Введение в алгебраические коды

Сагалович Ю.Л.

4 октября 2011 г.

Предисловие

Содержание этой книги составляет годовой курс "Алгебраические коды", который автор читал в течение ряда лет в Московском физико-техническом институте (государственном университете). Разумеется, за 120-130 академических часов, включая и семинарские занятия, можно сообщить студентам лишь ничтожную долю тех сведений, которые накоплены за полвека развития этой замечательной ветви теории информации. И находясь под влиянием богатства результатов, красоты и изящества теории алгебраических кодов, автор всегда испытывал чувство досады по поводу рамок учебной программы, ограничивающей желание лектора внушить своим слушателям во всей полноте значимость и стройность алгебраической теории кодирования.

Тем более автор был поставлен в тупик, когда ему порекомендовали написать для студентов учебное пособие под названием "Алгебраические коды". Можно ли согласиться на такой шаг, когда мировое сообщество специалистов и исследователей имеет в своем распоряжении книги У.У. Питерсона; В.Д. Колесника и Е.Т. Мирончикова; Э. Берлекэмпа; Э.Л. Блоха и В.В. Зяблова; У.У. Питерсона и Уэлдона; Т. Касами, Н. Токура, Ё. Ивадари и Я. Инагаки; Ф. Дж. Мак-Вильямс и Н.Дж.А. Слоэна; Э.М. Габидулина и В.Б. Афанасьева; Р. Блэйхута; Лидла и Нидеррейтера.

Однако по здравом размышлении автор уяснил два главных факта. Книги указанных выше авторов давно стали библиографической редкостью, и надежды на их переиздание нет. Сами же эти руководства настолько насыщены и полны, а подчас не просто полны, но полны энциклопедически, что для первого чтения трудны. Процесс преподавания, его временны́е ограничения, уровень подготовки студентов способствовали естественному отбору содержания курса. Естественный отбор, однако, был не столь беспощаден, как в животном мире. Всякое вынужденное отсечение безусловно интересных сведений, ко-

торые буквально просились на страницы книжки, было поистине болезненным. Когда логика рассуждений требовала продолжить изложение темы, но заданный объем курса, а значит и книжки, ставил понятное препятствие, приходилось идти на компромисс. Искушенный читатель без труда обнаружит соответствующие места.

В то же время в данную книжку включена специальная глава "Начальные сведения из теории чисел" , хотя в перечисленных выше источниках теоретико-числовые сведения использованы, но разрозненно.

Дело в том, что ни в один курс математики технических вузов не включается даже упоминания о теории чисел. Если в монографии можно ограничиться лишь ссылкой на источник, где можно прочитать, скажем, о теореме Эйлера, то в учебнике это было бы неуважением к читателю. Теория чисел настолько близка к алгебре, что в книге об алгебраических кодах она выглядит вполне органичной, так как является одним из исторических корней современной алгебры. Теоретико-числовые факты служат яркой иллюстрацией к большинству утверждений теории групп, а без нее теория кодов вовсе и не теория. Возможность доказать, например, (малую) теорему Ферма или теорему Вильсона различными способами, в зависимости от того, в каком месте книжки находится читатель, демонстрирует тесную связь между различными разделами курса. Наконец, читатель, решивший освоить основное содержание теории алгебраических кодов, вполне способен и достоин чести познакомиться с основами "царицы математики" как раздела, цементирующего математическую культуру.

Возвращаясь к теме объема, приведем, пожалуй, самый главный аргумент против расширения содержания книжки и доведения ее в конце концов до уровня энциклопедии по кодам. Такой уровень был выдающимся достижением тридцать лет тому назад, когда вышла в свет книга Ф. Дж. Мак-Вильямс и Н. Дж. А. Слоэна. Теперь другое время, и когда накоплен огромный запас методов построения кодов, удовлетворяющих разнообразным практическим нуждам, когда необычайным образом разрастается тематика исследований по теории кодирования, когда специалистам, овладевшим началами теории кодов, предоставлен широкий ассортимент способов построения реальных систем связи, назревает другая необходимость. Это — необходимость подняться в преподавании теории кодирования на новый уровень абстракции. Возможность к такому переходу обеспечена появлением монографии "Алгеброгеометрические коды (Основные понятия)" трех авторов: С.Г. Вледуца, Д.Ю. Ногина и М.А. Цфасмана. На мой взгляд, сформирована

и аудитория молодых людей с хорошей математической подготовкой, способная к восприятию новых серьезных сведений.

Данная же книжка преследует куда более скромную цель: помочь читателю освоить основы теории алгебраических кодов и более или менее осмысленно ориентироваться в обширной литературе по теории кодирования.

Автор искренне благодарит В.Б. Афанасьева и В.В. Зяблова за труд первого чтения, сопровождавшегося рядом важных советов и замечаний, которые послужили улучшению предлагаемой книги, Д.С. Осипова и Д.В. Лаконцева за помощь в оформлении оригинала-макета, В.С. Козякина за неоценимое содействие и помощь.

Предисловие ко второму изданию

В настоящем издании прежде всего устранены недочёты предыдущего издания; улучшено изложение в нескольких местах книги; внесены незначительные изменения в списки задач к главам; существенно расширена глава с указаниями к решению задач. Внесены дополнения в главу о конечных полях В теоретико-числовую главу помещен раздел об алгоритме Эвклида отыскания наибольшего общего делителя двух целых чисел. Это вызвано тем, что в разделах 7.7 — 7.10, написанных при участии В.Б.Афанасьева, подробно изложен метод декодирования, основанный на алгоритме Эвклида. Среди других методов этому методу отдано предпочтение по той причине, что алгоритм Эвклида, известный с незапямятных времён, через столетия оказался как будто специально приспособленным для весьма специфических нужд теории кодирования.

Введение

0.1. Система передачи информации. Двоичный симметричный канал

В простейшем случае система передачи сообщений состоит из передатчика, канала связи, и приемника. Схематически он выглядит, как показано на рис.1

Рис. 1. Простейшая система передачи информации

Передатчик передает сообщения в виде последовательностей одинаковой длины n, состоящих из нулей и единиц:

$$u = (u_1, u_2, \ldots, u_n),\ u_i = 0, 1. \tag{0.1.1}$$

(Вообще говоря, последовательности могут иметь и разные длины, но здесь этот случай рассматриваться не будет.)

В канале действует случайная помеха: каждый символ передаваемой последовательности независимо от других может быть искажен с вероятностью

$$\tau < 1/2. \tag{0.1.2}$$

Это означает, что с вероятностью $\tau < 1/2$ происходят переходы $0 \to 1$, $1 \to 0$ и с вероятностью $\varsigma = 1 - \tau$ компоненты 0 и 1 последовательности u сохраняют свое значение. Такой канал называется двоичным симметричным. Схематически он изображен на рис. 2.

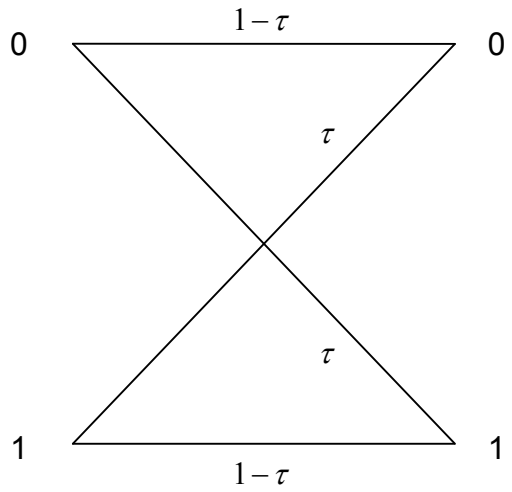

Рис. 2. Двоичный симметричный канал

Часто вместо "последовательности" применяют термины "слово" или "вектор". Для нас это синонимы, но для определенности и единообразия в дальнейшем употребляется термин "вектор".

Процесс действия помехи на передаваемый вектор можно описать следующим образом. Введем в рассмотрение вектор

$$e = (e_0, e_1, \ldots, e_{n-1}), \tag{0.1.3}$$

где $e_i = 1$, если символ u_i искажен, и $e_i = 0$ — в противном случае. Вектор (0.1.3) называется вектором-ошибкой.

П р и м е р 0. 1.

Пусть передавался вектор $u = (110001001)$, а на приемном конце получен вектор $v = (010001101)$. Это означает, что в канале подверглись искажению первый и седьмой символы передававшегося вектора. Это означает, что в векторе e единицы расположены на первом и седьмом местах, т.е. $e = (100000100)$.

Легко заметить, что вектор v получается поразрядным сложением векторов u и e по модулю два: 0+0=1+1=0, 0+1=1+0=1.

Вообще, если передается вектор (0.1.1), то на приёмном конце получают вектор

$$v = u + e = (u_1 + e_1, u_2 + e_2, \ldots, u_n + e_n).$$

Про такой канал говорят, что это канал с "аддитивной помехой".

В подавляющем большинстве случаев, когда непосвященному впервые задается вопрос, как поступить, если существует вероятность искажения передаваемого символа, следует ответ: "повторить передачу" , а значит, правильным считать тот символ, который встречается чаще. И чем больше повторять передачу, тем надежнее она будет. Такая уверенность основана на том, что в силу условия (0.1.2) правильных значений передаваемого символа окажется больше, чем неправильных.

Ясно однако, что хотя с ростом числа n повторений одного и того же передаваемого символа растет и степень уверенности в правильности определения истинного значения символа по правилу большинства, одновременно с этим падает скорость передачи, ибо на передачу одного символа тратится все больше времени. Нетрудно было бы подкрепить это утверждение несложными выкладками, однако в этом нет нужды, так как не его доказательству посвящается данное руководство, и интуитивного представления о процессе многократного повторения вполне достаточно.

Можно с уверенностью утверждать, что не будь разумной альтернативы методу многократного повторения передаваемого символа, наука, которая называется "Алгебраические коды" , не была бы известна.

Положим в общем случае, что ради надежной передачи информации, вместо k символов приходится передавать $n > k$ символов. Сопоставление некоторым способом n символов заданным k символам вектора называют *кодированием.* Отношение k/n называется *скоростью передачи.* В предыдущем случае $k = 1$, и скорость передачи $k/n = 1/n$.

Восстановление закодированных k символов по n переданным символам, каждый из которых мог подвергнуться искажению с вероятностью (0.1.2), называется *декодированием.* Не исключена возможность, что это декодирование окажется ошибочным.

Теперь систему передачи информации можно детализировать более подробно, как показано на рис. 3. Из источника сообщений извлекается вектор длины k, кодер преобразует его в

вектор длины n, который передается по каналу связи с помехой, а затем декодируется и поступает по назначению.

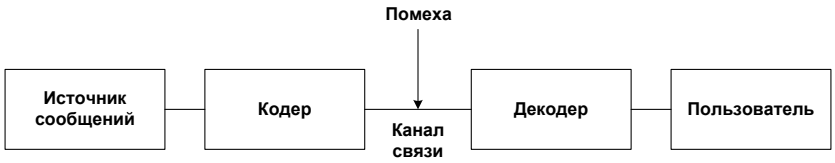

Рис. 3. Система передачи информации

Альтернативу методу многократного повторения доставляет знаменитая теорема Шеннона.

Она гласит, что *существует кодирование, позволяющее добиться сколь угодно достоверной передачи сообщения, лишь бы скорость передачи не превосходила некоторой величины, которая называется пропускной способностью канала.*

Достоверность передачи измеряется вероятностью P ошибки декодирования.

Основную роль в точном формулировании теоремы Шеннона играет функция энтропии $H(x)$. Она имеет вид

$$H(x) = -x \log x - (1-x) \log(1-x).$$

Если не оговорено иное, логарифмы берутся при основании 2. Натуральный и десятичный логарифмы, как обычно, пишутся соответственно $\ln$ и $\lg$. Если же логарифмы берутся при некотором основании q, то пользуются обозначением $H_q(x)$. Последнее имеет место в случае так называемого q-ичного канала, который будет обсуждаться в одной из следующих глав.

С помощью функции энтропии определяется упомянутая пропускная способность $C(\tau)$ канала, представленного на рис. 2.

$$C(\tau) = 1 - H(\tau).$$

Она зависит только от величины τ – вероятности искажения одного двоичного символа.

В этих терминах теорема Шеннона, которая первоначально доказывалась на случай двоичного симметричного канала, звучит следующим образом:

Каково бы ни было $\epsilon > 0$, *при достаточно большом* n, *и* $k/n < C(\tau) = 1 - H(\tau)$, *существует такое кодирование, при котором* $P < \epsilon$.

Теорема Шеннона доказана методом случайного выбора векторов и не говорит, как реализовать такое кодирование. Однако, если доказано существование "хорошего" кодирования, то это прямой сигнал к поиску конструктивных методов создания кодов с заданными свойствами. Значительная часть книги посвящена именно этой теме.

0.2. Кодовое расстояние

Важнейшую роль в теории кодирования играет понятие *кодового расстояния*. Под кодовым расстоянием чаще всего подразумевают так называемое расстояние Хэмминга. Расстоянием $d(a,b)$ между двумя векторами $a = (a_0, a_1, \ldots, a_{n-1})$ и $b = (b_0, b_1, \ldots, b_{n-1})$ называется число таких компонент i двух векторов, в которых последние не совпадают, т.е., в которых $a_i \neq b_i$.

Нетрудно показать, что так определенное расстояние обладает всеми свойствами метрики. Действительно, очевидно, что расстояние симметрично: $d(a,b) = d(b,a)$. Выполняется также неравенство треугольника. В самом деле, расположим три вектора a, b, c в виде таблицы, обозначив в ней числами $m_1 \geq 0, m_2 \geq 0, m_3 \geq 0, m_4 \geq 0, m_5 \geq 0, m_6 \geq 0$ количество столбцов соответственно[1]

	m_1	m_2	m_3	m_4	m_5	m_6
a	0	0	0	1	1	1
b	0	1	1	0	0	1
c	1	0	1	0	1	0

.

Тогда

$$\begin{aligned} d(a,b) &= m_2 + m_3 + m_4 + m_5, \\ d(a,c) &= m_1 + m_3 + m_4 + m_6, \\ d(b,c) &= m_1 + m_2 + m_5 + m_6. \end{aligned}$$

Отсюда получаем $d(a,b) + d(a,c) - d(b,c) = 2(m_3 + m_4) \geq 0$, что и требовалось.

