$ порядка $(mn \times mn)$:

$$A \times B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{bmatrix}.$$

Упражнение 46. Доказать свойства:

1) $A \times (B_1 + B_2) = A \times B_1 + A \times B_2$;

2) $(A \times B)(C \times D) = AC \times BD$.

Теорема 53. *Если существуют матрицы Адамара порядков m и n, то их прямое произведение есть матрица Адамара порядка $m \cdot n$.*

Доказательство. Пусть H_m и H_n — две матрицы Адамара порядков m и n соответственно. Тогда

$$(H_m \times H_n)(H_m \times H_n)^T = (H_m \times H_n)(H_m^T \times H_n^T) =$$

$$= H_m \cdot H_m^T \times H_n \cdot H_n^T = mE_m \times nE_n = mnE_{mn}. \qquad \blacktriangle$$

Отсюда легко вытекает следующий метод построения матриц Адамара.

Теорема 54. *Если H_n — матрица Адамара порядка n, то*

$$H_{2n} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times H_n = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$$

— матрица Адамара порядка $2n$.

Пример 10. Матрица Адамара порядка 4, полученная из матрицы Адамара $H(2)$:

$$H(4) = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array}\right).$$

Начиная с тривиальной матрицы $H(1) = (1)$, методом, описанным в теореме 54, получим последовательность матриц $H(1)$, $H(2)$, $H(4), \ldots, H(2^k) \ldots$ Такие матрицы называются *матрицами Сильвестра*. Они имеют порядки, равные степени двойки $2^k, k \geq 1$.

Упражнение 47. Построить матрицу Адамара H_8.

9.1.3. Матрица Адамара по типу Пэйли

Для описания конструкции Пэйли построения матриц Адамара, потребуются квадратичные вычеты и некоторые их свойства.

Определение. Пусть $p > 2$ — простое число. Числа $b \not\equiv 0 (\mathrm{mod}\ p)$ делятся на два класса, называемые *квадратичными вычетами* и *квадратичными невычетами* в зависимости от того, имеет сравнение

$$x^2 \equiv b \,(\mathrm{mod}\ p)$$

решение по модулю p или не имеет.

Пример 11. При $p = 7$:

$2^2 = 4 (\mathrm{mod}\ 7)$ — вычет;

$3^2 \equiv 2 (\mathrm{mod}\ 7)$ — вычет (число 2 представимо в виде 3^2, где $3 < 7$);

3 — невычет, так как сравнение $x^2 \equiv 3 (\mathrm{mod}\ 7)$ не имеет решения, в чем легко убедиться простым перебором;

$4^2 = 16 \equiv 2 (\mathrm{mod}\ 7)$;

$5^2 = 25 \equiv 4 (\mathrm{mod}\ 7)$;

$6^2 = 36 \equiv 1 (\mathrm{mod}\ 7)$.

Новых вычетов не получили, следовательно, 1, 2 и 4 — все вычеты по модулю 7.

Для того чтобы найти все вычеты по модулю p, нужно рассмотреть все числа $1, 2, \ldots, p-1$, но достаточно рассмотреть квадраты чисел от 1 до $(p-1)/2$, так как

$$(p-a)^2 \equiv a^2 (\mathrm{mod}\ p).$$

Свойства квадратичных вычетов

1. *Произведение двух квадратичных вычетов или невычетов является квадратичным вычетом, произведение квадратичного вычета на невычет является невычетом.*

2. Критерий Эйлера. *Пусть $p > 2$ — простое. Число a, взаимно простое с p, является квадратичным вычетом по модулю p тогда и только тогда, когда*

$$a^{(p-1)/2} \equiv 1 (\mathrm{mod}\ p)$$

и является невычетом по модулю p тогда и только тогда, когда

$$a^{(p-1)/2} \equiv -1 (\mathrm{mod}\ p).$$

3. *Если $p = 4k + 1$, то число -1 является квадратичным вычетом, если $p = 4k - 1$, то число -1 является невычетом.*

Для доказательства этого свойства полезен следующий факт.

4. *Если α — примитивный элемент поля Галуа $GF(p)$, где p — простое, то, нетрудно показать, что квадратичные вычеты задаются четными степенями элемента α.*

Определение. Пусть p — простое нечетное число. Функция $\chi(i)$, называемая *символом Лежандра*, определяется на множестве целых чисел следующим образом:

$\chi(i) = 0$, если $i \equiv 0 (\mathrm{mod}\ p)$;

$\chi(i) = 1$, если остаток от деления i на p является квадратичным вычетом по модулю p;

$\chi(i) = -1$, если остаток от деления i на p является квадратичным невычетом по модулю p.

Теорема 55. *Для любого $c \not\equiv 0 (\mathrm{mod}\ p)$ справедливо*

$$\sum_{b=0}^{p-1} \chi(b)\chi(b + c) = -1.$$

Доказательство. Так как произведение двух квадратичных вычетов или невычетов есть вычет, а произведение вычета на невычет есть невычет, то

$$\chi(xy) = \chi(x)\chi(y) \quad \text{при} \quad 0 \le x, y \le p - 1.$$

Если $b = 0$, то $\chi(0) = 0$ и вклад в сумму слагаемого при $b = 0$

$$\chi(0)\chi(c) = 0.$$

Пусть $b \neq 0$ и пусть z такое число, что $zb \equiv b + c (\mathrm{mod}\ p)$. Если b пробегает множество $\{1, 2, \ldots, p-1\}$, то z пробегает множество $\{0, 2, 3, \ldots, p-1\}$ (действительно,

так как $c \not\equiv 0 (\mathrm{mod}\ p)$, имеем $z \neq 1$). Разным числам b отвечают разные z. Пусть это не так, т. е.

$$\begin{array}{l} zb \equiv b + c\,(\mathrm{mod}\ p) \\ zb' \equiv b' + c\,(\mathrm{mod}\ p) \end{array}, \qquad b \neq b' \ \text{ и } \ b, b' \leq p-1.$$

Тогда $z(b-b') \equiv b-b'\,(\mathrm{mod}\ p)$, или $(z-1)(b-b') \equiv 0 (\mathrm{mod}\ p)$, т. е. в силу $b \neq b'$ имеем $z \equiv 1 (\mathrm{mod}\ p)$, что невозможно.

Рассмотрим исходную сумму:

$$\sum_{b=0}^{p-1} \chi(b)\chi(b+c) = \sum_{b=1}^{p-1} \chi(b)\chi(zb) = \sum_{b=1}^{p-1} (\chi(b))^2 \chi(z) = \sum_{b=1}^{p-1} \chi(b^2)\chi(z) =$$

$$= \sum_{\substack{z=0 \\ z\neq 1}}^{p-1} \chi(z) = \sum_{z=0}^{p-1} \chi(z) - \chi(1) = -1.$$

Здесь $\sum\limits_{z=0}^{p-1} \chi(z) = 0$, так как половина чисел от 1 до $p-1$ (p — нечетно) являются квадратичными вычетами и $\chi(z)$ для них равен 1, а половина — невычетами и $\chi(z)$ для них равен -1. ▲

Метод построения Пэйли дает построение матрицы Адамара порядка $n = p+1$, кратного 4 (или порядка $n = p^m + 1$, кратного 4, если используются квадратичные вычеты над полем $GF(p^m)$), где p — простое.

Для конструкции Пэйли потребуется матрица *Джекобстола* $Q = (q_{ij})$, которая является $(p \times p)$-матрицей, где $q_{ij} = \chi(j-i)$, строки и столбцы матрицы Q пронумерованы числами $0, 1, \ldots, p-1$.