Пусть в нашем распоряжении имеется некоторое множество A двоичных векторов длины n, и число $|A|$ векторов этого множества удовлетворяет условию $|A| = M$. Назовем это множество *кодом*. Положим $a, b \in A$, $d = \min_{a \neq b} d(a,b)$, и минимум

[1] Столбцы $(111)^T$, $(000)^T$ отсутствуют, так как никакого вклада в расстояние не вносят.

берется по всем C_M^2 парам векторов множества A. Эта величина носит название "минимальное расстояние кода" или "кодовое расстояние". Оба термина — синонимы.

Пусть минимальное расстояние кода есть $d = 2t + 1$. Если при передаче вектора a было искажено t, или менее символов, то принятый вектор $a^{'}$ будет $a^{'} = a + e$, где вектор-ошибка e содержит t или менее единиц. Тогда для произвольного вектора $x \in A$, $x \neq a$, будет справедливо неравенство $d(a^{'}, a) < d(a^{'}, x)$. Оно означает, что принятый вектор $a^{'}$ ближе к передававшемуся вектору a, чем к любому другому вектору $x \in A$, $x \neq a$.

Так как в силу (0.1.2) вероятность того, что произошло меньшее число ошибок, больше, чем вероятность того, что произошло большее число ошибок [2], то декодирование принятого вектора $a^{'}$ в вектор a будет правильным. Такой способ декодирования является частным случаем декодирования "по максимуму правдоподобия", который независимо от способа его реализации в своей расширительной трактовке преследует цель минимизировать ошибку декодирования.

В этом месте и проявляется теорема Шеннона.

"Правильное декодирование" есть синоним выражения "исправлена ошибка". Сказанное означает, что справедливо

Утверждение 0.2.1. *При*

$$d \geq 2t + 1 \tag{0.2.4}$$

код A исправляет все независимые ошибки кратности t и менее.

Пусть цель исправления ошибок не преследуется, а требуется только установить факт наличия ошибок. Ясно, что при $d = 2t + 1$ никакие независимые ошибки кратности $2t$ и менее не могут преобразовать передававшийся вектор a ни в один из векторов $x \in A$, $x \neq a$. Установление факта наличия ошибок называют "обнаружением" ошибок или "отказом от декодирования".

Таким образом, имеет место

Утверждение 0.2.2. *При $d \geq 2t + 1$ код A исправляет все независимые ошибки кратности t и менее, или обнаруживает все независимые ошибки кратности $2t$ и менее.*

[2] Нетрудно провести соответствующий расчет, однако читатель сделает это самостоятельно

Подчеркнем, что союз "или" здесь разделительный!

Пусть, наконец, кодовое расстояние четно: $d = 2t + 2$. Разумеется, по-прежнему исправляются все независимые ошибки кратности t и менее. Но если произойдет $t+1$ ошибок, то исправить их невозможно, зато можно обнаружить, так как они не смогут преобразовать ни один вектор $a \in A$ ни в один вектор $x \in A$.

Иначе говоря, справедливо

Утверждение 0.2.3. *При $d \geq 2t + 2$ код A исправляет все независимые ошибки кратности t и менее, и одновременно обнаруживает все независимые ошибки кратности $t + 1$.*

Изложенному здесь можно придать геометрическую форму. Рассмотрим рис. 4 и 5.

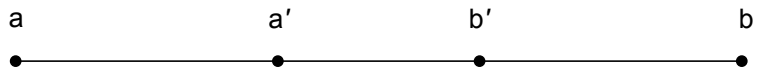

Рис. 4. Случай расстояния $d = 3$

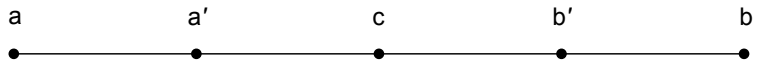

Рис. 5. Случай расстояния $d = 4$

Если был передан вектор a (или b), и произошла одна ошибка, то будет принят вектор $a^{'}$ или $((b^{'}))$. В случае расстояния $d = 3$ принятый вектор $a^{'}$ будет действительно ближе к истинно переданному a, а принятый вектор $b^{'}$ будет действительно ближе к истинно переданному b. Декодирование будет безошибочным. Если произойдут две ошибки, то при расстоянии $d = 3$ принятый вектор $a^{'}$ совпадет с вектором b, а принятый вектор $b^{'}$ совпадет с вектором a. Декодирование будет ошибочным.

В случае расстояния $d = 4$ при двух ошибках оба передававшихся вектора a и b превратятся в вектор c, который одинаково удален от a и b, и не будет декодирован ни в один из них. Обнаружен факт наличия ошибок. В принятой терминологии — "обнаружена ошибка".

0.3. Скорость передачи и минимальное расстояние

Интуитивно ясно, что создание расстояния между кодовыми векторами кода A *с необходимостью* влечет увеличение длины вектора, и как отмечалось выше, ведет к уменьшению скорости передачи k/n. Поэтому интересно выяснить, каков закон, связывающий скорость передачи кода с кодовым расстоянием. Будем выводить этот закон так называемым методом исчерпания, одновременно строя сам код A, содержащий M векторов длины n, находящихся на расстоянии не менее чем d друг от друга. Заметим при этом, что $k =]\log_2 M[$.

В качестве первого вектора u_1 выберем произвольный вектор из 2^n всех векторов длины n. Затем найдем все векторы, которые находятся на расстоянии $1, 2, \ldots, d-1$ от него. Вместе с вектором u_1 их будет всего

$$\sum_{i=0}^{d-1} C_n^i.$$

Удалим эти векторы из нашего арсенала выбора. Оставшиеся

$$2^n - \sum_{i=0}^{d-1} C_n^i$$

векторов находятся от вектора u_1 на расстоянии, не меньшем, чем d. Поэтому любой из них может быть выбран как вектор u_2 нашего кода. И в этом случае найдем все векторы, которые находятся на расстоянии $1, 2, \ldots, d-1$ теперь уже от вектора u_2. По построению они находятся на расстоянии, не меньшем, чем d и от вектора u_1.

Удалим и эти векторы из нашего арсенала выбора, не обращая внимания на то, что некоторые из них были удалены ранее. Оставшиеся не менее чем

$$2^n - 2\sum_{i=0}^{d-1} C_n^i$$

векторов находятся от векторов u_1 и u_2 на расстоянии, не меньшем, чем d. Любой из них может быть выбран как вектор u_3.

Поступая с ним так же, как и с предыдущими, а затем выбирая последовательно векторы кода до вектора u_{M-1}, удалим всего

$$(M-1)\sum_{i=0}^{d-1} C_n^i$$

векторов. Если количество оставшихся векторов будет удовлетворять неравенству

$$2^n - (M-1)\sum_{i=0}^{d-1} C_n^i > 0, \tag{0.3.5}$$

то можно выбрать еще один вектор, который находится на расстоянии, не меньшем, чем d от всех предыдущих, а это означает, что код A с параметрами n, d, M заведомо существует.

Заметим, что множество $\sum_{i=0}^{d-1} C_n^i$ векторов, отстоящих на расстоянии не менее, чем $d-1$ от вектора u_i, вместе с самим вектором u_i называется шаром радиуса $d-1$, а вектор u_i — его центром.

Неравенству (0.3.5), как легко видеть, можно придать иную форму:

$$\frac{k}{n} < 1 - \frac{1}{n}\log\sum_{i=0}^{d-1} C_n^i. \tag{0.3.6}$$

Читается это неравенство так: если скорость передачи k/n кода удовлетворяет неравенству (0.3.6), то код длины n с расстоянием d существует. Неравенство (0.3.6) называется *границей Гилберта*. Оба последних неравенства выражают типичную границу существования. Способ ее вывода — исчерпание — основан на переборе всех 2^n векторов длины n, ничего общего не имеет с алгеброй, и ни в коем случае не может быть способом построения кода. Довольно грубое обращение с шарами, которые могут пересекаться, и потому один и тот же вектор может удаляться не один раз, создает впечатление о возможности улучшения границы существования. Разумеется, учитывая возможность многократного выбрасывания каких-нибудь векторов, в неравенстве (0.3.6) можно получить величину, меньшую, чем та, которая стоит под знаком суммы. Но когда говорят об улучшении границы, имеют в виду асимптотическую форму границы (см. 8.3). Так вот ее-то, как показала история, улучшить не удается.

Создание расстояния d между векторами играет важную роль не только в случае помех, свойственных двоичному симметричному каналу, показанному на рис. 2. На рис. 6 показан так называемый двоичный стирающий канал.

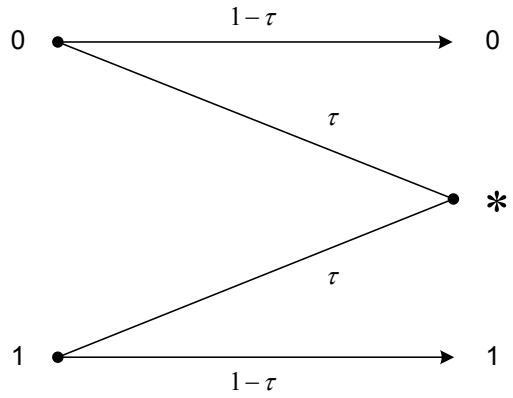

Рис. 6. Двоичный стирающий канал

Передаваемые символы принимаются неискаженными с вероятностью $1-\tau$, но помеха действует таким образом, что подвергшийся ей символ с вероятностью τ принимает третье значение x, отличное от 0 и от 1. Такая помеха называется стиранием. При этом номера позиций, где произошли стирания известны. В этом главное отличие стираний от ошибок, когда номера искаженных символов неизвестны.

Тривиальным методом исправления l стираний является подстановка на (известные!) стертые позиции всех возможных 2^l комбинаций из нулей и единиц и выбор из них единственной правильной комбинации, которая заведомо существует. Имеет место

Утверждение 0.3.1. *Для исправления любых l стираний необходимо и достаточно чтобы, минимальное расстояние кода удовлетворяло условию*

$$d \geq l + 1. \tag{0.3.7}$$

Доказательство. Пусть выполняется условие $d \geq l+1$. Заменим все t стертых символов всеми возможными 2^l способами символами 0 и 1. Тогда одна и только одна комбинация l нулей и единиц восстановит принятый вектор в вектор

переданный. Но любая другая ошибочная замена будет обнаружена, так как она будет находиться на расстоянии не более, чем $d-1$ от истинной. Наоборот, при $d < l+1$ найдется такая ошибочная замена, которая не будет обнаружена, и произойдет ошибка декодирования.

Можно воспользоваться таким рассуждением. Пусть принятый вектор a содержит $l \leq d-1$ стираний, и пусть код A имеет минимальное расстояние d. Такой код содержит в точности один вектор u, совпадающий с принятым вектором a в неискажённой его части. Действительно, если найдётся ещё один вектор v с таким свойством, то окажется, что $d(u,v) = l \leq d-1 < d$, что противоречит условию.

Рассмотрим случай, когда в канале при передаче могут возникнуть искажения обоих видов, т.е. одновременно, ошибки и стирания. Имеет место

Утверждение 0.3.2. *Для исправления любых t независимых ошибок и одновременно любых l (независимых) стираний необходимо и достаточно чтобы, минимальное расстояние кода удовлетворяло условию*

$$d \geq 2t + l + 1. \tag{0.3.8}$$

Доказательство. Пусть в принятом векторе a, кроме t ошибок, имеется l стёртых символов. Удалим из вектора a, равно как и из всех остальных векторов кода A все те компоненты, в которых размещены стирания вектора a. Получится новый код A' с параметрами $n' = n - l,\ d' \geq d - l$. Этот код исправляет любые независимые ошибки кратности $t \leq [(d-l-1)/2]$ и менее. Когда ошибки будут исправлены, мы возвратимся к коду A, которому принадлежит принятый вектор a содержащий теперь уже только стирания. Их мы умеем исправлять, применяя предыдущие соображения.

Сравним три знакомых нам соотношения (0.3.8),(0.2.4), (0.3.7) Второе получается из первого, если есть только ошибки, и нет стираний, а третье получается из первого, если есть только стирания, и нет ошибок.

Иными словами, если минимальное расстояние кода есть d, то при отсутствии стираний код исправляет любые $t \leq (d-1)/2$ независимых ошибок, а при отсутствии ошибок исправляет любые $l \leq d-1$ стираний. Наконец, при наличии не более чем l стираний код с расстоянием d исправляет любые $t \leq [(d-l-$

1)/2] или менее независимых ошибок. [3]

Исправление стираний будет обсуждаться также в гл. 6.

0.4. Код Хэмминга

Рассмотрим матрицу

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \tag{0.4.9}$$

Все семь столбцов матрицы различны, и из всех возможных восьми столбцов длины 3 отсутствует нулевой столбец.

Пусть код A состоит из всех таких векторов, скалярное произведение которых на каждый из трех векторов-строк

$$(0111100) = z_1, (1110010) = z_2, (1101001) = z_3$$

матрицы H равно нулю. Этому условию удовлетворяет, например, вектор $u = (1000011)$:

$$\begin{aligned} (1000011)z_1 &= 1\cdot 0 + 0\cdot 1 + 0\cdot 1 + 0\cdot 1 + 0\cdot 1 + 1\cdot 0 + 1\cdot 0 = 0, \\ (1000011)z_2 &= 1\cdot 1 + 0\cdot 1 + 0\cdot 1 + 0\cdot 0 + 0\cdot 0 + 1\cdot 1 + 1\cdot 0 = 0, \\ (1000011)z_3 &= 1\cdot 1 + 0\cdot 1 + 0\cdot 0 + 0\cdot 1 + 0\cdot 0 + 1\cdot 0 + 1\cdot 1 = 0. \end{aligned} \tag{0.4.10}$$

На самом деле (0.4.10) изображает операцию $uH^T = 0$, т.е. умножение вектора u на матрицу H^T, где индекс T означает транспонирование.

Положим, что при передаче вектора u данного примера произошла одиночная ошибка, например, в четвертом разряде вектора, и на приемном конце принят вектор $v = u + e = (1000011) + (0001000) = (1001011)$. Тогда, выполнив на приемном конце умножение принятого вектора на транспонированную матрицу H, получим

$$vH^T = uH^T + eH^T = eH^T = (101),$$

а это есть четвертый слева столбец матрицы H. Его номер указывает на то, что при передаче был искажен четвертый символ

[3] Когда ниже в тексте будет введён в обращение $q-$ичный, а не только двоичный, канал, читатель поймёт, что предыдущие рассуждения об исправлении ошибок и стираний, а также обнаружении ошибок, останутся справедливыми.