Пример 12. При $p = 7$ числа $1, 2$ и 4 являются квадратичными вычетами, следовательно, $\chi(1) = \chi(2) = \chi(4) = 1$. Матрица Джекобстола имеет вид

$$Q = \begin{bmatrix} 0 & 1 & 1 & -1 & 1 & -1 & -1 \\ -1 & 0 & 1 & 1 & -1 & 1 & -1 \\ -1 & -1 & 0 & 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 0 & 1 & 1 & -1 \\ -1 & 1 & -1 & -1 & 0 & 1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 0 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 0 \end{bmatrix}.$$

При $p = 4k-1$ число -1 является квадратичным невычетом по свойству 2 и $\chi(-1) = -1$. Поэтому

$$q_{ij} = \chi(j-i) = \chi(-1)\chi(i-j) = -\chi(i-j) = -q_{ij},$$

следовательно, матрица Q — кососимметрическая, т. е. $Q^T = -Q$.

Имеет место следующее утверждение.

Лемма 9. *Справедливо* $QQ^T = pE - J$ *и* $QJ = JQ = \mathbf{0}_p$, *где* J — *матрица, элементами которой являются единицы, а* $\mathbf{0}_p$ — *матрица, состоящая из нулей.*

Доказательство. Пусть $P = (p_{ij}) = QQ^T$. Тогда

$$p_{ii} = \sum_{k=0}^{p-1} q_{ik}^2 = p - 1.$$

Если $i \neq j$, то

$$p_{ij} = \sum_{k=0}^{p-1} q_{ik}q_{jk} = \sum_{k=0}^{p-1} \chi(k-i)\chi(k-j) = \sum_{k=0}^{p-1} \chi(k-i)\chi(k-i+(i-j)) =$$

$$= \sum_{b=0}^{p-1} \chi(b)\chi(b+c) = -1,$$

где $b = k - i$ и $c = i - j$ по предыдущей теореме.

Легко проверяется утверждение $QJ = JQ = \mathbf{0}_p$ (действительно, каждый столбец и строка матрицы Q содержит равное число $(p-1)/2$ элементов 1 и -1). ▲

Построим, используя Q, матрицу H. Обозначим, как и прежде, через $\mathbf{1}$ и $\mathbf{0}$ векторы длины p, состоящие из единиц и нулей соответственно.

Теорема 56. *Матрица*

$$H = \begin{pmatrix} 1 & \mathbf{1} \\ \mathbf{1}^T & Q - E_p \end{pmatrix}$$

является матрицей Адамара (типа Пэйли).

Доказательство. Рассмотрим произведение

$$H \cdot H^T = \begin{pmatrix} 1 & \mathbf{1} \\ \mathbf{1}^T & Q - E_p \end{pmatrix} \cdot \begin{pmatrix} 1 & \mathbf{1} \\ \mathbf{1}^T & Q^T - E_p \end{pmatrix} = \begin{pmatrix} p+1 & \mathbf{0} \\ \mathbf{0}^T & J + (Q - E_p)(Q^T - E_p) \end{pmatrix}$$

Используя предыдущую лемму и тот факт, что по определению матрицы Q выполняется $Q + Q^T = \mathbf{0}$, получаем

$$J + (Q - E_p)(Q^T - E_p) = J + Q \cdot Q^T - (Q + Q^T) + E_p = J + (pE_p - J) + E_p = (p+1)E_p.$$

Следовательно,

$$H \cdot H^T = \begin{pmatrix} p+1 & \mathbf{0} \\ \mathbf{0}^T & (p+1)E_p \end{pmatrix} = (p+1)E_{p+1}.$$

▲

Другими словами, получили нормализованную матрицу Адамара типа Пэйли порядка $p + 1$.

9.1.4. Коды Адамара

Пусть H_n — нормализованная матрица Адамара. Заменим всюду 1 на 0, а -1 на 1, тогда H_n превращается в *двоичную* матрицу Адамара A_n. Так как строки H_n ортогональны, то любые две строки матрицы A_n совпадают в $n/2$ позициях, следовательно, расстояние Хэмминга между ними равно $n/2$.

Из матрицы A_n построим три кода Адамара.

1. Код $\mathcal{A}_n$ с параметрами $(n-1, n, n/2)$, состоящий из строк матрицы A_n с выколотой первой координатой.

Пример 13. Код Адамара $\mathcal{A}_8$ с параметрами $(7, 8, 4)$, полученный из матрицы Адамара типа Пэйли. Кодовые слова — это все циклические сдвиги вектора (1001011) и нулевой вектор длины 7. Иначе говоря,

$$[(1001011)] = \begin{matrix}(1&0&0&1&0&1&1)\\(1&1&0&0&1&0&1)\\(1&1&1&0&0&1&0)\\(0&1&1&1&0&0&1)\\(1&0&1&1&1&0&0)\\(0&1&0&1&1&1&0)\\(0&0&1&0&1&1&1)\end{matrix}.$$

Пример 14. Код Адамара $\mathcal{A}_{12}$ с параметрами $(11, 12, 5)$, полученный из матрицы Адамара типа Пэйли:

$$[(11011100010)],$$

содержащий также нулевой вектор длины 11.

2. Код $\mathcal{B}_n$ с параметрами $(n-1, 2n, (n/2)-1)$, состоящий из векторов кода $\mathcal{A}_n$ и их дополнений.

Пример 15. Код Адамара $\mathcal{B}_{12}$ с параметрами $(11, 24, 5)$ со следующими кодовыми словами:

$$\mathbf{0}^{11}, \mathbf{1}^{11}, [(11011100010)], [(00100011101)].$$

3. Код $\mathcal{C}_n$, $(n, 2n, n/2)$, состоящий из строк матрицы A_n и их дополнений.

Код $\mathcal{A}_n$ является *симплексным*. Это означает, что расстояние между любыми двумя кодовыми словами равно $n/2$.

Если H — матрица Сильвестра, то $\mathcal{A}_n$, $\mathcal{B}_n$ и $\mathcal{C}_n$ — групповые коды. Если H — матрица Адамара типа Пэйли, то полученные коды нелинейны для любого $n > 8$. Линейные оболочки этих кодов принадлежат классу так называемых квадратично-вычетных кодов.

9.1.5. Связь кодов Адамара с кодом Хэмминга

Рассмотрим связь кодов Адамара $\mathcal{A}_n$, $n = 2^k - 1$ с кодом Хэмминга. Для этого вспомним определение ортогонального кода.

Определение. Если C — линейный $[n,k]$-код над F, то *дуальный* или *ортогональный* к нему код $C^\perp$ определяется как множество всех векторов, ортогональных всем кодовым словам кода C:

$$C^\perp = \{u \mid u \cdot v = 0 \text{ для любого } v \in C\}.$$

Если код C имеет проверочную матрицу H и порождающую G, то дуальный код $C^\perp$ имеет проверочную матрицу G и порождающую H.

Теорема 57. *Код, дуальный к коду Хэмминга, является кодом Адамара $\mathcal{A}_n$, построенным из матрицы Сильвестра. Справедливо и обратное.*

Доказательство будем проводить индукцией по m, где $n = 2^m - 1$.

Пусть $m = 2$. Рассмотрим код Хэмминга длины 3, заданный следующей проверочной матрицей

$$H_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Она является порождающей матрицей ортогонального кода к коду Хэмминга. Кодовые слова имеют вид

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Добавляя проверку на четность и заменяя 0 на 1 и 1 на -1, получаем

$$\begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{matrix} \overset{\substack{0\to 1 \\ 1\to -1}}{\longrightarrow} H(4) = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{array}\right],$$

т. е. матрицу Сильвестра (см. пример 10).