кодового вектора. Вектор-ошибка при умножении принятого вектора v на транспонированную матрицу H "вырезал" именно тот столбец матрицы, номер которого указывает на номер искаженного символа кодового вектора.

Изменим теперь слегка матрицу (0.4.9), расположив ее столбцы в лексикографическом порядке.

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}. \qquad (0.4.11)$$

Примечательной чертой такого нового расположения столбцов является то, что при любой одиночной ошибке в принятом векторе v результатом умножения vH^T будет в точности двоичный номер искаженного разряда кодового вектора. В общем случае матрица H содержит m строк и $n = 2^m - 1$ ненулевых столбцов, которые все различны. Код A, все векторы которого, умноженные скалярно на строки матрицы H, дают нуль, называется *кодом Хэмминга.* Он содержит 2^{2^m-1-m} кодовых векторов, его кодовое расстояние равно трем, он исправляет все одиночные ошибки, или обнаруживает все двойные независимые ошибки.

Подробнее этот код будет обсуждаться в главе 4.

Для построения кодов с более высокими корректирующими возможностями придется обратиться к накоплению новых алгебраических средств.

Предварительно зададимся вопросом: нельзя ли, все-таки, для повышения надежности передачи сообщений обойтись иными средствами, а не алгебраическими? Быть может, стоит всегонавсего увеличить энергию сигнала — и помеха будет подавлена. Порой это сделать нельзя из-за дефицита энергии, но даже и при избытке таковой обойтись без средств кодирования, в том числе и алгебраического, бывает невозможно. Например, при использовании спутниковой связи, когда неизбежна концентрация нескольких каналов связи в одной точке, усиление энергии сигнала может привести к взаимному влиянию каналов и наведению помех одним каналом в другом.

0.5. Задачи к введению

0.1. Доказать утверждение: для того, чтобы код исправлял любые комбинации t или менее ошибок, и одновременно обнаруживал все комбинации $t' \geq t$ или менее ошибок, необходимо и

достаточно, чтобы кодовое расстояние d удовлетворяло условию $d \geq t + t'$. 0.2. Доказать, что код Хэмминга исправляет любые два или менее стираний.

Глава 1.

Начальные сведения из теории чисел

1.1. Предварительные замечания

В этой главе представлен даже не минимально достаточный, а минимально необходимый набор элементарных теоретико-числовых фактов. Главная цель — наряду с изяществом доказательств основополагающих теорем теории чисел накопить иллюстративный материал к теории групп, колец и полей. Опущено почти все, что относится к теории делимости, решению сравнений разных степеней, систем сравнений, теории квадратичных вычетов, символов Лежандра и Якоби и т.д.

Опыт показал однако, что безусловная красота теории чисел увлекает молодых людей и побуждает их к расширению теоретико-числовых знаний. На первый случай им можно порекомендовать такие источники, как "Основы теории чисел" И.М. Виноградова и "Теория чисел" А.А. Бухштаба.

Итак, считаются известными понятия:

наибольшего общего делителя

$$(a, b) = d,$$

и наименьшего общего кратного

$$M(a, b)$$

двух чисел a и b. Между ними имеется связь

$$M(a, b) = ab/(a, b).$$

Каноническое разложение числа a на сомножители имеет вид

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k},$$

где p_i – простое число, α_i — натуральное, и это разложение единственно.

Простых чисел бесконечно много. Действительно, если найдены все первые k простых чисел

$$p_1, p_2, \ldots, p_k \tag{1.1.1}$$

до p_k включительно, то число

$$p_1 p_2 \ldots p_k + 1$$

либо само будет простым, либо любой простой его делитель, деля всю сумму, не будет совпадать ни с одним из чисел (1.1.1).

1.2. Наибольший общий делитель. Алгоритм Эвклида

Разумеется, для определения общего наибольшего делителя

$$(a, b) = d,$$

двух целых чисел a и $b, a > b$ можно воспользоваться их каноническими разложениями и выделить их максимальную общую часть. Однако для этой цели существует замечательное средство — алгоритм Эвклида. В теории чисел и алгебре алгоритм Эвклида имеет дело только с вычислением наибольшего общего делителя $d = (a, b)$ двух чисел a и b. Именно, в чистом виде

$$\begin{array}{ll} & 0 < b < a \\ a = bq_0 + r_0, & 0 < r_0 < b \\ b = r_0 q_1 + r_1, & 0 < r_1 < r_0 \\ r_0 = r_1 q_2 + r_2, & 0 < r_2 < r_1 \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ r_k = r_{k+1} q_{k+2} + r_{k+2}, & 0 < r_{k+2} < r_{k+1} \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ r_{n-2} = r_{n-1} q_n & 0 = r_n < r_{n-1} \end{array} \tag{1.2.2}$$

Последнее равенство, где $r_n = 0$, неизбежно ввиду уменьшения остатков на каждом следующем шаге деления. Из $a = bq_0 + r_0$ следует, что совокупность общих делителей чисел a и b совпадает с совокупностью общих делителей чисел b и r_0. Это означает, что $(a, b) = (b, r_0)$. На том же основании $((b, r_0) = (r_0, r_1))$.

Отсюда последовательно

$$(a, b) = (b, r_0) = (r_0, r_1) = \ldots = (r_{n-1}, 0) = r_{n-1}. \tag{1.2.3}$$

Однако посредством алгоритма Эвклида можно решить одну важную задачу: найти такие целые числа s и t, чтобы выполнялось равенство

$$as + bt = d, \tag{1.2.4}$$

где $d = (a, b)$). Оно получается следующим образом. Определим начальные условия

$$\begin{array}{l} u_{-2} = 0,\ u_{-1} = 1. \\ v_{-2} = 1,\ v_{-1} = 0 \end{array} \tag{1.2.5}$$

и будем вычислять q_k, r_k, u_k и v_k по формулам

$$\begin{array}{l} r_{k-2} = q_k r_{k-1} + r_k \\ u_k = q_k u_{k-1} + u_{k-2} \\ v_k = q_k v_{k-1} + v_{k-2}. \end{array} \tag{1.2.6}$$

Вычисления обрываются при $k = n$, так как $r_n = 0$.

Непосредственно проверяется, что

$$\begin{array}{l} v_k r_{k+1} + v_{k+1} r_k = v_k(-q_{k+1} r_k + r_{k-1}) + \\ +(q_{k+1} v_k + v_{k-1}) r_k = v_{k-1} r_k + v_k r_{k-1}, \\ u_k r_{k+1} + u_{k+1} r_k = u_k(-q_{k+1} r_k + r_{k-1}) + \\ +(q_{k+1} u_k + u_{k-1}) r_k = u_{k-1} r_k + u_k r_{k-1}, \end{array} \tag{1.2.7}$$

и что

$$\begin{array}{l} v_{k+1} u_k - u_{k+1} v_k = (q_{k+1} v_k + v_{k-1}) u_k - \\ -(q_{k+1} u_k + u_{k-1}) v_k = -(v_k u_{k-1} - u_k v_{k-1}). \end{array} \tag{1.2.8}$$

Нетрудно заметить, что в правых частях равенств (1.2.7) и (1.2.8) все нижние индексы на единицу меньше чем нижние индексы у одноименных величин в левых частях. Кроме того, в правой части равенства (1.2.8) знак изменился на обратный. Воспользуемся этим обстоятельством и из (1.2.7) получим последовательно

$$\begin{array}{l} v_k r_{k+1} + v_{k+1} r_k = v_{k-1} r_k + v_k r_{k-1} = \\ v_{k-2} r_{k-1} + v_{k-1} r_{k-2} = \ldots = v_{-2} r_{-1} + v_{-1} r_{-2} = r_{-1}, \end{array} \tag{1.2.9}$$

а также

$$\begin{array}{l} u_k r_{k+1} + u_{k+1} r_k = u_{k-1} r_k + u_k r_{k-1} = \\ u_{k-2} r_{k-1} + u_{k-1} r_{k-2}) = \ldots = u_{-2} r_{-1} + u_{-1} r_{-2} = r_{-2}, \end{array} \tag{1.2.10}$$

Наконец, из (1.2.8) получим

$$\begin{aligned} v_{k+1}u_k - u_{k+1}v_k = -(v_k u_{k-1} - u_k v_{k-1}) = \dots \\ \dots = v_{-1}u_{-2} - u_{-1}v_{-2} = (-1)^k \end{aligned} \tag{1.2.11}$$

.

Все эти равенства справедливы при всех $k \geq -1$, и правые их части получаются благодаря (1.2.5).

Таким образом,

$$\begin{aligned} &v_{k-1}r_k + v_k r_{k-1} = r_{-1}, \\ &u_{k-1}r_k + u_k r_{k-1} = r_{-2}, \\ &v_k u_{k-1} - u_k v_{k-1} = (-1)^k. \end{aligned} \tag{1.2.12}$$

Так как $r_n = 0$, последние три равенства дают

$$\begin{aligned} &v_n r_{n-1} = r_{-1}, \\ &u_n r_{n-1} = r_{-2}, \\ &r_{n-1}(v_n u_{n-1} - u_n v_{n-1}) = (-1)^n r_{n-1}. \end{aligned} \tag{1.2.13}$$

Умножив первое равенство (1.2.13) на u_{n-1}, а второе – на v_{n-1}, вычтем почленно из первого второе. Воспользовавшись третьим равенством (1.2.13), получим окончательно

$$r_{-1}u_{n-1} - r_{-2}v_{n-1} = (-1)^n r_{n-1} = (-1)^n (r_{-2}, r_{-1}). \tag{1.2.14}$$

Вспомним, что в наших обозначениях $r_{-2} = a,\ r_{-1} = b$. Отсюда и получается (1.2.4). Заметим, что равенство (1.2.4) можно получить, и не обращаясь к алгоритму Эвклида (см. задачу 1.6 к данной главе.)

Из равенства (1.2.11) немедленно получается, что $(u_k,\ v_k) = 1$.

Как увидит читатель, алгоритм Эвклида играет важную и интересную роль в проблеме декодирования.

1.3. Сравнения

В связи с данным положительным числом m все целые числа можно разбить на классы. К одному классу отнесём все целые числа, которые при делении на m дают один и тот же остаток.

Если числа a и b принадлежат к одному классу, то это записывается так

$$a \equiv b (\mathrm{mod} m). \tag{1.3.15}$$

Число m называется модулем. Числа, принадлежащие к одному классу, называются равноостаточными или сравнимыми по модулю m .

Утверждение 1.3.1. *Формула* (1.3.15) *равносильна следующим:*

$$a = b + mt, \quad a - b = mt. \tag{1.3.16}$$

Д о к а з а т е л ь с т в о. Пусть выполнено (1.3.15) Тогда

$a = mt_1 + r, b = mt_2 + r, 0 \leq < r < m$, откуда $a - b = m(t_1 - t_2) = mt$.

Наоборот, пусть выполняется (1.3.16). Тогда, если $b = mt_2 + r$, то $a = mt_2 + mt + r = m(t_2 + t) + r = mt_1 + r$, и $a \equiv b(\mathrm{mod} m)$

1.4. Свойства сравнений

Свойства сравнений доказываются в основном сведéнием сравнений к равенствам и наоборот на основании (1.3.15) и (1.3.16) в утверждении 1.3.1.

Утверждение 1.4.1. *Два числа, сравнимые с третьим, сравнимы между собой.*

Действительно, из $a \equiv b(\mathrm{mod} m)$ и $a \equiv c(\mathrm{mod} m)$ следует $a = b + mt_1, a = c + mt_2, \quad b + mt_1 = c + mt_2, \quad b - c = m(t_2 - t_1), \quad b = c + mt_3, \quad b \equiv c(\mathrm{mod} m)$.

Утверждение 1.4.2. *Сравнения можно почленно складывать.*

В самом деле,

$$a_1 \equiv b_1(\mathrm{mod} m), \ldots, a_k \equiv b_k(\mathrm{mod} m),$$

$$a_1 - b_1 = mt_1, \ldots, a_k - b_k = mt_k,$$

$$a_1 + \ldots + a_k = b_1 + \ldots + b_k + m(t_1 + \ldots + t_k),$$

$$a_1 + \ldots + a_k \equiv b_1 + \ldots + b_k(\mathrm{mod} m), \tag{1.4.17}$$

что и требовалось.

Утверждение 1.4.3. *Слагаемые, стоящие в какой-либо части сравнения, можно переносить в другую, меняя знак на обратный.*

Действительно, пользуясь утверждением 1.4.2, без ограничения общности, прибавим к сравнению (1.4.17) почленно тривиальное сравнение $-b_k \equiv -b_k(\mathrm{mod}m)$. Получим

$$a_1 + \ldots + a_k - b_k \equiv b_1 + \ldots + b_k - b_k(\mathrm{mod}m),$$

т.е.

$$a_1 + \ldots + a_k - b_k \equiv b_1 + \ldots + b_{k-1}(\mathrm{mod}m).$$

Утверждение 1.4.4. *К каждой части сравнения можно прибавить любое целое, кратное модуля.*

В самом деле, в силу утверждения 1.4.2 очевидное сравнение $mk_1 \equiv mk_2(\mathrm{mod}m)$ можно почленно прибавить к сравнению $a \equiv b(\mathrm{mod}m)$. В результате $a + mk_1 \equiv b + mk_2(\mathrm{mod}m)$, что и требовалось.

Не прибегая к формальным выкладкам, то же самое можно высказать так: прибавление к делимому числа, кратного делителя, не изменяет остатка.

Утверждение 1.4.5. *Сравнения можно почленно перемножать.*

Действительно, пусть $a_1 \equiv b_1(\mathrm{mod}m)$, $a_2 \equiv b_2(\mathrm{mod}m)$. Это означает, что $a_1 = b_1 + mt_1$, $a_2 = b_2 + mt_2$. Отсюда после перемножения равенств получим $a_1a_2 = b_1b_2 + mT$, т.е. $a_1a_2 \equiv b_1b_2(\mathrm{mod}m)$.

Утверждение 1.4.6. *Обе части сравнения можно возводить в одну и ту же степень.*

Это простое следствие из утверждения 1.4.5.

Обобщение утверждений 1.4.2, 1.4.5 и 1.4.6 выглядит следующим образом:

Утверждение 1.4.7. *Если в выражении целой рациональной функции*

$$S = \sum Ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_k^{\alpha_k}$$

заменить A на B, x_i на y_i, такие, что $A \equiv B(\mathrm{mod}\, m)$ $x_i \equiv y_i(\mathrm{mod}m)$, то получим:

$$\sum Ax_1^{\alpha_1}x_2^{\alpha_2}\ldots x_k^{\alpha_k} \equiv \sum By_1^{\alpha_1}y_2^{\alpha_2}\ldots y_k^{\alpha_k}(\mathrm{mod}m).$$

Аналогично, из

$$a_i \equiv b_i(\mathrm{mod} m), i = 0, 1, \ldots, n, \quad u \quad x \equiv y(\mathrm{mod} m)$$

получим

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_n \equiv b_0 y^n + b_1 y^{n-1} + \ldots + b_n(\mathrm{mod} m).$$

Утверждение 1.4.8. *Обе части сравнения можно разделить на их общий дедитель, если он взаимно прост с модулем.*

Пусть

$$a \equiv b(\mathrm{mod} m), \quad a = a_1 d, \quad b = b_1 d, \quad (m, d) = 1.$$

Тогда

$$a_1 d - b_1 d = mt, \quad (a_1 - b_1)d = mt.$$

Так как $(m, d) = 1$, то

$$a_1 - b_1 = mt', \quad a_1 - b_1 \equiv 0(\mathrm{mod} m), \quad a_1 \equiv b_1(\mathrm{mod} m).$$

Требование взаимной простоты использовано, и оно существенно. Например:

$$6 \equiv 18(\mathrm{mod} 4), \quad \text{но} \quad 3 \not\equiv 9(\mathrm{mod} 4).$$

Это были свойства сравнений, аналогичные свойствам равенств.

1.5. Дальнейшие свойства сравнений

Утверждение 1.5.1. *Обе части сравнения и модуль можно умножить на одно и то же число.*

Последовательно имеем

$a \equiv b(\mathrm{\ mod}\ m)$, $a - b = mt$, $m_1(a - b) = m_1 mt$, $m_1 a = m_1 b + m_1 mt$.

Это значит

$$m_1 a \equiv m_1 b(\mathrm{mod} m_1 m).$$

Утверждение 1.5.2. *Обе части сравнения и модуль можно разделить на их общий делитель.*

Если

$$a \equiv b(\mathrm{mod} m), \quad \text{и} \quad a = a_1 d, \quad b = b_1 d, \quad m = m_1 d,$$

то

$$a_1 d = b_1 d + m_1 dt, \quad a_1 = b_1 + m_1 t, \quad a_1 \equiv b_1(\mathrm{mod} m_1).$$

Утверждение 1.5.3. *Если сравнение имеет место по нескольким модулям $m_1, m_2, \ldots, m_k$, то оно имеет место и по модулю, который равен наименьшему общему кратному модулей: $M(m_1, m_2, \ldots, m_k)$.*

В самом деле, пусть

$$a \equiv b(\mathrm{mod} m_i), \quad i = 1, 2, \ldots, k; \quad \text{тогда} \quad a - b = m_i t.$$

Будучи кратным каждого из модулей, разность $a - b$, кратна и их наименьшего общего кратного.

П р и м е р 1. 1.

$$812 \equiv 1622(\mathrm{mod} 6, 9, 15), \quad 812 \equiv 1622(\mathrm{mod} 90)$$

Утверждение 1.5.4. *Если сравнение имеет место по модулю m, то оно имеет место и по модулю d, равному любому делителю числа m.*

Имеем последовательно

$$a \equiv b(\mathrm{mod} m), \quad a - b = mt, \quad a - b = m_1 dt = dt_1,$$

где

$$t_1 = m_1 t, \quad m = m_1 d.$$

Утверждение 1.5.5. *Если одна часть сравнения и модуль делятся на какое-нибудь число, то и другая часть сравнения делится на это число.*

Действительно, пусть

$$a \equiv b(\mathrm{mod} m); \quad \text{тогда} \quad a - b = mt.$$

Если

$a = a_1 z$ и $m = m_1 z$, то $a_1 z - m_1 zt = b$; $z(a_1 - m_1 t) = b = zb_1$.

Утверждение 1.5.6. *Если*

$$a \equiv b(\mathrm{mod} m), \quad \textit{то} \quad (a, m) = (b, m).$$

Доказательство. Согласно утверждению 1.5.5 совокупность общих делителей чисел a и m совпадает с совокупностью общих делителей чисел b и m. Значит, совпадают и наибольшие общие делители этих чисел, что и требуется.

1.6. Полная система вычетов

Числа, сравнимые по модулю m, образуют класс чисел по модулю m. Всего имеется m классов; любое число можно представить в виде

$$mq + r, \quad r = 0, 1, \ldots, m - 1.$$

Любое число класса называется вычетом по модулю m. Вычет, полученный при $q = 0$, равный самому остатку r, называется наименьшим неотрицательным вычетом. Вычет ρ, самый малый по абсолютной величине, называется абсолютно наименьшим вычетом. При $r < m/2 \quad \rho = r$; при $r > m/2 \quad \rho = r - m$;

Если m четное, и $r = m/2$, то $\rho = m/2$, или $\rho = -m/2$.

Взяв от каждого класса по одному вычету, получим полную систему вычетов по модулю m.

В качестве полной системы вычетов употребляют наименьшие неотрицательные вычеты:

$$0, 1, \ldots, m - 1,$$

или также абсолютно наименьшие вычеты

$$-\frac{m-1}{2}, \ldots, -1, 0, 1, \ldots, \frac{m-1}{2}$$

при нечётном m, и

$$-\frac{m}{2} + 1, \ldots, -1, 0, 1, \ldots, \frac{m}{2}$$

или

$$-\frac{m}{2}, \ldots, -1, 0, 1, \ldots, \frac{m}{2} - 1$$

при чётном m.

Утверждение 1.6.1. *Любые m чисел, попарно несравнимые по модулю m, образуют полную систему вычетов.*

Действительно, будучи несравнимыми, они принадлежат различным классам, а так как их m штук, то в каждый класс попадёт в точности по одному числу.

Теорема 1.6.2. *Если $(a, m) = 1$, и x пробегает полную систему вычетов по модулю m, то $ax + b$, где b произвольное целое, пробегает полную систему вычетов.*

Действительно, чисел $ax + b$ столько, сколько чисел x, т.е. m. Покажем, что $ax_1 + b$ и $ax_2 + b$ несравнимы, если несравнимы x_1 и x_2. Положим противное, т.е., что

$$ax_1 + b \equiv ax_2 + b(\mathrm{mod}m).$$

Тогда

$$ax_1 \equiv ax_2(\mathrm{mod}m),$$

и так как $(a, m) = 1$, то в силу утверждения 1.4.8

$$x_1 \equiv x_2(\mathrm{mod}m),$$

что противоречит условию.

П р и м е р 1. 2.

Пусть $m = 8$. $a = 11, b = 5$, $(11, 8) = 1$. Полная система вычетов есть 0,1,2,3,4,5,6,7;

$$11 \cdot 0 + 5 \equiv 5(\mathrm{mod}8),$$

$$11 \cdot 1 + 5 \equiv 0(\mathrm{mod}8),$$

$$11 \cdot 2 + 5 \equiv 3(\mathrm{mod}8),$$

$$11 \cdot 3 + 5 \equiv 6(\mathrm{mod}8),$$

$$11 \cdot 4 + 5 \equiv 1(\mathrm{mod}8),$$

$$11 \cdot 5 + 5 \equiv 4(\mathrm{mod}8),$$

$$11 \cdot 6 + 5 \equiv 7(\mathrm{mod}8),$$

$$11 \cdot 7 + 5 \equiv 2(\mathrm{mod}8).$$

Теорема 1.6.3. *Если* $(a, b) = 1$, *и* x *пробегает полную систему вычетов по модулю* b, *а* y *пробегает полную систему вычетов по модулю* a, *то* $ax + by + c$, *где* c *произвольное целое, пробегает полную систему вычетов по модулю* ab.

Доказательство. В условиях теоремы сумма $ax+by+c$ принимает в точности ab значений. Покажем, что они попарно несравнимы по модулю ab. Пусть наоборот, нашлись две таких пары x_1, y_1; x_2, y_2, что не выполняются сравнения $x_1 \equiv x_2$, $y_1 \equiv y_2$, но при этом $ax_1+by_1+c \equiv ax_2+by_2+c(\bmod ab)$. Отсюда имеем последовательно: на основании утверждения 1.4.3

$$ax_1 + by_1 \equiv ax_2 + by_2(\mathrm{mod} ab),$$

на основании утверждения 1.5.4

$$ax_1 + by_1 \equiv ax_2 + by_2(\mathrm{mod} b),$$

на основании утверждения 1.4.4

$$ax_1 \equiv ax_2(\mathrm{mod} b),$$

а так как $(a, b) = 1$, то

$$x_1 \equiv x_2(\mathrm{mod} b),$$

что противоречит условию. Аналогичные выкладки без труда производятся для y_1, y_2.

1.7. Приведённая система вычетов

Согласно утверждению 1.5.6, числа одного и того же класса по модулю m имеют с модулем один и тот же наибольший общий делитель. Особенно важны классы, для которых этот делитель равен единице, т.е. числа взаимно простые с модулем. Взяв от каждого класса по одному вычету, получим *приведённую* систему вычетов.

Обычно приведённую систему вычетов выбирают из наименьших неотрицательных вычетов

$$0, 1, \ldots, m - 1.$$

Обозначим через c количество чисел ряда $0, 1, \ldots, m - 1$, взаимно простых с m.

Обозначим $c = \varphi(m)$. Эта функция называется функцией Эйлера.

Имеют место следующие факты:

Утверждение 1.7.1. *Любые $\varphi(m)$ чисел, попарно несравнимые по модулю m и взаимно простые с ним, образуют приведённую систему вычетов.*

Действительно, будучи несравнимыми и взаимно простыми с модулем, они принадлежат различным классам чисел, взаимно простых с модулем, а так как их $\varphi(m)$ штук, то в каждый класс попадёт в точности по одному числу.

Теорема 1.7.2. *Если $(a, m) = 1$, и x пробегает приведённую систему вычетов по модулю m, то произведение ax также пробегает приведённую систему вычетов.*

Д о к а з а т е л ь с т в о. Действительно, чисел ax столько, сколько чисел x, т.е. c. Покажем, что ax_1 и ax_2 несравнимы, если несравнимы x_1 и x_2. Положим противное, т.е., что

$$ax_1 \equiv ax_2(\mathrm{mod}m).$$

Тогда, так как $(a, m) = 1$, то в силу 1.4.8

$$x_1 \equiv x_2(\mathrm{mod}m),$$

что противоречит условию.

П р и м е р 1. 3.

Пусть $m = 8$. $a = 11$, $(11, 8) = 1$. Приведённая система вычетов есть 1, 3, 5, 7;

$$11 \cdot 1 \equiv 3(\mathrm{mod}8),$$
$$11 \cdot 3 \equiv 1(\mathrm{mod}8),$$
$$11 \cdot 5 \equiv 7(\mathrm{mod}8),$$
$$11 \cdot 7 \equiv 5(\mathrm{mod}8).$$

Пусть $a = 13$, $(13, 8) = 1$. Получим:

$$13 \cdot 1 \equiv 5(\mathrm{mod}8),$$
$$13 \cdot 3 \equiv 7(\mathrm{mod}8),$$
$$13 \cdot 5 \equiv 1(\mathrm{mod}8),$$
$$13 \cdot 7 \equiv 3(\mathrm{mod}8).$$

Следствие 1.7.3. *Пусть $(a, m) = 1$, и $ax \equiv 1(\mathrm{mod}m)$. Тогда, согласно* 1.3.16*, $ax = 1 + mt, ax - mt = 1$, и мы получили соотношение вида* 1.2.4*. Отсюда посредством процедуры раздела* 1.2 *отыскивается число x.*

1.8. Теоремы Эйлера и Ферма

Теорема 1.8.1 (Эйлер). *При* $m > 1$ *и* $(a, m) = 1$

$$a^{\varphi(m)} \equiv 1(\mathrm{mod} m).$$

Доказательство. Пусть x пробегает приведённую систему вычетов, т.е.

$$x = r_1, r_2, \ldots, r_c; \quad c = \varphi(m).$$

Тогда ax_1 также пробегает приведённую систему вычетов, быть может, в другом порядке:

$$ar_i \equiv \rho_i(\mathrm{mod} m).$$

Сравнения можно почленно перемножать

$$a^c \prod_{i=1}^{c} r_i \equiv \prod_{i=1}^{c} \rho_i(\mathrm{mod} m).$$

Если r_i и ρ_i принадлежат наименьшим неотрицательным вычетам, а это всегда можно сделать, то обе части сравнения можно сократить на одно и то же число $\prod_{i=1}^{c} r_i = \prod_{i=1}^{c} \rho_i$, взаимно простое с модулем, откуда

$$a^c \equiv 1(\mathrm{mod} m),$$

что и требовалось.

При $m = p \quad \varphi(p) = p - 1$.

Отсюда следует

Теорема 1.8.2 (Ферма). *При простом* p, *и* a, *не делящемся на* p,

$$a^{p-1} \equiv 1(\mathrm{mod} p).$$

Сравнение

$$a^p \equiv a(\mathrm{mod} p)$$

верно при всех a, *так как оно верно и при* a, *делящемся на* p.

Стоит показать, что это сравнение справедливо и при a, не делящемся на p. Действительно, из теоремы Ферма получаем (см. утверждение 1.3.1) $a^{p-1} = tp+1$, $a^p = atp+a = Tp+a$, $a^p \equiv a(\mathrm{mod} p)$.

1.9. Мультипликативность функции Эйлера

Теорема 1.9.1. *Если* $(a, b) = 1$, *то* $\varphi(ab) = \varphi(a)\varphi(b)$.

Доказательству теоремы предпошлем следующую лемму:

Лемма 1.9.2. *Сумма*

$$ax + by, \tag{1.9.18}$$

принимает все значения из приведенной системы вычетов по модулю ab *тогда и только тогда, когда* x *пробегает приведенную систему вычетов по модулю* b, *а* y *пробегает приведенную систему вычетов по модулю* a.

Доказательство. *Достаточность* Легко видеть, что

$$(ax + by, ab) = 1. \tag{1.9.19}$$

Действительно, $(ax, b) = 1$ по условию, а $b|by$. Поэтому $(ax + by, b) = 1$. (В противном случае, т.е. если $(ax + by, b) = d > 1$, то $d|(ax + by)$, $d|b$. Отсюда $d|by$, а значит, $d|(ax + by - by)$, т.е, $d|ax$, что противоречит тому, что $(ax, b) = 1$.) Из соображений симметрии также $(ax + by, a) = 1$. Отсюда следует (1.9.19).