Пусть для $n = 2^{m-1}-1$ имеем: код, ортогональный коду Хэмминга, является кодом Адамара $\mathcal{A}_{2^{m-1}}$ с порождающей матрицей H_n, являющейся проверочной матрицей кода Хэмминга длины n. Покажем, что матрица вида

$$H = \left[\begin{array}{ccc|c|ccc} 0 & \dots & 0 & 1 & 1 & \dots & 1 \\ & & & 0 & & & \\ & H_n & & \vdots & & H_n & \\ & & & 0 & & & \end{array}\right]$$

является проверочной матрицей кода Хэмминга порядка $2n-1$. Действительно, это так, поскольку число столбцов равно $2n-1$ и все они являются различными векторами длины m (в H_n все столбцы — различные векторы длины $m-1$).

Строки матрицы

$$\left[\begin{array}{ccc|c|ccc} & & & 0 & & & \\ & H_n & & \vdots & & H_n & \\ & & & 0 & & & \end{array}\right],$$

порождают код с кодовой матрицей

$$\left[\begin{array}{c|c|c} & 0 & \\ \mathcal{A}_n & \vdots & \mathcal{A}_n \\ & 0 & \end{array}\right].$$

Добавляя вектор $(0\ldots01\ldots1)$, получаем код, порожденный матрицей H, кодовая матрица которого имеет вид

$$\left[\begin{array}{c|c|c} & 0 & \\ \mathcal{A}_n & \vdots & \mathcal{A}_n \\ & 0 & \\ \hline & 1 & \\ \mathcal{A}_n & \vdots & \overline{\mathcal{A}_n} \\ & 1 & \end{array}\right] = \mathcal{A}_{2n+1}.$$

Отсюда, добавив столбец из нулей, имеем матрицу

$$\left[\begin{array}{c|c|c|c} 0 & & 0 & \\ \vdots & \mathcal{A}_n & \vdots & \mathcal{A}_n \\ 0 & & 0 & \\ \hline 0 & & 1 & \\ \vdots & \mathcal{A}_n & \vdots & \overline{\mathcal{A}_n} \\ 0 & & 1 & \end{array}\right] \overset{\substack{0\to1\\1\to-1}}{\longrightarrow} H(2n) = \left[\begin{array}{c|c} H(n) & H(n) \\ \hline H(n) & \overline{H(n)} \end{array}\right],$$

где $H(n)$ — матрица Сильвестра порядка n, следовательно, по теореме 54 матрица $H(2n)$ также является матрицей Сильвестра.

Доказательство в обратную сторону проводится также по индукции, начиная с кода Адамара. ▲

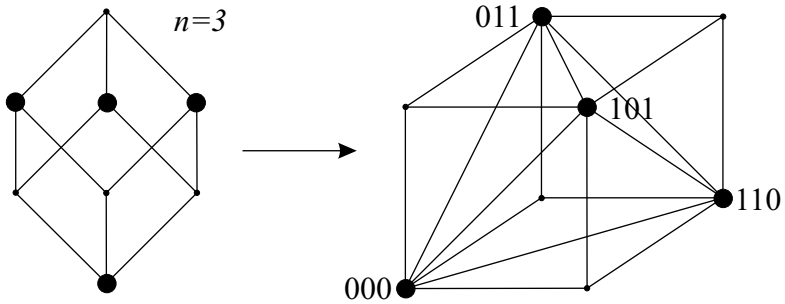

Рис. 9. Симплексный код длины 3

Код Адамара $\mathcal{A}_n$ называется также *симплексным*, поскольку его вершины образуют в E^n правильный симплекс — расстояние между любыми двумя кодовыми словами одно и то же и равно $\frac{n+1}{2}$. На рис. 9 приведен пример кода $\mathcal{A}_3$, вершины которого образуют правильный тетраэдр.

9.2. Коды Рида-Маллера

Определение кодов Рида-Маллера ($\mathcal{RM}$-кодов) удобнее всего приводить в терминах булевых функций. Рассмотрим произвольную булеву функцию $f(x_1, \ldots, x_m)$ от m переменных, тождественно не равную нулю. В качестве кодовых слов кода Рида-Маллера длины $n = 2^m$ будут выбраны двоичные векторы, являющиеся наборами значений булевых функций специального вида. Но прежде чем задать способ выбора этих функций, приведем некоторые определения и утверждения. Каждой строке таблицы истинности произвольной функции f, для которой значение функции равно единице, можно поставить в соответствие элементарную конъюнкцию длины m, равную единице на этом наборе, т. е. произведение всех переменных $x_1, \ldots, x_m$, взятых с отрицанием или без. Дизъюнкция этих элементарных коньюнкций называется *совершенной дизъюнктивной нормальной формой (СДНФ)* булевой функции f.

Далее, используя закон де Моргана

$$x \vee y = \neg(\neg x \& \neg y),$$

удалим все дизъюнкции в представлении функции $f(x_1, \ldots, x_m)$, другими словами, функция $f(x_1, \ldots, x_m)$ будет выражена через функции из множества $\{\&, \neg\}$. Так как f — произвольная булева функция, тождественно не равная нулю, то отсюда следует, что система функций $\{\&, \neg\}$ полна. Напомним, что система булевых функций $\{f_1, \ldots, f_s, \ldots\}$ называется *полной*, если любая булева функция может быть записана в виде формулы через функции этой системы.

Теорема 58 (Жегалкин). *Каждая булева функция может быть однозначно представлена посредством полинома по модулю* 2.

Этот полином носит название *полинома Жегалкина.*

Доказательство. Рассмотрим произвольную булеву функцию $f(x_1, \ldots, x_m)$, представленную системой функций $\{\&, \neg\}$. Выражая всякий раз отрицание через сложение: $\neg x = x + 1$ и опуская знаки &: $x\&y = xy$, после раскрытия скобок и приведения подобных членов $x + x = 0$, $xx = x$, используя дистрибутивность, получим многочлен

$$\sum_{1 \le i_1, \ldots, i_s \le m} a_{i_1 \ldots i_s} x_{i_1} \cdot \ldots \cdot x_{i_s} + a,$$

где $a_{i_1 \ldots i_s}, a \in \{0, 1\}$ — константы для различных $i_1, \ldots, i_s$. Полученный многочлен является полиномом Жегалкина. Теперь покажем, что произвольная булева функция однозначно представляется полиномом Жегалкина. Подсчитаем число различных полиномов Жегалкина

$$\sum_{1 \le i_1, \ldots, i_s \le m} a_{i_1 \ldots i_s} x_{i_1} \cdot \ldots \cdot x_{i_s} + a.$$

Для фиксированного $s, s \le m$, имеем $\begin{pmatrix} m \\ s \end{pmatrix}$ возможностей выбора произведения $x_{i_1} \cdot \ldots \cdot x_{i_s}$. Так как $0 \le s \le m$, то всего возможно

$$\sum_{s=0}^{m} \begin{pmatrix} m \\ s \end{pmatrix} = 2^m$$

различных элементарных конъюнкций от переменных $x_1, \ldots, x_m$, взятых без отрицания. Далее, так как коэффициенты $a_{i_1 \ldots i_s}, a$ могут принимать два значения 0 или 1, получаем 2^{2^m} различных полиномов Жегалкина, каждому из которых отвечает единственная булева функция. С другой стороны, число всех булевых функций от m переменных также равно 2^{2^m}. ▲

Из теоремы Жегалкина получаем, что функции

$$1, x_1, \ldots, x_m, x_1x_2, \ldots, x_{m-1}x_m, \ldots, x_1x_2 \ldots x_m$$

образуют базу пространства E^{2^m} всех булевых функций от m переменных. Теперь перейдем к определению кодов Рида-Маллера.