Необходимость. Если в сумме (1.9.18) хотя бы одно x (или также y) не принадлежит приведенной системе вычетов по модулю b, (по модулю a,), то такая сумма не принадлежит приведенной системе вычетов по модулю ab. Действительно, пусть, например, $(x, b) = d > 1$. Тогда d делит оба слагаемые в (1.9.18), и, следовательно, сумма (1.9.18) не взаимно проста с ab.

Лемма доказана.

Следствие 1.9.3. *Сумма* (1.9.18) *не взаимно проста с* ab *тогда и только тогда, когда хотя бы одно* x *(или такжде* y*) не принадлежит приведенной системе вычетов по модулю* b, *(по модулю* a,*)*

Теперь из полной системы вычетов по модулю ab удалим все Z вычетов, не взаимно простых с ab. Останутся $ab - Z = \varphi(ab)$ вычетов, взаимно простых с ab и только они. Из общего числа ab сумм $ax + by$ удалим те из них, в которых выполняется хотя бы одно из соотношений $(x, b) \neq 1$, $(y, a) \neq 1$. Их будет ровно

столько же, т.е. Z. Оставшихся $ab - Z$ сумм будет в точности $\varphi(a)\varphi(b)$.

Таким образом, произведение $\varphi(a)\varphi(b)$ и число $\varphi(ab)$ выражают одну и ту же величину, а потому $\varphi(ab) = \varphi(a)\varphi(b)$, и теорема доказана.

Приведём другое доказательство теоремы 1.9.1.

Для этого рассмотрим таблицу

$$\begin{array}{rrrl} 1 & 2 & 3 & \dots b \\ b+1 & b+2 & b+3 & \dots 2b \\ 2b+1 & 2b+2 & 2b+3 & \dots 2b \\ \dots & \dots & \dots & \dots \\ (a-1)b+1 & (a-1)b+2 & (a-1)b+3 & \dots ab \end{array} \tag{1.9.20}$$

На основании теоремы 1.6.2 числа каждого столбца пробегают полную систему вычетов по модулю a, и, таким образом, в каждом столбце содержится в точности $\varphi(a)$ чисел, взаимно простых с a. Каждый столбец имеет вид

$$r,\ b+r,\ 2b+r,\ \dots,\ (a-1)b+r. \tag{1.9.21}$$

Числа, взаимно простые с b содержатся в тех и только в тех столбцах, в которых $(b, r) = 1$. Это означает, что во всех $\varphi(b)$ столбцах (1.9.21), где $(b, r) = 1$, и толко в них содержится $\varphi(a)\varphi(b)$ чисел, взаимно простых одновременно и с a и с b, а значит, и с их произведением ab. Но всего в таблице (1.9.20) содержится $\varphi(ab)$ таких чисел. Отсюда $\varphi(ab) = \varphi(a)\varphi(b)$, что и требоваось.

1.10. Вычисление функции Эйлера

Теорема 1.10.1. *Пусть*

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k},$$

где числа

$$p_1, p_2, \dots, p_k$$

простые и попарно различные. Тогда

$$\varphi(a) = a \prod_{i=1}^{k} (1 - \frac{1}{p_k}). \tag{1.10.22}$$

Д о к а з а т е л ь с т в о. Так как числа

$$p_1, p_2, \ldots, p_k$$

взаимно просты, то согласно теореме 1.9.1

$$\varphi(a) = \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2}) \ldots \varphi(p_k^{\alpha_k}). \tag{1.10.23}$$

Но

$$\varphi(p_i^{\alpha_i}) = p_i^{\alpha_i} - p_i^{\alpha_i - 1} = p_i^{\alpha_i - 1}(p_i - 1) = p_i^{\alpha_i}(1 - \frac{1}{p_i}).$$

Действительно, среди $p_i^{\alpha_i}$ чисел $1, 2, \ldots, p_i^{\alpha_i}$ ровно $p_i^{\alpha_i - 1}$ чисел делится на p_i. Остальные

$$p_i^{\alpha_i} - p_i^{\alpha_i - 1} = p_i^{\alpha_i}(1 - \frac{1}{p_i})$$

чисел не делятся на p_i, а значит, взаимно просты с ним и с $p_i^{\alpha_i}$. Отсюда и из (1.10.23) получается (1.10.22).

П р и м е р 1. 4.

1). $a = 1000 = 2^3 5^3$, $\varphi(1000) = 1000(1 - 1/2)(1 - 1/5) = 400$;

2). $a = 255 = 3 \cdot 5 \cdot 17, \varphi(255) = 255(1 - 1/3)(1 - 1/5)(1 - 1/17) = 128$;

3). $a = 13068 = 4 \cdot 27 \cdot 121 = 2^2 3^3 11^2, \varphi(13068) = 13068(1 - 1/2)(1 - 1/3)(1 - 1/11) = (1/2)(2/3)(10/11) = 1980$;

4). $a = 1001 = 7 \cdot 11 \cdot 13, \varphi(1001) = 6 \cdot 10 \cdot 12 = 720$.

Если бы перед выводом теоремы Эйлера была предложена, например, задача — найти остаток от деления числа 729^{720} на 1001 — она могла вызвать недоумение. На самом деле $1001 = 7 \cdot 11 \cdot 13$, $\varphi(1001) = 720$, $(729, 1001) = 1$, $729^{720} \equiv 1 (\mathrm{mod}\ 1001)$.

1.11. Первообразные корни

При $(a, m) = 1$ существуют положительные целые числа γ с условием

$$a^{\gamma} \equiv 1 (\mathrm{mod} m). \tag{1.11.24}$$

Например, (теорема Эйлера)

$$a^{\varphi(m)} \equiv 1 (\mathrm{mod} m).$$

Определение 1.11.1. *Наименьшее из чисел γ, удовлетворяющее* (1.11.24) *называется показателем, которому a принадлежит по модулю m.*

П р и м е р 1. 5.

$$2^1 = 2 \equiv 2(\mathrm{mod}5),$$
$$2^2 = 4 \equiv 4(\mathrm{mod}5),$$
$$2^3 = 8 \equiv 3(\mathrm{mod}5),$$
$$2^4 = 16 \equiv 1(\mathrm{mod}5).$$

Таким образом, 4 – есть показатель, которому 2 принадлежит по модулю 5, $\varphi(5) = 4$.

В то же время

$$5^1 = 5 \equiv 5(\mathrm{mod}12),$$
$$5^2 = 25 \equiv 1(\mathrm{mod}12).$$

Таким образом, число 2 – есть показатель, которому 5 принадлежит по модулю 12, хотя $\varphi(12) = 4$. Более того, все элементы приведённой системы вычетов по модулю 12 принадлежат показателю 2.

Утверждение 1.11.2. *Если a по модулю m принадлежит показателю δ, то числа $1 = a^0, a, \ldots, a^{\delta-1}$ попарно несравнимы по модулю m.*

Действительно, из сравнения $a^l \equiv a^k(\mathrm{mod}m),\ (0 \le k < l < \delta)$ следовало бы $a^{l-k} \equiv 1(\mathrm{mod}m),\ (0 \le l - k < \delta)$, что противоречит определению величины δ : ведь δ есть наименьшее из чисел γ, для которых $a^\gamma \equiv 1(\mathrm{mod}m)$.

Утверждение 1.11.3. *Если a по модулю m принадлежит показателю δ, то $a^{\gamma_1} \equiv a^{\gamma_2}(\mathrm{mod}\, m)$ тогда и только тогда, когда $\gamma_1 \equiv \gamma_2(\mathrm{mod}\delta)$, т.е. $\gamma_1 - \gamma_2 = \delta t$.*

Сначала перед доказательством приведем

П р и м е р 1. 6.

$$2^1 = 2 \equiv 2(\mathrm{mod}15),$$
$$2^2 = 4 \equiv 4(\mathrm{mod}15),$$
$$2^3 = 8 \equiv 8(\mathrm{mod}15),$$
$$2^4 = 16 \equiv 1(\mathrm{mod}15).$$

Таким образом, число 2 принадлежит показателю $\delta = 4$ по модулю 15.

Далее

$$2^5 = 32 \equiv 2(\mathrm{mod}15),\ 5 \equiv 1(\mathrm{mod}4),$$
$$2^6 = 64 \equiv 2^2(\mathrm{mod}15),\ 6 \equiv 2(\mathrm{mod}4),$$
$$2^7 = 128 \equiv 2^3(\mathrm{mod}15),\ 7 \equiv 3(\mathrm{mod}4),$$
$$2^8 = 256 \equiv 2^0(\mathrm{mod}15),\ 8 \equiv 0(\mathrm{mod}4).$$

Иначе говоря, сравнимость последовательных степеней числа a, принадлежащего показателю δ, наступает с периодичностью δ.

Доказательству предпошлем следующее построение. Пусть r_1 и r_2 – наименьшие неотрицательные вычеты по модулю δ; $r_1 < \delta, r_2 < \delta$, тогда при некоторых q_1 и q_2 имеем $\gamma_1 = \delta q_1 + r_1,\ \gamma_2 = \delta q_2 + r_2$. Отсюда и из сравнения $a^\delta \equiv 1(\mathrm{mod}m)$ получаем

$$a^{\gamma_1} = (a^\delta)^{q_1} a^{r_1} \equiv a^{r_1}(\mathrm{mod}m), \quad a^{\gamma_2} = (a^\delta)^{q_2} a^{r_2} \equiv a^{r_2}(\mathrm{mod}m). \tag{1.11.25}$$

Д о к а з а т е л ь с т в о. *Необходимость.* Пусть $a^{\gamma_1} \equiv a^{\gamma_2}(\mathrm{mod}m)$.

Тогда $a^{r_1} \equiv a^{r_2}(\mathrm{mod}m)$, и в силу утверждения 1.11.2 $r_1 = r_2 = r$, (так как в противном случае a^{r_1} и a^{r_2} были бы несравнимы), а потому $\gamma_1 = \delta q_1 + r,\ \gamma_2 = \delta q_2 + r$, т.е. $\gamma_1 \equiv \gamma_2(\mathrm{mod}\delta)$.

Достаточность. Пусть $\gamma_1 \equiv \gamma_2(\mathrm{mod}\delta)$. Тогда $r_1 = r_2$, и из (1.11.25) получаем, что $a^{\gamma_1} \equiv a^{\gamma_2}(\mathrm{mod}m)$. Теорема доказана.

Имеем далее

$$a^{\varphi(m)} \equiv 1(\mathrm{mod}m).$$

т.е.

$$a^{\varphi(m)} \equiv a^0(\mathrm{mod}m).$$

Отсюда и из предыдущего рассуждения следует, что

$$\varphi(m) \equiv 0(\mathrm{mod}\delta).$$

Таким образом, справедливо

Утверждение 1.11.4. *Показатели, которым числа принадлежат по модулю m, суть делители числа $\varphi(m)$. Наибольший из этих делителей есть само $\varphi(m)$.*

Определение 1.11.5. *Числа, принадлежащие показателю $\varphi(m)$, называются первообразными корнями по модулю m,*

(если они существуют).

Иначе говоря, a будет первообразным корнем по модулю m, когда ни при каком $\varepsilon < \varphi(m)$ не выполняется сравнение

$$a^{\varepsilon} \equiv 1(\mathrm{mod} m).$$

Все случаи, когда существуют первообразные корни по модулю m, суть

$$m = 2,\ 4,\ p^n,\ 2p^n.$$

Доказательство этого факта здесь опущено. Читатель найдёт его в руководствах по теории чисел.

Утверждение 1.11.6. *В приведенной системе вычетов по модулю m количество чисел, принадлежащих показателю δ, есть $\varphi(\delta.)$*

Действительно, покажем, что если a принадлежит показателю δ по модулю m, и $(\gamma, \delta) = 1$, то a^{γ} также принадлежит показателю δ по модулю m.

Предположим противное, т.е., что $(a^{\gamma})^{\delta_1} \equiv 1(\mathrm{mod} m)$, где $\delta_1 < \delta$. Тогда, согласно утверждению 1.11.3, $\gamma\delta_1 \equiv 0(\mathrm{mod}\delta)$. В свою очередь это означает, что $\gamma\delta_1$ делится на δ, что противоречит условию $\delta_1 < \delta$, а потому $\delta_1 = \delta$, что и требовалось.

Пусть теперь $(\gamma, \delta) = d > 1$. Положим $\delta/d = \delta_1 < \delta$. Тогда $(a^{\gamma})^{\delta_1} = (a^{\gamma})^{\frac{\delta}{d}} = (a^{\delta})^{\frac{\gamma}{d}} = 1$. Это означает, что a принадлежит показателю $\delta_1 < \delta$ по модулю m.

В частности, из утверждения 1.11.6 следует, что *число первообразных корней есть*

$$\varphi(c) = \varphi(\varphi(m)).$$

Пусть $c = \varphi(m)$, и $q_1, q_2, \ldots, q_k$ – различные простые делители числа c.

Утверждение 1.11.7. *Для того, чтобы g, взаимнопростое с m, было первообразным корнем по модулю m, необходимо и достаточно, чтобы g не удовлетворяло ни одному из сравнений $g^{\frac{c}{q_i}} \equiv 1(\mathrm{mod} m)$.*

Д о к а з а т е л ь с т в о. *Необходимость.* Если

$$g^{\frac{c}{q_i}} \equiv 1,$$

то $c = \varphi(m)$ не показатель, которому принадлежит g по модулю m.

Достаточность. Пусть не выполняется ни одно из сравнений $g^{\frac{c}{q_i}} \equiv 1(\mathrm{mod} m)$, и пусть показатель $\delta < c$. Тогда c/δ – целое, и пусть $c/\delta = qt,\ c/q = \delta t,\ g^{\delta t} = (g^{\delta})^t \equiv 1 \bmod m$. Это значит, $g^{c/q} \equiv 1 \bmod m$, что противоречит условию.

Более простое доказательство этого факта получается из элементарных сведений о циклических группах (см. след. гл.)

1.12. Индексы

Пусть $p-$ простое нечетное, $n \geq 1,\ m$ – одно из чисел $2,\ 4,\ p,\ 2p^n$ и пусть $c = \varphi(m)$.

Пусть g — первообразный корень по модулю m. Заметим, что $(g, m) = 1$, так как g принадлежит приведенной системе вычетов по модулю m.

Утверждение 1.12.1. *Если γ пробегает полную систему наименьших неотрицательных вычетов $\gamma = 0, 1, \ldots, c-1$ по модулю c, то g^{γ} пробегает приведенную систему вычетов по модулю m.*

Д о к а з а т е л ь с т в о. g^{γ} пробегает c чисел взаимнопростых с m, так как $(g, m) = 1$. Остается применить утверждение 1.11.2.

Определение 1.12.2. *Пусть $(a, m) = 1$. Если $a \equiv g^{\gamma}(\mathrm{mod}\ m)$, то γ называется индексом числа a по модулю m при основании g и обозначается*

$$\gamma = \mathrm{ind}_g a.$$

Ввиду утверждения 1.12.1 всякое такое a, что $(a, m) = 1$, имеет единственный индекс γ_0 среди чисел $\gamma = 0, 1, \ldots, c-1$.

В самом деле, если одновременно

$$a \equiv g^{\gamma_1}(\mathrm{mod} m),$$

и

$$a \equiv g^{\gamma_2}(\mathrm{mod} m),$$

то

$$g^{\gamma_1} \equiv g^{\gamma_2} (\mathrm{mod} m),$$

что невозможно, так как тогда

$$g^{\gamma_1 - \gamma_2} \equiv 1 (\mathrm{mod} m),$$

и g не первообразный корень.

Из определения индекса следует, что числа с одинаковым индексом γ образуют класс вычетов по модулю m, так как два числа, сравнимые с третьим, сравнимы между собой. (Сравнимые числа принадлежат к одному классу.)

Зная γ_0, можно указать все индексы числа a :

$$g^{\gamma_1} \equiv g^{\gamma_2} (\mathrm{mod} m),$$

тогда и только тогда, когда

$$\gamma_1 \equiv \gamma_2 (\mathrm{mod} c).$$

Утверждение 1.12.3. *Имеет место следующее сравнение:*

$$\mathrm{ind}\, ab \cdots l \equiv \mathrm{ind}\, a + \mathrm{ind}\, b + ... + \mathrm{ind}\, l (\mathrm{mod} m).$$

Действительно,

$a \equiv g^{\mathrm{ind}\, a} (\mathrm{mod} m),\ b \equiv g^{\mathrm{ind}\, b} (\mathrm{mod} m),\ \ldots\ , l \equiv g^{\mathrm{ind}\, l} (\mathrm{mod} m).$

Сравнения можно почленно перемножать, поэтому:

$ab \cdots l \equiv g^{\mathrm{ind}\, a + \mathrm{ind}\, b + \cdots + \mathrm{ind}\, l}$.

Следовательно, $\mathrm{ind}\, a + \mathrm{ind}\, b + ... + \mathrm{ind}\, l$ есть один из индексов произведения $ab \cdots l$

Аналогия с логарифмами очевидна.

1.13. Приложения к криптографии

Было бы неверным сказать, что криптографическая тематика далека от теории кодирования. Однако для начального курса алгебраических кодов криптографические вкрапления могут показаться несколько чужеродными. Поэтому целью данного

раздела является не столько изучение методов защиты информации от несанкционированного доступа, сколько демонстрация некоторых интересных практических возможностей функции Эйлера.[1]

Для примера рассмотрим криптографический метод RSA, названный так по начальным буквам фамилий его авторов Р. Ривеста, А. Шамира и Л. Адлмана.

В процессе защиты информации от несанкционированного доступа с одной стороны, и попыток взломать систему защиты – с другой, участвуют два субъекта. Это криптограф, который шифрует информацию и легальным образом дешифрует ее, и криптоаналитик , который пытается взломать систему защиты.

И криптографу и криптоаналитику известно следующее: Модуль N и такое число e, что $e, (\varphi(N), e) = 1$. Это так называемый открытый ключ, содержащийся в справочнике, который можно купить в газетном киоске.

Разложение $N = P \times Q$, где P, Q — простые числа, это закрытый, секретный ключ.

Он известен только криптографам, но неизвестен криптоаналитикам. Утверждение, что разложение $N = P \times Q$, может быть кому-то неизвестным, на первый взгляд кажется нелепым. Кто не знает, как разложить число на множители!? Однако именно в этом утверждении заключен весь корень прочности системы защиты.

Имеем $\varphi(N) = \varphi(P)\varphi(Q) = (P-1)(Q-1)$.

Шифрование происходит следующим образом:

Пусть следует передать число M. В действительности передается

$$C \equiv M^e (\mathrm{mod} N).$$

Так как $(e, \varphi(N)) = 1$, то легко заранее найти такое x, что $e \times x \equiv 1 (\mathrm{mod} \varphi(N))$.

В самом деле, на основании теоремы 1.7.2 в разделе 1.7 имеем:

[1] В лекциях завершение доказательства теоремы Эйлера всякий раз производило впечатление неожиданностью результата, появлявшегося после довольно скучного рассказа о свойствах сравнений. Простота и изящество теоремы, казалось, придавали ей очевидную эстетическую значимость, усомниться в которой было невозможно. Но однажды лектор был обескуражен вопросом: "А каково практическое значение этой теоремы?" С тех пор теорема Эйлера непременно сопровождалась рассказом о ее криптографической ценности. Потенциальный критик теоремы переходил в ряды ее почитателей, ибо ничто так не утверждает человека в собственной значимости (а заодно и значимости теоремы), как сознание приобщенности к секретам.

Если при $(a, m) = 1$ *элемент* x *пробегает приведенную систему вычетов по модулю* m, *то и* ax *пробегает приведенную систему вычетов.*

Здесь $m = \varphi(N), a = e$. Для отыскания x можно с успехом воспользоваться равенством (1.2.4), применив для этого процедуру раздела 1.2. В (1.2.4) следует положить $a = e,\ b = \varphi(N),\ d = 1$. Тогда $s = x$.

Из $e \times x \equiv 1(\mathrm{mod}\varphi(N))$ на основании (1.3.16) следует $e \times x = y\varphi(N) + 1$, где y целое.

Дешифрование:

Получив $C \equiv M^e(\mathrm{mod}N)$, поступают следующим образом:

$$C^x \equiv M^{ex}(\mathrm{mod}N) \equiv M^{y\varphi(N)+1}(\mathrm{mod}N) \equiv (M^{\varphi(N)})^y M(\mathrm{mod}N).$$

Существуют две возможности.

$$1.\quad (M, N) = 1, M < N. \tag{1.13.26}$$

В этом случае

$$(M^{\varphi(N)})^y M \equiv M(\mathrm{mod}N),$$

так как в силу (1.13.26)

$$M^{\varphi(N)} \equiv 1(\mathrm{mod}N).$$

Отсюда $C^x \equiv M(\mathrm{mod}N)$, и на основании (1.3.16) $C^x = M + uN$. Поэтому в силу (1.13.26) немедленно получается $u = 0$, $C^x = M$.

$$2.\quad (M, N) > 1, M < N. \tag{1.13.27}$$

Так как $N = PQ$ и P, $Q-$ простые числа, то вследствие (1.13.27) число M может быть кратным только одного из чисел P, Q. В самом общем случае можно считать $M = P^t L$, $L < Q$, $(L, N) = 1$, t – натуральное число. Имеем

$$M^{ex} = P^{tex} L^{ex} \equiv P^{t(y\varphi(N)+1)} L^{y\varphi(N)+1}(\mathrm{mod}N). \tag{1.13.28}$$

Рассмотрим множители в левой части сравнения (1.13.28) раздельно.

В силу условия $(L, N) = 1$

$$L^{y\varphi(N)+1} \equiv L(\mathrm{mod}N). \tag{1.13.29}$$

Далее, зная, что

$$\varphi(N) = \varphi(Q)\varphi(P),$$

положим

$$P^{ty\varphi(Q)\varphi(P)}P^t \equiv z(\mathrm{mod}PQ).$$

Если сравнение имеет место по модулю m, (см. утвержд. 1.5.4), то оно имеет место и по модулю d, равному любому делителю числа m.

Отсюда

$$P^{ty\varphi(Q)\varphi(P)}P^t \equiv z(\mathrm{mod}Q). \tag{1.13.30}$$

С другой стороны, так как $(P, Q) = 1$, то по теореме Ферма

$$P^{\varphi(Q)} \equiv 1(\mathrm{mod}Q),$$

и потому

$$P^t \equiv z(\mathrm{mod}Q). \tag{1.13.31}$$

Из (1.13.30), (1.13.31) в силу утверждения 1.4.1 следует

$$P^{ty\varphi(N)}P^t \equiv P^t(\mathrm{mod}Q).$$

Тривиальным образом

$$P^{ty\varphi(N)}P^t \equiv P^t(\mathrm{mod}P).$$

Если сравнение имеет место по нескольким модулям (см. утвер. 1.5.3) то оно имеет место и по модулю, равному наименьшему общему кратному модулей, т.е.

$$P^{ty\varphi(N)}P^t \equiv P^t(\mathrm{mod}PQ). \tag{1.13.32}$$

Из (1.13.28), (1.13.29) и (1.13.32) получаем окончательно

$$C^x \equiv M^{ex} = P^{tex}L^{ex} \equiv P^tL(\mathrm{mod}N).$$

Это означает, что

$$C^x = P^tL + vN.$$

Как и выше, ввиду $M < N$, $v = 0$, $C^x = P^tL = M$.

Центральную роль в вычислениях играет равенство

$$\varphi(N) = \ = \varphi(P)\varphi(Q) = (P-1)(Q-1).$$

При очень больших и почти равных P, Q найти разложение $N = PQ$ — весьма трудоемкая операция, что и лежит в основе надежды на длительность взлома системы защиты криптоаналитиком. (Попытку взлома называют атакой).

П р и м е р 1. 7.

$P = 17, Q = 31, N = 17 \times 31 = 527, \varphi(N) = \varphi(P)\varphi(Q) = (P-1)(Q-1) = 16 \times 30 = 480$. Пусть $e = 7$, $(7, 480) = 1$. Отсюда $x = 343$.

Действительно, из $e \times x = y\varphi(N)+1$ следует $7x = 480y+1 = (68 \times 7 + 4)y + 1$. Отсюда $(4y+1)/7-$ целое, т.е. $4y + 1 = 7z$. Отсюда $y = (7z-1)/4-$ целое, т.е. $3z - 1 = 4v$.

Отсюда $z = (4v+1)/3-$ целое, т.е. $v+1 = 3$, и $v = 2$. Далее $3z = 8+1$, $z = 3$, $4y+1 = 21$, $y = 5$, $7x = 480 \times 5+1 = 2401$, $x = 343$, что и требовалось.

Положим $M = 5$. Тогда $C = 5^7 \equiv 129(\mathrm{mod}527.)$

Получив в виде сообщения число 129, мы должны извлечь из него корень седьмой степени по mod527.

Согласно вышесказанному, $C^x = 129^{343}(\mathrm{mod}527.)$

Имеем

$129^2 = 16641 = 527 \times 31+304$, и $343 = 256+64+16+4+2+1$

Отсюда

$129^{343} = 129^{256} \times 129^{64} \times 129^{16} \times 129^4 \times 129^2 \times 129 \equiv 304^{128} \times$
$\times\, 304^{32} \times 304^8 \times 304^2 \times 304 \times 129 \equiv$

$$[\text{так как } 304^2 = 92416 = 527 \times 175 + 191]$$

$$\equiv 191^{64} \times 191^{16} \times 191^4 \times 191 \times 304 \times 129 \equiv$$

$$[\text{так как } 191^2 = 36481 = 527 \times 69 + 118]$$

$$\equiv 118^{32} \times 118^8 \times 118^2 \times 191 \times 304 \times 129 \equiv$$

$$[\text{так как } 118^2 = 13924 = 527 \times 26 + 222]$$

$$\equiv 222^{16} \times 222^4 \times 222 \times 191 \times 304 \times 129 \equiv$$

$$[\text{так как } 222^2 = 49284 = 527 \times 93 + 273]$$

$$\equiv 273^8 \times 273^2 \times 222 \times 191 \times 304 \times 129 \equiv$$

$$[\text{так как } 273^2 = 74529 = 527 \times 141 + 222]$$

$\equiv 222^4 \times 222 \times 222 \times 191 \times 304 \times 129 \equiv 273^3 \times 191 \times 304 \times 129 \equiv 273 \times 222 \times 191 \times 304 \times 129 \equiv 5 (\mathrm{mod} 527)$,

что и требовалось.

Подчеркнем, что не зная $\varphi(N)$, криптоаналитик не может вычислить x. Знание $\varphi(N)$ требует знания разложения $N = P \times Q$, и получение этого разложения — самая трудоемкая процедура. Объем $L(N)$ вычислений для разложения $N = PQ$ имеет порядок

$$L(N) \sim e^{\sqrt{\ln N \times \ln \ln N}}.$$

Обычно N не менее, чем семисот-, а то и тысячезначное число. Прямая и обратная функции, т.е., с одной стороны, вычисление $C \equiv M^e (\mathrm{mod} N)$, а с другой — получение $C^x (\mathrm{mod} N)$, где главную роль играет отыскание неизвестного криптоаналитику разложения $N = PQ$, по трудоемкости — несоизмеримы. Такие функции принято называть односторонними.

Не доказана неизбежность знания разложения $N = PQ$. Можно ли избрать другой путь дешифрования, не использующий соотношение $\varphi(N) = \varphi(P)\varphi(Q)$?

1.14. Задачи к главе 1

1.1. Построить таблицы сложения и умножения наименьших неотрицательных вычетов по mod 7.

1.2. Пусть $m > 0, (a, m) = 1, b$ – целое и x пробегает полную, а ξ приведенную систему вычетов по модулю m. Доказать, что $\sum_x \{(ax + b)/m\} = \frac{1}{2}(m - 1)$, $\sum_\xi \{(a\xi)/m\} = \frac{1}{2}\varphi(m)$.

1.3. Найти наименьшее положительное число, при делении которого на 17, 13 и 10 получаются остатки соответственно 15, 11 и 3.

1.4. Доказать, что показатель, с которым данное простое число p входит в произведение $n!$, равен $[n/p] + [n/p^2] + [n/p^3] + \ldots$.

1.5. Найти все первообразные корни по модулям 9, 11, 14, 17, 18.

1.6. Не прибегая к алгоритму Эвклида, доказать, что если $(a, b) = d$, то найдутся два таких целых s и t, которые удовлетворяют условию $as + bt = d$.