Определение. *Двоичный код Рида-Маллера* $\mathcal{RM}(r, m)$ *порядка* r, $0 \leq r \leq m$, — это совокупность векторов длины 2^m, отвечающих полиномам от m переменных степени не больше r.

Код Рида-Маллера $\mathcal{RM}(1, m)$ первого порядка ортогонален расширенному коду Хэмминга и совпадает с кодом Адамара $\mathcal{B}_m$. В свою очередь, расширенный код Хэмминга является кодом Рида-Маллера $\mathcal{RM}(m-2, m)$ порядка $m-2$. Из определения кода Рида-Маллера порядка r вытекает, что он состоит из всех линейных комбинаций векторов, соответствующих произведениям

$$1, x_1, \ldots, x_m, x_1x_2, \ldots, x_{m-1}x_m, \ldots, x_{m-r+1}x_{m-r+2} \ldots x_m.$$

Эти произведения задают базис кода Рида-Маллера порядка r. Отсюда следует, что размерность кода равна

$$k = 1 + \left(\begin{array}{c} m \\ 1 \end{array} \right) + \cdots + \left(\begin{array}{c} m \\ r \end{array} \right),$$

все кодовые слова имеют четный вес.

Код Рида-Маллера $\mathcal{RM}(r, m)$ любого порядка r, $0 \leq r \leq m$, может быть описан с помощью конструкции Плоткина.

Теорема 59. *Справедливо*

$$\mathcal{RM}(r+1, m+1) = \{ (u, u+v) \mid u \in \mathcal{RM}(r+1, m), v \in \mathcal{RM}(r, m) \}.$$

Доказательство. Рассмотрим произвольное кодовое слово f из $\mathcal{RM}(r+1, m+1)$. С одной стороны, оно представляет собой двоичный вектор длины 2^{m+1}, с другой стороны — многочлен от $(m+1)$ переменных, степень которого не больше $r+1$. Перепишем многочлен $f(x_1, \ldots x_{m+1})$ следующим образом

$$f(x_1, \ldots, x_{m+1}) = g(x_1, \ldots, x_m) + x_{m+1}h(x_1, \ldots, x_m),$$

где степень $g(x_1, \ldots, x_m)$ не больше $r+1$, а степень $h(x_1, \ldots, x_m)$ не больше r. Пусть g и h — векторы длины 2^m, отвечающие многочленам $g(x_1, \ldots, x_m)$ и $h(x_1, \ldots, x_m)$ соответственно, иными словами, g и h — это наборы значений булевых функций (от переменных $x_1, \ldots, x_m$), представленных этими многочленами. Поскольку степень $g(x_1, \ldots, x_m)$ не больше $r+1$, то, по определению кода Рида-Маллера, вектор g

принадлежит коду $\mathcal{RM}(r+1, m)$ порядка $r+1$. Аналогично, h принадлежит коду $\mathcal{RM}(r, m)$.

Рассмотрим многочлен $g(x_1, \ldots, x_m)$ как многочлен от переменных $x_1, \ldots, x_{m+1}$:

$$g'(x_1, \ldots, x_{m+1}) = g(x_1, \ldots, x_m).$$

Вектор длины 2^{m+1}, отвечающий этому многочлену, равен (g, g). Действительно, из булевой функции $g(x_1, \ldots, x_m)$ от m переменных мы получили булеву функцию $g'(x_1, \ldots, x_{m+1})$ от $(m+1)$ переменных, где переменная x_{m+1} является фиктивной. Вторая половина таблицы истинности булевой функции $g'(x_1, \ldots, x_{m+1})$, отвечающая значению 1 переменной x_{m+1}, дублирует первую.

Рассмотрим также многочлен $h'(x_1, \ldots, x_{m+1})$ как многочлен от переменных $x_1, \ldots, x_{m+1}$, определенный с помощью многочлена $h(x_1, \ldots, x_m)$ следующим образом

$$h'(x_1, \ldots, x_{m+1}) = x_{m+1}h(x_1, \ldots, x_m) =$$
$$= 0 \cdot h(x_1, \ldots, x_m) + 1 \cdot h(x_1, \ldots, x_m).$$

Нетрудно видеть, что этому многочлену отвечает двоичный вектор длины 2^{m+1} вида $(\mathbf{0}^{2^m}, h)$. Таким образом, получили

$$f = (g, g) + (\mathbf{0}^{2^m}, h) = (g, g + h).$$

▲

Теорема 60. *Минимальное расстояние кода Рида-Маллера $\mathcal{RM}(r, m)$ порядка r, $0 \le r \le m$, равно $d = 2^{m-r}$.*

Доказательство. Доказательство будем проводить индукцией по m.

Пусть $m = 1$. В этом случае имеем следующие два тривиальных кода Рида-Маллера порядков 0 и 1 соответственно:

$\mathcal{RM}(0, 1) = \{(00), (11)\}$ с кодовым расстоянием 2;
$\mathcal{RM}(1, 1) = \{(00), (01), (11), (10)\}$ с кодовым расстоянием 1.

Пусть для любого $(m-1)$ теорема верна и минимальное расстояние кода $\mathcal{RM}(r, m-1)$ порядка r равно 2^{m-1-r} для любого r, удовлетворяющего условию $0 \le r \le m-1$. Рассмотрим конструкцию Плоткина для кода $\mathcal{RM}(r, m)$, изложенную в предыдущей теореме:

$$\mathcal{RM}(r, m) = \{(u, u+v) | u \in \mathcal{RM}(r, m-1), v \in \mathcal{RM}(r-1, m-1)\}.$$

По предположению индукции, кодовое расстояние кода $\mathcal{RM}(r-1, m-1)$ равно $2^{m-1-(r-1)} = 2^{m-r}$, а кода $\mathcal{RM}(r, m-1)$ равно 2^{m-r-1}. Нетрудно видеть, что

$$d((u, u+v), (u', u'+v')) \ge 2^{m-r}$$

для произвольных кодовых слов $(u, u+v)$ и $(u', u'+v')$ из $\mathcal{RM}(r, m)$ (см. доказательство теоремы Плоткина в гл. 1). Очевидно, что между кодовыми словами кода $\mathcal{RM}(r, m)$ достижимо расстояние, равное 2^{m-r}. ▲

Таким образом, код Рида-Маллера $\mathcal{RM}(r,m)$ порядка r имеет параметры:

- длина кода равна $n = 2^m$;
- мощность кода равна 2^k, где $k = 1 + \left(\begin{array}{c} m \\ 1 \end{array} \right) + \cdots + \left(\begin{array}{c} m \\ r \end{array} \right)$;
- кодовое расстояние $d = 2^{m-r}$.

Коды Рида-Маллера имеют быстрые процедуры кодирования и декодирования и широко используются на практике, в частности, коды Рида-Маллера третьего порядка представляют собой хороший экземпляр кодов, которые могут быть использованы в криптографии для шифрования информации в криптосистемах с открытыми ключами, а именно в криптосистеме МакЭлиса.

Теорема 61. *Для любых* $r, 0 \leq r \leq (m-1)$ *код Рида-Маллера* $\mathcal{RM}(m-r-1,m)$ *ортогонален коду Рида-Маллера* $\mathcal{RM}(r,m)$.

Доказательство. Рассмотрим произвольные векторы f и g, где $f \in \mathcal{RM}(m-r-1,m)$, $g \in \mathcal{RM}(r,m)$. По определению кодов Рида-Маллера, степени многочленов $f(x_1,\ldots,,\, x_m)$ и $g(x_1,\,\ldots,\,x_m)$ не превосходят $m-r-1$ и r соответственно. Следовательно, степень многочлена $f(x_1,\ldots,\,x_m)\cdot g(x_1,\ldots,\,x_m)$ не превосходит $m-1$ и отвечающий этому многочлену вектор принадлежит коду $\mathcal{RM}(m-1,m)$. Поскольку в коде $\mathcal{RM}(m-1,m)$ все вершины имеют четный вес, то есть скалярное произведение векторов f и g равно нулю, то получаем $\mathcal{RM}(m-r-1,m) \subseteq \mathcal{RM}(r,m)^{\perp}$. Но

$$\dim\mathcal{RM}(m-r-1,m) + \dim\mathcal{RM}(r,m) = 2^m,$$

откуда следует, что

$$\mathcal{RM}(m-r-1,m) = \mathcal{RM}(r,m)^{\perp},$$

что доказывает теорему. ▲

9.2.1. Коды с параметрами кодов Рида-Маллера

Выколотый код Рида-Маллера $\mathcal{RM}^*(r,m)$ получается удалением какой-либо координаты и имеет параметры

$(n = 2^m - 1, M = 2^k, d = 2^{m-r} - 1)$, где $k = 1 + \left(\begin{array}{c} m \\ 1 \end{array} \right) + \cdots + \left(\begin{array}{c} m \\ r \end{array} \right)$.

Покажем, как, используя свитчинговые и каскадные конструкции, можно строить мощные классы кодов с параметрами кодов Рида-Маллера и выколотых кодов Рида-Маллера. Двоичные коды (не обязательно линейные) с параметрами выколотого кода Рида-Маллера порядка r будем обозначать $\mathcal{LRM}^*(r,m)$ и называть *выколотыми кодами Рида-Маллера*. Линейный код $\mathcal{RM}^*(r,m)$ является частным случаем кодов $\mathcal{LRM}^*(r,m)$.

Сначала рассмотрим применение конструкции Васильева из теоремы 13. Посредством этого метода построения кодов можно получить много нелинейных выколотых кодов Рида-Маллера $\mathcal{LRM}^*(r,m))$ порядка r.

Пусть $\mathcal{LRM}^*(r,m-1)$ и $\mathcal{LRM}^*(r-1,m-1)$ — произвольные выколотые коды Рида-Маллера порядков r и $r-1$ соответственно длины $n = 2^{m-1} - 1$. Справедливо следующее утверждение:

Теорема 62 (Пулатов А. К., 1962). *Множество векторов*

$$\{\,(x, |x| + \lambda(y), x + y) \mid x \in \mathcal{LRM}^*(r, m-1), y \in \mathcal{LRM}^*(r-1, m-1)\,\},$$

где λ — произвольная функция из кода $\mathcal{LRM}^*(r-1, m-1)$ в множество $\{0, 1\}$, является кодом с параметрами выколотого кода Рида-Маллера порядка r длины $n = 2^m - 1$, т. е. $\mathcal{LRM}^*(r, m)$-кодом.

Доказательство теоремы аналогично доказательству теоремы 13.

Замечание. Этот класс кодов оказался хорошим представителем для реализаций π-схемами. А. К. Пулатовым была получена квадратичная нижняя оценка сложности π-схем, реализующих коды с параметрами выколотых кодов Рида-Маллера.

Используя теорему 62, можно показать, что существуют нетривиальные разбиения выколотых кодов Рида-Маллера $\mathcal{LRM}^*(r, m-1)$ порядка r длины $n = 2^{m-1} - 1$ на выколотые коды Рида-Маллера порядка $r-1$ длины $n = 2^{m-1} - 1$:

$$\mathcal{LRM}^*(r, m-1) = \bigcup_{i=1}^{t} \mathcal{LRM}_i^*(r-1, m-1), \quad \text{где } t = 2^{\binom{m}{r}}$$

(доказательство этого факта аналогично доказательству теоремы 21).

Теорема 63 (Соловьева Ф. И., 1981). *Пусть даны произвольные разбиения выколотых кодов Рида-Маллера порядка r:*

$$\mathcal{LRM}_1^*(r, m-1) = \bigcup_{i=1}^{t} \mathcal{LRM}_{1,i}^*(r-1, m-1),$$

$$\mathcal{LRM}_2^*(r, m-1) = \bigcup_{i=1}^{t} \mathcal{LRM}_{2,i}^*(r-1, m-1),$$

где $t = 2^{\binom{m}{r}}$. Пусть π — произвольная подстановка на t элементах. Тогда множество

$$\{(x, y, |y|) : x \in \mathcal{LRM}_{1,i}^*(r-1, m-1), y \in \mathcal{LRM}_{2,\pi(i)}^*(r-1, m-1), i = 1, \ldots, t\}$$

является $\mathcal{LRM}^*(r, m)$-кодом.

Доказательство теоремы аналогично доказательству теоремы 22.

Замечание. Легко видеть, что применение проверки на четность к $\mathcal{LRM}^*(r, m)$-кодам позволяет получить широкие классы неэквивалентных нелинейных кодов с параметрами кодов Рида-Маллера порядка r длины $n \geq 16$. Среди этих кодов могут оказаться коды с полезными свойствами, например с транзитивной группой автоморфизмов (код называется *транзитивным*, если его группа автоморфизмов действует транзитивно на всех его кодовых словах), а также коды, являющиеся $\mathbb{Z}_4$-линейными (линейными над кольцом $\mathbb{Z}_4$). Последние могут найти применение в криптографии, например, в криптосистемах МакЭлиса с открытыми ключами.

9.3. Коды Препараты

Определение. Максимальный двоичный код длины $n = 4^m$, где $m \geq 2$ с кодовым расстоянием $d = 6$ называется *кодом Препараты*. Его мощность равна $2^n/n^2$.

Первый такой код (для каждой допустимой длины) построил Франко Препарата в 1968 г.

В 1971–1973 гг. Н. В. Семаков, В. А. Зиновьев и Г. В. Зайцев исследовали свойства кодов с параметрами кодов Препараты и связь их с двоичными совершенными кодами, исправляющими одну ошибку. Ими было показано, что коды Препараты, помимо максимальности, обладают некоторыми весьма нетривиальными свойствами, а именно эти коды *равномерно упакованы* и *дистанционно инвариантны*. Кроме того, было показано, что любой код Препараты является подкодом некоторого расширенного совершенного кода той же длины с расстоянием 4 (причем единственного) и в некотором смысле плотно упакован в нем.

В 1972 г. Я. М. Геталс и С. Л. Сновер показали, что линейных кодов Препараты не существует. В 1976 г. И. Думер и позднее, в 1983 г. Р. Д. Бакер, Я. Х. Ван Линт и Р. М. Вилсон построили семейство кодов Препараты, содержащее оригинальный код Препараты. В 1994 г. А. Р. Хэммонс, П. В. Кумар, А. Р. Кальдербанк, Н. Дж. А. Слоэн и П. Соле предложили конструкцию первого кода Препараты с групповой структурой — $\mathbb{Z}_4$-*линейного* кода Препараты.