1.7. Доказать теорему Вильсона: $(p - 1)! \equiv -1 (\mathrm{mod}\, p)$, где p простое.

1.8. Пусть целое $a > 1$. Доказать, что простые нечетные делители числа $a^p - 1$ делят $a - 1$ или имеют вид $2px + 1$, где p – простое нечетное число.

1.9. Пусть целое $a > 1$. Доказать, что простые нечетные делители числа $a^p + 1$ делят $a + 1$ или имеют вид $2px + 1$, где p – простое нечетное число.
1.10. Придумать доказательство теоремы Ферма, отличное от доказательства в тексте.
1.11. Вывести теорему Эйлера из теоремы Ферма.
1.12. Полагая $N = 3 \cdot 5,\ 3 \cdot 7,\ 5 \cdot 7,\ 3 \cdot 11,$ провести процедуру шифрования-дешифрования для разнообразных значений M.

Глава 2.

Элементы теории групп, колец и полей

2.1. Множество с операцией

Пусть дано некоторое множество M элементов. Говорят, что в M определена *алгебраическая операция* $*$, если всяким двум (различным или одинаковым) элементам множества M, взятым в определенном порядке, по некоторому закону ставится в соответствие вполне определенный элемент, принадлежащий этому же множеству M.

В таком случае говорят также, что множество M *замкнуто относительно некоторой операции* $*$, если для любых $a \in M$ и $b \in M$ имеет место $a * b \in M$ и $b * a \in M$.

Вообще говоря, операция $*$ не обязана быть коммутативной. Изображать ее мы будем либо в виде умножения ab, либо в виде сложения $a + b$, смотря по случаю. Примеры некоммутативной операции приведены ниже.

2.2. Обратная операция

Говорят, что для операции $*$, заданной в M, существует обратная операция, если при любых $a \in M$ и $b \in M$ каждое из уравнений

$$a * x = b, \quad x * a = b \tag{2.2.1}$$

имеет решение и при том единственное.

П р и м е р 2. 1.

Решить сравнение $3x \equiv 2(\mod 11)$. Нетрудно проверить, что решением будет $x \equiv 8(\mathrm{mod}11)$. Имеем

$$3x = 2{+}11x_1 = 9x_1{+}2x_1{+}2,\ 2x_1{+}2 = 3x_2, \quad 3x_2{-}2 \equiv 0(\mod 2),$$

$$3x_2 = 2t, \quad x_2 = 2t_1, \quad x_2 = x_1 = 2, \quad x \equiv 8(\mathrm{mod}11)$$

Легко видеть, что не для всякой операции в множестве M существует обратная операция. Например, для сложения натуральных чисел и умножения целых чисел.

2.3. Группа

Без ограничения общности опустим знак операции *, перейдя, таким образом, к употреблению операции, которую будем называть умножением.

Определение 2.3.1. *Непустое множество G называется группой, если выполняются следующие условия:*

1.1). *Множество замкнуто относительно некоторой операции.*

1.2). *Операция ассоциативна:*

$$(ab)c = a(bc).$$

1.3). *В G выполняется обратная операция.*

Группа называется *абелевой*, если всегда

$$ab = ba.$$

Некоммутативную группу даёт

П р и м е р 2. 2.

Некоммутативная операция — умножение перестановок в симметрической группе. Вообще, перестановкой степени n называется

$$\begin{pmatrix} 1, 2, \ldots, n \\ i_1, i_2, \ldots, i_n \end{pmatrix}$$

заменяющая в множестве элементов $1, 2, \ldots, n$ элемент l на i_l.

Имеется $n!$ перестановок из n элементов. Последовательное выполнение двух перестановок есть снова перестановка, и эта операция называется произведением перестановок.

$$\begin{pmatrix} 1, 2, \ldots, n \\ i_1, i_2, \ldots, i_n \end{pmatrix} \begin{pmatrix} 1, 2, \ldots, n \\ j_1, j_2, \ldots, j_n \end{pmatrix}$$

Ниже левые части двух равенств отличаются порядком сомножителей, а потому отличаются и правые части:

$$\begin{pmatrix} 1234 \\ 2341 \end{pmatrix} \begin{pmatrix} 1234 \\ 4213 \end{pmatrix} = \begin{pmatrix} 1234 \\ 2134 \end{pmatrix},$$

$$\begin{pmatrix}1234\\4213\end{pmatrix}\begin{pmatrix}1234\\2341\end{pmatrix}=\begin{pmatrix}1234\\1324\end{pmatrix}.$$

Читателю известна по крайней мере еще одна некоммутативная операция — умножение матриц. Некоммутативность заключена в несимметричности самой фразы, выражающей принцип умножения: "строка на столбец", где одно слово — подлежащее, а другое — дополнение. Умножение матриц будет коммутативным, только если они симметричны относительно главной диагонали.

Существуют другие определения группы, например:

Определение 2.3.2. *Непустое множество G называется группой, если выполняются следующие условия:*

2.1) *Множество G замкнуто относительно некоторой операции.*

2.2) *Операция ассоциативна,*

2.3) *Для каждого $a \in G$ существует хотя бы одна (левая) единица*

$$1a = a,$$

т.е. такой элемент, который оставляет элемент a неизменным.

2.4) *Для каждого $a \in G$ существует хотя бы один (левый) обратный элемент*

$$a^{-1}a = 1.$$

Легко показать, однако, что из первого определения следует второе.

Действительно, пусть уравнение $xa = b$ разрешимо при *любых* a и b.

Тогда разрешимо и уравнение

$$xc = c.$$

Обозначим решение этого уравнения через $x = e$, т.е. $ec = c$. Покажем, что это решение и есть (левая) единица, каково бы ни было c.

Опираясь на разрешимость уравнения $ax = b$, составим уравнение

$$cx = a,$$

где a произвольный элемент, и умножим его слева почленно на e :

$$ecx = ea.$$

Но выше было

$$ec = c,$$

и потому

$$ecx = cx = ea,$$

Однако одновременно по условию

$$cx = a,$$

и, значит,

$$ea = a.$$

Отсюда следует, что

$$e = 1,$$

для любого $a \in G$, (так как a произвольно).

Условие 2.4 непосредственно следует из разрешимости уравнения $xa = 1$.

Легко также показать, что в 2.3 и 2.4 левый обратный элемент является также и правым, а левая единица — правой.

Дано

$$a^{-1}a = 1.$$

Умножим это равенство сначала справа на

$$a^{-1}:$$

$$(a^{-1}a)a^{-1} = 1a^{-1} = a^{-1},$$

а затем слева на

$$(a^{-1})^{-1}.$$

Это можно сделать, так как по доказанному левый обратный элемент $(a^{-1})^{-1}$ для a^{-1} существует.

Теперь воспользуемся ассоциативностью операции умножения:

$$(a^{-1})^{-1}(a^{-1}a)a^{-1} = (a^{-1})^{-1}(1a^{-1}) = (a^{-1})^{-1}a^{-1} = 1.$$

С другой стороны,

$$(a^{-1})^{-1}(a^{-1}a)a^{-1} = 1aa^{-1}.$$

Поэтому

$$1aa^{-1} = 1,$$

т.е.

$$aa^{-1} = 1.$$

и левый обратный элемент есть и правый обратный.

Далее, в выражении $a1$ заменим 1 на

$$a^{-1}a.$$

Получим

$$aa^{-1}a = 1a,$$

так как по доказанному только что a^{-1} является также и правым обратным элементом. Иначе говоря, $a1 = 1a$, т.е левая единица является одновременно и правой, что и требовалось.

Доказательство того, что из второго определения следует первое — тривиально. Разрешимость уравнений

$$a * x = b \quad \text{и} \quad x * a = b,$$

т.е. наличие обратной операции, следует из наличия правого и одновременно левого обратного элемента a^{-1} :

$$a^{-1}ax = a^{-1}b,\ x = a^{-1}b;\ xaa^{-1} = ba^{-1};\ x = ba^{-1}.$$

Так вводится деление, как операция, обратная умножению.

Обратный элемент произведения равен произведению обратных элементов, взятых в обратном порядке:

$$(ab)^{-1} = b^{-1}a^{-1}, \quad \text{так как} \quad (b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = 1.$$

Можно иначе.

Найдём x в уравнении $(ab)x = 1$. Для этого умножим обе части слева сначала на a^{-1}:

$$a^{-1}abx = a^{-1}, \quad bx = a^{-1},$$

а затем на b^{-1}:

$$b^{-1}bx = b^{-1}a^{-1}, \quad \text{и, значит,} \quad x = b^{-1}a^{-1}.$$

Аналогично, пусть $x(ab) = 1$. Умножим обе части справа сначала на b^{-1}:

$$xabb^{-1} = b^{-1}, \quad xa = b^{-1},$$

а затем на a^{-1} :

$$xaa^{-1} = b^{-1}a^{-1}, \quad \text{и значит,} \quad x = b^{-1}a^{-1}.$$

Если условиться,что

$$a_1 a_2 \dots a_m = \prod_{i=1}^{m} a_i,$$

то легко видеть, что

$$\prod_{i=1}^{m} a_i \prod_{j=1}^{n} a_{m+j} = \prod_{i=1}^{m+n} a_i.$$

Далее, полагая

$$a^n = \underbrace{aa \dots a}_{n \text{ раз}},$$

получим

$$a^n a^m = a^{n+m}$$

и

$$(a^m)^n = a^{mn}.$$

Группа по умножению называется *мультипликативной* группой, а группа по сложению — *аддитивной.* Обратному элементу a^{-1} в мультипликативной группе отвечает противоположный элемент $-a$ в аддитивной, а единице отвечает нуль. Для абелевых групп иногда целесообразно записывать групповую операцию аддитивно, т.е. писать $a + b$ вместо ab. Вместо $a + (-b)$ можно кратко писать $a - b$, так как разность $a - b$ есть решение уравнения $x + b = a$. В самом деле, положив $x = a - b$, и подставив это в уравнение, получим $(a-b)+b = a+(-b+b) = a+0 = a$, т.е. $a - b$ удовлетворяет уравнению. Так вводится вычитание, как операция, обратная сложению.

2.4. Порядок группы и порядок элемента группы

Определение 2.4.1. *Число элементов конечной группы называется порядком группы.*

Если все степени элемента a группы являются различными элементами группы, то a называется элементом бесконечного порядка. Если же среди них имеются одинаковые (в случае конечной группы это обязательно имеет место), например,

$$a^k = a^l, \quad k > l,$$

то

$$a^{k-l} = 1.$$

Это означает, что существуют степени, равные 1.

Определение 2.4.2. *Порядком элемента a группы называется наименьшее целое $n > 0$, при котором*

$$a^n = 1, \tag{2.4.2}$$

Иными словами, если n есть порядок элемента a, то
1)

$$a^n = 1$$

при $n > 0$, и
2) *если*

$$a^k = 1, \ k > 0,$$

то $k \geq n$.

(Ср. с определением 1.11.1)

В этом случае говорят, что a есть *элемент порядка n.*

Отсюда, умножив обе части (2.4.2) на a^{-1}, сразу имеем $a^{-1} = a^{n-1}$.

Теорема 2.4.3. *Если элемент a имеет порядок n, то*

1) *Все элементы* $1, a, \ldots, a^{n-1}$ *различны.*(Ср. с утверждением 1.11.2)*)*

2) *Всякая другая степень элемента a, положительная или отрицательная, равна одному из этих элементов.*

Действительно, если

$$a^k = a^l, \quad \text{и} \quad n - 1 \geq k > l \geq 1, \quad \text{то} \quad a^{k-l} = 1,$$

что противоречит условию теоремы, так как $k - l < n$.

Далее, если

$$k = nq + r, \quad \text{где} \quad 0 \leq r < n, \quad \text{то} \quad a^k = (a^n)^q a^r = a^r.$$

Наконец, если $a^k = 1$, то и $a^r = 1$. Отсюда $a^k = (a^n)^q$, и k кратно n.

Определение 2.4.4. *Наибольший из порядков элементов группы называется показателем группы.*

В случае бесконечной группы говорят о группе бесконечной мощности, употребляя иногда словосочетание "группа бесконечного порядка".

2.5. Примеры групп

1. Аддитивная группа целых чисел $\mathbb{Z}$.

2. Аддитивная группа всех рациональных чисел $\mathbb{Q}$.

3. Аддитивная группа действительных чисел $\mathbb{R}$.

4. Аддитивная группа комплексных чисел $\mathbb{C}$.

5. Аддитивная группа всех четных чисел.

Не будет аддитивной группой множество всех нечетных чисел, множество всех неотрицательных четных. Целые числа не образуют группы по умножению (мультипликативной группы.) Не образуют группу по умножению и все рациональные числа ввиду невозможности деления на нуль.

6. Все отличные от нуля рациональные числа уже образуют группу по умножению: мультипликативную группу рациональных чисел.

7. Мультипликативная группа положительных рациональных чисел $\mathbb{Q}$.

8. Мультипликативная группа положительных действительных чисел $\mathbb{R}$.

9. Аддитивная группа классов вычетов по модулю m.

10. Мультипликативная группа классов вычетов приведенной системы по модулю m.

11. Все комплексные числа, являющиеся корнями из единицы степени n, образуют мультипликативную группу порядка n.

12. Невырожденные квадратные матрицы порядка n образуют некоммутативную мультипликативную группу.

13. Группа вращений правильных многогранников.

14. Симметрическая группа, т.е. группа S_n подстановок (некоммутативная) порядка $|S_n| = n!$ и степени n.

15. Группа всех p^n векторов длины n с основанием p и операцией поразрядного сложения по модулю p.

2.6. Подгруппы

Определение 2.6.1. *Подмножество H группы G называется подгруппой этой группы, если оно само является группой*

относительно операции, определенной в группе G.

Для установления факта, является ли (непустое) подмножество H элементов группы G подгруппой группы G, достаточно проверить

1) Замкнутость множества относительно данной операции.

2) Содержится ли в H вместе с a также и a^{-1}.

Ассоциативность следует автоматически, так как она имеет место в G.

Наличие единицы следует из 1) и 2).

В случае конечной группы условие 2) автоматически следует из 1): вместе с a вследствие замкнутости в H содержится и произведение $aa\ldots a$. Так как группа конечна, то найдется такое n, что $a^n = 1$, и

$$\underbrace{aa\ldots a}_{n-1 \text{ раз}} = a^{n-1} = a^{-1}.$$

Для абелевой группы, где групповая операция записывается аддитивно, достаточно потребовать, чтобы вместе с a и b содержалась и разность $a - b$. (Вместо требования, чтобы содержались $a + b$ и $-a$.)