Конструкция Думера-Бакера

Соответствие между множествами и векторами

Пусть m — нечетное целое число, $m \geq 3$ и $n = 2^m - 1$. Пусть α — примитивный элемент поля $GF(2^m)$. Имеем

$$GF(2^m) = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}.$$

Рассмотрим множество

$$I = \{-\infty, 0, 1, 2, \ldots, n-1\},$$

состоящее из логарифмов элементов поля по основанию α (полагаем, как и раньше в гл. 6, что $\alpha^{-\infty} = 0$). Произвольному подмножеству X множества элементов поля $GF(2^m)$ взаимно однозначно сопоставим двоичный вектор c длины $n+1$, координаты которого занумерованы с помощью множества I:

$$X \longleftrightarrow c = (c_{-\infty}, c_0, c_1, c_2, \ldots, c_{n-1}) \tag{9.1}$$

по правилу

$$\alpha^i \in X \Longleftrightarrow c_i = 1.$$

Другими словами, вектор c является *характеристическим вектором* множества X: его ненулевые позиции указывают элементы поля, включаемые в множество X. Например, пустому множеству $\emptyset$ отвечает нулевой вектор. Введем некоторые обозначения для множеств $X, Y \in GF(2^m)$:

— как обычно через $|X|$ обозначается мощность множества X;

— $X \triangle Y$ — симметрическая разность множеств X и Y. Заметим, что эта операция на множествах отвечает двоичному сложению соответствующих характеристических векторов;

— $X + a$ — сдвиг множества X на элемент поля a, т. е. $X + a = \{x + a \,|\, x \in X\}$;

— aX — умножение X на ненулевой элемент поля a, т. е. $aX = \{ax \,|\, x \in X\}$.

Используя соответствие (9.1), произвольный двоичный вектор длины $2(n+1) = 2^{m+1}$ можно описать парой (X, Y), где X, Y — подмножества множества элементов поля $GF(2^m)$. Далее мы не будем делать различия между парами множеств (X, Y) и двоичными векторами длины 2^{m+1}. Из контекста далее будет ясно, речь идет о множествах или о векторах.

Построение кодов Препараты

Группа автоморфизмов поля $GF(2^m)$ состоит из всех отображений вида

$$x \to x^\sigma,$$

где $\sigma = 2^k$, $k = 0, 1, \ldots, m-1$. Рассмотрим из них такие, что отображения

$$x \to x^{\sigma-1},$$
$$x \to x^{\sigma+1}$$

взаимно однозначны. Другими словами, для этого выберем такие σ, что выполняются условия

$$\begin{aligned} (\sigma - 1,\, 2^m - 1) &= 1, \\ (\sigma + 1,\, 2^m - 1) &= 1. \end{aligned} \tag{9.2}$$

Считаем далее, что $\sigma = 2^k$ удовлетворяет условиям (9.2).

Теорема 64. *Код $P(\sigma)$ длины 2^{m+1}, состоящий из всех двоичных векторов, описываемых парами (X, Y), удовлетворяющими условиям:*

$$\begin{aligned} &1) \quad |X| \text{ — четно}, \ |Y| \text{ — четно}; \\ &2) \quad \sum_{x \in X} x = \sum_{y \in Y} y; \\ &3) \quad \sum_{x \in X} x^{\sigma+1} + \left(\sum_{x \in X} x \right)^{\sigma+1} = \sum_{y \in Y} y^{\sigma+1}, \end{aligned} \tag{9.3}$$

является кодом Препараты.

Для доказательства этой теоремы потребуется ввести некоторые дополнительные определения и изучить свойства кодов $P(\sigma)$. Сначала приведем следующие простые свойства поля $GF(2^m)$.

Лемма 10. *Для любых элементов a, b поля $GF(2^m)$ и любого $\sigma = 2^k$ имеет место тождество*

$$(a+b)^{\sigma+1} = a^{\sigma+1} + a^\sigma b + ab^\sigma + b^{\sigma+1}.$$

Доказательство. Согласно теореме 35 (см. разд. 6.2.), имеем

$$(a+b)^\sigma = a^\sigma + b^\sigma.$$

Тогда, преобразовав $(a+b)^{\sigma+1}$ к виду $(a+b)^\sigma(a+b)$, получим искомое тождество. ▲

Несложно доказывается

Лемма 11. *Для любого множества X из носителя поля $GF(2^m)$ и любого $\sigma = 2^k$ выполняется*

$$\left(\sum_{x\in X} x\right)^\sigma = \sum_{x\in X} x^\sigma.$$

Напомним, что двоичный код называется *дистанционно-инвариантным*, если число кодовых слов, находящихся на некотором расстоянии l от фиксированной кодовой вершины, не зависит от выбора этой вершины, а зависит только от l.

Утверждение 16. *Код $P(\sigma)$ дистанционно-инвариантен.*

Доказательство. По построению код $P(\sigma)$ содержит вектор $(\emptyset, \emptyset)$. Пусть (X_0, Y_0) — произвольный вектор из $P(\sigma)$. Для доказательства утверждения достаточно построить взаимно однозначное соответствие между множествами кодовых векторов, находящихся на расстоянии l от вершин $(\emptyset, \emptyset)$ и (X_0, Y_0) соответственно, где l — некоторое фиксированное расстояние.

Зададим следующее взаимно однозначное отображение:

$$f : (X, Y) \to (U, V),$$

по правилу

$$U = X \triangle X_0 + a, \text{ где } a = \sum_{x\in X_0} x;$$
$$V = Y \triangle Y_0.$$

Заметим, что расстояние от (X, Y) до (X_0, Y_0) равно весу вектора $(X\triangle X_0, Y\triangle Y_0)$ и равно расстоянию от (U, V) до $(\emptyset, \emptyset)$. Покажем, что если (X, Y) принадлежит коду, то и вектор (U, V) ему принадлежит. В силу взаимной однозначности отображения f, этого будет достаточно для доказательства дистанционной инвариантности $P(\sigma)$.

Пусть (X, Y) — кодовый вектор. Проверим, что вектор (U, V) также кодовый, т. е. для него выполняются условия (9.3):

1) $|X|$, $|X_0|$, $|Y|$, $|Y_0|$ — четны, следовательно, $|X\triangle X_0 + a|$ и $|Y\triangle Y_0|$ четны.

2) Докажем равенство

$$\sum_{x\in U} x = \sum_{y\in V} y.$$

Действительно, имеем

$$\sum_{x\in U} x = \sum_{x\in X\triangle X_0} (x+a) = \left(\sum_{x\in X\triangle X_0} x\right) + a|X\triangle X_0| =$$

$$= \sum_{x\in X\triangle X_0} x = \sum_{x\in X} x + \sum_{x\in X_0} x = \sum_{y\in Y} y + \sum_{y\in Y_0} y = \sum_{y\in V} y.$$

3) Проверим выполнение равенства

$$\sum_{x\in U} x^{\sigma+1} + \left(\sum_{x\in U} x\right)^{\sigma+1} = \sum_{y\in V} y^{\sigma+1}.$$

Преобразуя первое слагаемое, с учетом леммы 10, имеем

$$\sum_{x\in U} x^{\sigma+1} = \sum_{x\in X\triangle X_0} (x+a)^{\sigma+1} =$$

$$= \sum_{x\in X\triangle X_0} (x^{\sigma+1} + x^\sigma a + x a^\sigma + a^{\sigma+1}) = \sum_{x\in X\triangle X_0} x^{\sigma+1} + \left(\sum_{x\in X\triangle X_0} x^\sigma\right) a + \left(\sum_{x\in X\triangle X_0} x\right) a^\sigma =$$

$$= \sum_{x\in X} x^{\sigma+1} + \sum_{x\in X_0} x^{\sigma+1} + \left(\sum_{x\in X} x^\sigma\right) a + a^{\sigma+1} + \left(\sum_{x\in X} x\right) a^\sigma + a^{\sigma+1}.$$

Используя лемму 11, получаем

$$\sum_{x\in U} x^{\sigma+1} = \sum_{x\in X} x^{\sigma+1} + \sum_{x\in X_0} x^{\sigma+1} + \left(\sum_{x\in X} x\right)^\sigma a + \left(\sum_{x\in X} x\right) a^\sigma.$$

Для второго слагаемого, используя лемму 10, имеем

$$\left(\sum_{x\in U} x\right)^{\sigma+1} = \left(\sum_{x\in X\triangle X_0} (x+a)\right)^{\sigma+1} = \left(\sum_{x\in X\triangle X_0} x\right)^{\sigma+1} = \left(\sum_{x\in X} x + a\right)^{\sigma+1} =$$

$$= \left(\sum_{x\in X} x\right)^{\sigma+1} + \left(\sum_{x\in X} x\right)^\sigma a + \left(\sum_{x\in X} x\right) a^\sigma + a^{\sigma+1}.$$

Тогда

$$\sum_{x\in U} x^{\sigma+1} + \left(\sum_{x\in U} x\right)^{\sigma+1} = \sum_{x\in X} x^{\sigma+1} + \left(\sum_{x\in X} x\right)^{\sigma+1} + \sum_{x\in X_0} x^{\sigma+1} + \left(\sum_{x\in X_0} x\right)^{\sigma+1} =$$

$$= \sum_{y\in Y} y^{\sigma+1} + \sum_{y\in Y_0} y^{\sigma+1} = \sum_{y\in V} y^{\sigma+1}.$$

Таким образом, дистанционная инвариантность кода $P(\sigma)$ доказана. ▲

Утверждение 17. *Группа автоморфизмов кода $P(\sigma)$ содержит отображения:*

$$\begin{array}{ll} 1. & (X,Y) \to (X+a, Y+a) \text{ для любого } a \in GF(2^m); \\ 2. & (X,Y) \to (Y,X); \\ 3. & (X,Y) \to (aX, aY) \text{ для любого ненулевого } a \in GF(2^m); \\ 4. & (X,Y) \to (\varphi(X), \varphi(Y)) \text{ для любого } \varphi \in \operatorname{Aut} GF(2^m). \end{array} \tag{9.4}$$

Доказательство получается непосредственной проверкой выполнения условий (9.3) для образов кодового вектора (X,Y) под действием отображений 1–4. ▲

Утверждение 18. *Код $P(\sigma)$ имеет кодовое расстояние* 6.

Доказательство. Поскольку код $P(\sigma)$ содержит нулевую вершину и в силу утверждения 16 дистанционно инвариантен, достаточно показать, что минимальный ненулевой вес его кодовых слов равен 6.

Заметим, что все кодовые векторы (X,Y), в силу п. 1 условия (9.3), имеют четный вес. Из пунктов 1 и 2 условия (9.3) следует, что в коде нет слов веса 2. Покажем, что в коде $P(\sigma)$ нет слов и веса 4. Предположим обратное. Возможны два случая:

Случай 1. Вектор $(\{a,b\},\{c,d\})$ — кодовый, где $a \neq b$ и $c \neq d$.

Из утверждения 17 следует, что без ограничения общности в качестве a можно взять 0. Действительно, рассмотрим кодовое слово веса 4, полученное из выбранного действием автоморфизма $(X,Y) \to (X-a, Y-a)$. По п. 3 условия (9.3) имеем

$$(0^{\sigma+1} + b^{\sigma+1}) + (0+b)^{\sigma+1} = 0 = c^{\sigma+1} + d^{\sigma+1}.$$

Тогда $c^{\sigma+1} = d^{\sigma+1}$. Поскольку σ удовлетворяет условиям (9.2), отображение

$$x \to x^{\sigma+1}$$

взаимно однозначно, и, следовательно, $c = d$, что противоречит выбору c и d.

Случай 2. Вектор $(\{a,b,c,d\},\emptyset)$ — кодовый, где все элементы a,b,c,d различны.

В силу утверждения 17, считаем $a = 0$. Из пп. 2 и 3 условия (9.3) имеем

$$\begin{aligned} b + c + d &= 0, \\ b^{\sigma+1} + c^{\sigma+1} + d^{\sigma+1} &= 0. \end{aligned}$$

Тогда, выражая из первого равенства d и подставляя во второе, с учетом леммы 10 получаем

$$b^{\sigma+1} + c^{\sigma+1} + (b+c)^{\sigma+1} = b^\sigma c + bc^\sigma = bc(b^{\sigma-1} + c^{\sigma-1}) = 0.$$

В силу условия (9.2), отображение

$$x \to x^{\sigma-1}$$

взаимно однозначно, поэтому $b = c$. Противоречие с выбором b, c.

Покажем теперь, что код $P(\sigma)$ содержит слово веса 6. Рассмотрим попарно различные элементы a,b,c поля $GF(2^m)$. Определим d и e из системы уравнений

$$\begin{cases} a^{\sigma+1} + b^{\sigma+1} + c^{\sigma+1} = d^{\sigma+1}, \\ a + b + c + d = e. \end{cases}$$

Несложно убедиться, что вектор $(\{0, e\}, \{a, b, c, d\})$ удовлетворяет условиям (9.3) и, следовательно, является кодовым. ▲

Утверждение 19. *Мощность кода $P(\sigma)$ длины $N = 2^{m+1}$ равна $2^N/N^2$.*

Доказательство. Определим число пар (X, Y), удовлетворяющих условиям (9.3).

Выберем множество X так, чтобы мощность $|X|$ была четной. Это можно сделать 2^{2^m-1} способами. Выберем множество Y из ненулевых элементов поля так, чтобы выполнялись пункты 2 и 3 условия (9.3). Затем, если необходимо, добавим в множество Y элемент 0, чтобы мощность $|Y|$ была четной.

Пусть $X = \emptyset$. Тогда для Y выполняется система равенств

$$\begin{cases} \sum\limits_{y \in Y} y = 0, \\ \\ \sum\limits_{y \in Y} y^{\sigma+1} = 0, \end{cases} \tag{9.5}$$

см. пункты 2 и 3 в условии (9.3). Эту систему над $GF(2^m)$ можно рассматривать как систему из $2m$ уравнений над полем $GF(2)$, заменив каждый элемент поля $GF(2^m)$ на вектор-столбец длины m над $GF(2)$.

Пусть α — примитивный элемент поля $GF(2^m)$. Поскольку σ удовлетворяет условиям (9.2), элемент $\alpha^{\sigma+1}$ имеет порядок $2^m - 1$, и следовательно, также является примитивным элементом поля $GF(2^m)$. Минимальные многочлены $M^{(1)}(x)$, $M^{(\sigma+1)}(x)$, согласно свойствам 1 и 6 из разд. 7.6., неприводимы и имеют степень m. Тогда множество Y, удовлетворяющее системе (9.5), соответствует кодовому слову циклического кода длины $n = 2^m - 1$ с порождающим многочленом $M^{(1)}(x) \cdot M^{(\sigma+1)}(x)$. Мощность такого циклического кода равна 2^{n-2m}.

Выберем другое множество X четной мощности. Тогда всевозможным множествам Y таким, чтобы выполнялись условия (9.3), будут соответствовать векторы некоторого смежного класса циклического кода длины $n = 2^m - 1$ с порождающим многочленом $M^{(1)}(x) \cdot M^{(\sigma+1)}(x)$. Значит для фиксированного множества X число способов выбора подходящего множества Y равно 2^{2^m-1-2m}.

Таким образом, число различных пар (X, Y), удовлетворяющих условиям (9.3), а значит принадлежащих коду $P(\sigma)$, равно

$$\underbrace{2^{2^m-1}}_{\text{выбор } X} \cdot \underbrace{2^{2^m-1-2m}}_{\text{выбор } Y \text{ для } X} = 2^{2^{m+1}-2(m+1)} = 2^N/N^2.$$

Утверждение доказано. ▲

Из утверждений 18 и 19 немедленно вытекает теорема 64.

В литературе кодом Препараты иногда называют максимальный двоичный код длины $n - 1$ с кодовым расстоянием 5. Очевидно, что каждый такой код получается

выкалыванием одной координаты некоторого кода Препараты. Заметим, что операция расширения кода с помощью добавления проверки на четность, вообще говоря, не является взаимно обратной для операции выкалывания. Однако, можно показать, что в случае совершенных кодов и кодов Препараты эти операции всегда взаимно обратны.

Заключение

Безусловно, представленный вниманию читателя материал не претендует на полноту освещения всех областей теории кодов, корректирующих ошибки в каналах связи с шумами. Цель настоящих лекций — познакомить читателя с математическими основами современной теории кодирования, подготовить его к чтению специальной литературы по теории кодирования, а также таких смежных дисциплин, как криптография и сжатие данных. Для понимания изложенного материала достаточно знаний линейной алгебры, основ теории чисел, комбинаторики и теории вероятностей. Впрочем, все необходимые для понимания основного материала определения и утверждения приведены в тексте. Практически все главы и разделы сопровождаются упражнениями, решение которых позволит читателю глубже понять теорию.

Пособие предназначено для студентов математических факультетов и факультетов информационных технологий университетов, а также может быть полезно студентам физических факультетов, интересующимся математическими основами проблем передачи данных по каналам связи с помехами.

Следует отметить, что выбор материала, представленного в настоящем пособии, отвечает в некотором смысле вкусам автора. В частности, больше внимания уделено различным методам построения известных кодов, нежели процедурам декодирования. Автор со всей ответственностью осознает это и в дальнейших переизданиях данного пособия возможно восполнение этих пробелов. Кроме того, следует отметить, что имеются также другие классические интересные и красивые темы по теории кодирования, не затронутые в пособии, такие, как преобразование Адамара, теоремы Мак-Вильямс, теория равновесных кодов, глубокие связи теории кодирования с теорией блок-схем, теорией групп, теорией графов, а также не рассмотрены схемы отношений, сверточные коды и др.

Изложенный материал опробирован при чтении лекций в течение ряда лет в Новосибирском государственном университете. Автор выражает признательность всем студентам, которые помогали шлифовать в дискуссиях презентации многих тем, теорем, лемм, а также своим детям Ване Могильных и Маше Соловьевой за постоянную моральную поддержку в период написания этого пособия и помощь в напечатании лекций. Глубокая благодарность моим коллегам Валентину Афанасьеву, Алексею Пережогину, Владимиру Потапову, Денису Кротову за тщательное чтение пособия и полезные замечания, позволившие улучшить текст; моим рецензентам В. В. Зяблову, С. А. Малюгину, Ю. Л. Сагаловичу; моему ученику аспиранту Института математики СО РАН Антону Лосю за помощь в напечатании текста, сопровождение текста красивыми рисунками, считывание глав, оригинальную обложку; Н. Н. Токаревой, оказавшей всестороннюю техническую помощь при подготовке настоящего пособия — в формировании буклета, в печатании текста, подготовке рисунков, основательной и творческой считке текста, в написании раздела 9.3, посвященного кодам Препараты; О. Г. Заварзиной за подготовку верстки к печати.

Библиографический список основной литературы

[1] *Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А.* Теория кодов, исправляющих ошибки: Пер. с англ. М.: Связь, 1979. 744 с.

[2] *Берлекэмп Е. Р.* Алгебраическая теория кодирования: Пер. с англ. М.: Мир, 1971. 477 с.

[3] *Блейхут Р.* Теория и практика кодов, контролирующих ошибки: Пер. с англ. М.: Мир, 1986. 576 с.

[4] *Галлагер Р. Г.* Теория информации и надежная связь: М.: Сов. радио, 1974.

[5] *Касами Т., Токура Н., Ивадари Е., Инагаки Я.* Теория кодирования: Пер. с яп. М.: Мир, 1978. 576 с.

[6] *Колесник В. Д., Полтырев Г. Ш.* Курс теории информации. М.: Наука, 1982. 416 с.

[7] *Конвей Дж. Н. , Слоэн Н. Дж. А.* Упаковки шаров, решетки и группы: Пер. с англ. М.: Мир, 1990. Т. 1, 2.

[8] *Питерсон У., Уэлдон Э.* Коды, исправляющие ошибки: Пер. с англ. М.: Мир, 1976. 594 с.

[9] *Фано Р. М.* Передача информации. Статистическая теория связи. М.: Мир, 1965.

[10] *Шеннон К. Е.* Работы по теории информации и кибернетике. М.: Иностр. лит., 1963.

[11] *Шоломов Л. А.* Основы теории дискретных логических и вычислительных устройств. М.: Наука, 1980. 399 с.

[12] Handbook on coding theory, Amsterdam: North-Holland, 1998.

[13] *Solov'eva F. I.* "On perfect codes and related topics" Com^{2}Mac Lecture Note Series 13, Pohang 2004. 80 p. (доступна по адресу http://www.codingtheory.gorodok.net/literature/lecture-notes.pdf)

Библиографический список дополнительной литературы

[14] *Аршинов М. Н., Садовский Л. Е.* Коды и математика (рассказы о кодировании). М.: Наука, 1983. 144 с.

[15] *Васильев Ю. Л.* О негрупповых плотно упакованных кодах. Проблемы кибернетики. М: Физматгиз, 1962. Вып. 8. С. 337–339.

[16] *Гоппа В. Д.* Введение в алгебраическую теорию информации. М.: Наука: Физматлит, 1995.

[17] *Calderbank A. R.* The Art of Signaling: Fifty Years of Coding Theory. IEEE Trans. Inform. Theory. 1998. Vol. 44, № 6. P. 2561–2595. (доступна по адресу http://citeseer.ist.psu.edu/calderbank98art.html)

[18] *Камерон П., Ван Линт Дж. Х.* Графы. Коды и схемы: Пер. с англ. М.: Наука, 1980. 140 с.

[19] *Cohen G., Honkala I., Litsyn S., Lobstein A.* Covering codes. Netherlands: Elsevier, 1997. 542 p.

[20] *Колмогоров А. Н.* Теория информации и теория алгоритмов. М.: Наука, 1987. 304 с.

[21] *van Lint J. H.* Introduction to Coding Theory. Springer; Verlag; Berlin; Heidelberg, 1999.

[22] *Марков А. А.* Введение в теорию кодирования. М.: Наука: Физматлит, 1982.

[23] *Реньи А.* Дневник: Записки студента по теории информации: Пер. с венг. М.: Мир, 1980.

[24] *van Tillborg H.* Error correcting codes – a first course. Sweden, Chartwell Bratt Ltd., 1997.

[25] *Яглом А. М., Яглом И. М.* Вероятность и информация. М.: Наука, 1973.

[26] Сайт "Теория кодирования в Новосибирском государственном университете" по адресу http://www.codingtheory.gorodok.net.

[27] Сайт "Math Tree" : каталог математических интернет ресурсов (раздел "Теория кодирования") по адресу http://www.mathtree.ru.

Предметный указатель

Оглавление

Ф. И. Соловьева

ВВЕДЕНИЕ В ТЕОРИЮ КОДИРОВАНИЯ

Учебное пособие

Редактор С. Д. Андреева

Подписано в печать 14.06.2006 г.
Формат 84×120 / 8. Офсетная печать.
Уч.-изд. л. 17,7. Усл. печ. л. 14,4. Тираж 200 экз.

Заказ № .

Редакционно-издательский центр НГУ. 630090, Новосибирск-90, ул. Пирогова, 2.