Примеры подгрупп

1. Единичная группа, т.е. группа, состоящая только из единицы, и вся группа — несобственные подгруппы. Остальные — собственные, или истинные.

2. Подгруппой аддитивной группы целых чисел является множество всех целых чисел, кратных некоторому m.

3. Подгруппа группы всех p^n векторов с основанием p.

4. Подгруппа всех четных подстановок симметрической группы, так называемая, знакопеременная группа.

5. Подгруппа симметрической группы, оставляющая на месте фиксированный элемент.

2.7. Порождающие элементы и циклические группы

Утверждение 2.7.1. *Пересечение* $G_0 = G_1 \cap G_2 \cap \ldots \cap G_n$ *групп* $G_1, G_2, \ldots, G_n$ *также будет группой.*

Действительно, во-первых, G_0 не пусто, так как всем подгруппам G_1, G_2, $\ldots$, G_n принадлежит единица группы. Если единица — единственный элемент в G_0, то G_0 несобственная подгруппа. Пусть, далее, G_0 содержит элементы, отличные от единицы, т.е. $a, b \in G_0$. Тогда $a, b \in G_i$, $i = 1, 2, \ldots n$, а значит, $ab \in G_i$, и потому $ab \in G_0$. Таким образом, G_0 замкнуто. Далее, если $a \in G_0$, то $a \in G_i$, $i = 1, 2, \ldots n$, а значит, $a^{-1} \in G_i$, и потому $a^{-1} \in G_0$ Таким образом, G_0 вместе с a содержит a^{-1}. Если a является единственным отличным от единицы элементом в G_0, то это означает, что он обратный самому себе, т.е. $a = a^{-1}$. следовательно, a есть элемент второго порядка: $a^2 = 1$. Этим завершается доказательство.

Пусть теперь $a, b, \ldots, c$ — любые элементы группы G. Возьмем пересечение всех подгрупп группы G, которые содержат все эти элементы. Это пересечение есть подгруппа. Она называется подгруппой, порожденной элементами $a, b, \ldots, c$, и состоит из всех произведений и степеней (в том числе и отрицательных) элементов $a, b, \ldots, c$. Если взять только один единственный элемент $a \neq 1$ ($a \neq 0$, на случай аддитивной группы), то он порождает группу всех степеней $a^{\pm i}$, включая $a^0 = 1$ (всех кратных $\pm ia$, включая $0a = 0$ на случай аддитивной группы).

Определение 2.7.2. *Группа, порожденная одним элементом, называется циклической.*

Оговорка $a \neq 1$ ($a \neq 0$, на случай аддитивной группы), которая обычно подразумевается, необходима, так как в противном случае единственным элементом может оказаться только $a = 1$ ($a = 0$,), и группа будет состоять только из одного элемента.

Циклическую группу, порожденную элементом b, обычно обозначают символом $\{b\}$.

Всякая циклическая группа — коммутативна (абелева).

Порядок элемента и порядок группы, порожденной этим элементом, совпадают.

2.8. Подгруппы циклической группы

Пусть G — циклическая группа, и H — ее истинная подгруппа. Пусть G — есть совокупность степеней элемента a . Тогда и подгруппа H есть совокупность степеней элемента a.

Пусть a^l — наименьшая из положительных степеней. В таком случае в H содержатся и все степени $(a^l)^q$. Покажем, что ничего другого подгруппа H не содержит. В самом деле, пусть имеется такой элемент a^k, что $k = ql + r, \quad (0 < r < l)$. Имеем, $a^k = a^{ql+r} = a^{ql}a^r$, $a^{ql} \in H$; $a^{-ql} \in H, \quad a^{-ql}a^{ql}a^r \in H$, $a^r \in H$, и получается, что не l, а r есть наименьший из показателей в подгруппе H, чего быть не может.

Отсюда следует

Теорема 2.8.1. *Подгруппа циклической группы сама циклическая. Она состоит либо из единицы группы, либо из всех степеней элемента a^l, имеющего наименьший возможный положительный показатель l, (где a — элемент, порождающий всю исходную группу G.)*

Пусть a — элемент конечного порядка n, то есть $a^n = 1$, Тогда порядок группы есть n, и n должно делиться на l. Действительно, так как 1 принадлежит подгруппе H, и так как $1 = a^n$, то a^n также принадлежит подгруппе H. А так как все элементы подгруппы H суть степени элемента a^l, то n кратно l.

Если же a — элемент бесконечного порядка, то и группа H имеет бесконечный порядок и состоит из элементов

$$a^{\pm l},\ a^{\pm 2l},\ \ldots\ , a^{\pm il}, \ldots$$

Следовательно, в случае группы бесконечного порядка число l произвольно, а в случае группы конечного порядка n каждому его делителю l соответствует циклическая подгруппа $\{a^l\}$ порядка $q = n/l$ циклической группы G.

Элемент a называется порождающим или первообразным элементом группы G.

Вспомним пример с первообразными корнями по модулям

$m = 2,\ 4,\ p^\alpha,\ 2p^\alpha, \quad$ т.е., например, по модулям $\quad m = 9,\ 11,\ 18.$

Следовательно, приведённые системы вычетов по этим модулям оказываются циклическими группами, а делители числа $\varphi(m)$ суть порядки их подгрупп..

Рассмотрим, какие элементы циклической группы также являются первообразными.

Пусть опять a первообразный элемент, и $b = a^m$, где m не только не делит n, но и $(m, n) = 1$.

Если бы элемент $b = a^m$ принадлежал какой-нибудь истинной подгруппе H группы G, то m было бы кратно какому-нибудь делителю l числа n, что невозможно из-за $(m, n) = 1$.

Отсюда следует, что все степени

$$b^k,\ k = 1, 2, \ldots, n-1$$

различны, так как в противном случае элемент $b = a^m$ порождает истинную подгруппу H, которой и принадлежит, вопреки доказанному выше.

Иначе говоря, вместе с элементом a порождающими элементами группы будут все такие степени a^m, что $(m, n) = 1$.

Этот же факт устанавливается следующим простым рассуждением. Все степени $b^j = a^{jm}$ таковы, что благодаря условию $(m, n) = 1$, показатели jm пробегают вместе с j полную систему вычетов по модулю n (см. теорему 1.6.2). [1]

Если же $(m, n) = d > 1$, то $m = m_1 d$ и $(m_1, n) = 1$. Тогда, положив $b = a^{m_1}$, получим что b — первообразный элемент, и $b^d = a^{m_1 d} = a^m$ — суть d-я степень порождающего элемента a. А она по доказанному выше порождает циклическую подгруппу порядка $n/m_1 = d$.

Таким образом, первообразных элементов циклической группы порядка n имеется ровно столько, сколько чисел, взаимнопростых с n, имеется среди чисел $1, 2, \ldots, n-1$, т.е. $\varphi(n)$.

Покажем теперь, что для каждого делителя l порядка n циклической группы G существует *только одна* подгруппа H порядка n/l, хотя, быть может, делитель l входит в число n как степень l^x.

Действительно, предположим противное, и пусть имеется еще одна подгруппа H', состоящая из $l-$х степеней элемента $b = a^m$, где $(m, n) = 1$, и потому b вместе с a является также порождающим элементом группы G. Иначе говоря, пусть

$$H = \{a^l,\ a^{2l},\ \ldots, a^{\delta l}\}$$

и

$$H' = \{(a^m)^l,\ (a^m)^{2l},\ \ldots, (a^m)^{\delta l}\}$$

[1] Подчеркнем, что последнее чисто "теоретико-числовое" рассуждение не было бы возможным, не будь здесь главы 1, которой нет в большинстве руководств по теории групп.

две циклические подгруппы порядка $\delta = n/l$. Переписав

$$H' = \{a^{ml}, a^{2ml}, \ldots, a^{\delta ml}\},$$

заметим, что из-за $(m, n) = 1$, последовательность m, $2m, \ldots \delta m$, по теореме 1.6.2 пробегает полную систему вычетов по модулю δ. Пусть $im \equiv j(\mathrm{mod}\delta)$, где j — это одно из чисел 1, 2, ... , δ, составляющих полную систему наименьших положительных вычетов по модулю δ. Имеем $im = j + t\delta$.

Отсюда $a^{iml} = a^{(j+t\delta)l} = a^{jl}a^{t\delta l} = a^{jl}a^{tn} = a^{jl}(a^n)^t = a^{jl}$, а это означает, что подгруппы H и H' отличаются только порядком следования их элементов, а потому совпадают.

П р и м е р 2. 3.

Пусть группа G есть циклическая группа порядка $n = 63$, и пусть β — её порождающий (первообразный, образующий) элемент. Числа 3, 7, 9, 21 — суть делители порядка группы G. Подгруппа $H_3 \subset G$ порождается элементом β^3. Имеем

$$H_3 = \{\beta^3, \beta^6, \beta^9, \beta^{12}, \beta^{15}, \beta^{18}, \beta^{21}, \beta^{24}, \beta^{27}, \beta^{30},$$

$$\beta^{33}, \beta^{36}, \beta^{39}, \beta^{42}, \beta^{45}, \beta^{48}, \beta^{51}, \beta^{54}, \beta^{57}, \beta^{60}, \beta^{63} = 1\}.$$

В свою очередь

$$H_{21} \subset H_3,\ H_9 \subset H_3, \text{ где } H_9 = \{\beta^9, \beta^{18}, \beta^{27}, \beta^{36}, \beta^{45}, \beta^{54}, \beta^{63} = 1\},$$

$$H_{21} = \{\beta^{21}, \beta^{42}, \beta^{63} = 1, \}.$$

Далее

$$H_{21} \subset H_7 \subset G,\ H_7 = \{\beta^7, \beta^{14}, \beta^{21}, \beta^{28}, \beta^{35}, \beta^{42}, \beta^{49}, \beta^{56}, \beta^{63} = 1\}.$$

В четырёх подгруппах содержится 27 различных элементов. Так как $\varphi(63) = 36$, то остальные 36 элементов являются порождающими (первообразными, образующими) элементами группы G, и потому никакой истинной подгруппе принадлежать не могут.

На случай приведенных систем вычетов это означает, что первообразных элементов (корней) циклической группы приведенной системы вычетов по модулю m имеется

$$\varphi(\varphi(m)),$$

где $m = 2$, 4, p^α, $2p^\alpha$.

2.9. Смежные классы. Разложение группы по подгруппе

Пусть $H = \{h_1, h_2, \ldots, h_i, \ldots\}$ подгруппа группы G, и $a \in G$.

Определение 2.9.1. *Множество $aH = \{ah_1, ah_2, \ldots, ah_i, \ldots\}$ называется левосторонним (левым) смежным классом группы G, по подгруппе H. Множество Ha называется правосторонним, (правым) смежным классом группы G, по подгруппе H.*

Рассмотрим последовательность смежных классов:

$$a_0H, a_1H, a_2H, \ldots, a_jH, \ldots$$

где

$$a_0 \in H, a_i \notin H,\ i = 1, 2, \ldots$$

Утверждение 2.9.2. *Всякий смежный класс определяется любым своим элементом.*

Действительно, пусть

$$aH = \{ah_1, ah_2, \ldots, ah_i, \ldots\}. \tag{2.9.3}$$

Возьмём произвольный элемент $ah_j \in aH$. Тогда

$$ah_jH = \{ah_jh_1, ah_jh_2, \ldots, ah_jh_i, \ldots\} \tag{2.9.4}$$

Так как $h_jh_i \in H$, и так как по определению группы разрешимо уравнение $h_jx = h_k$, т.е. для любого $h_k \in H$ найдётся такое $h_i \in H$, что $h_k = h_j^{-1}h_i$, то правые части в (2.9.3) и (2.9.4) совпадают с точностью до порядка следования (перестановки) элементов. Отсюда следует

Утверждение 2.9.3. *Смежные классы либо не пересекаются, либо совпадают. Это значит, что при заданной подгруппе $H \subset G$ каждый элемент $a \in G$ принадлежит в точности одному смежному классу.*

Вся группа G распадается на непересекающиеся смежные классы по подгруппе H.

$$G = a_0H \cup a_1H \cup a_2H \cup \ldots \cup a_jH \cup \ldots,$$

Утверждение 2.9.4. *Все смежные классы равномощны,*

так как соответствием $ah_i \Leftrightarrow bh_i$, имеющим место для каждого i, устанавливается взаимно однозначное отображение aH в bH.

Термин "равномощны" применяется в том смысле, что смежные классы имеют одинаковый порядок, если они конечны, или имеют одинаковую мощность, если они бесконечны. Например, бесконечная группа целых чисел имеет подгруппу чисел, кратных числа m, а конечное множество классов вычетов по модулю m есть разложение группы целых чисел по упомянутой подгруппе. Каждый класс вычетов есть смежный класс, и он бесконечен. Все эти смежные классы имеют одинаковую мощность: они счетны.

Утверждение 2.9.5. *Два элемента a и b принадлежат одному и тому же смежному классу тогда и только тогда, когда $a^{-1}b \in H$.*

Доказательство. Пусть $a^{-1}b \in H$, тогда это значит, что

$$a(a^{-1}b) = b \in aH,$$

т.е. b принадлежит смежному классу aH, определяемому элементом a.

Обратно, пусть $h_i \in H$. Положим $b = ah_i$, т.е. b принадлежит смежному классу aH, определяемому элементом a. (Это и означает, что a и b принадлежат одному и тому же смежному классу.) Тогда

$$a^{-1}b = a^{-1}ah_i = h_i \in H,$$

т.е. $a^{-1}b \in H$.

Утверждение 2.9.6. *За исключением самой подгруппы H смежные классы по ней не являются группами.*

Действительно, если $a \in G$, но $a \notin H$, то и $a^{-1} \notin H$. Поэтому и $1 \notin aH$.

Утверждение 2.9.7. *Для произвольной подгруппы H элементы, обратные к элементам левого смежного класса, образуют правый смежный класс.*

Действительно, пусть

$$a \in G, \quad b \in H, \quad \text{т.е.} \quad b^{-1} \in H, \quad \text{и} \quad ab \in aH.$$

Тогда, так как

$$(ab)^{-1} = b^{-1}a^{-1}, \quad \text{то из-за} \quad b^{-1} \in H, \quad \text{имеем} \quad b^{-1}a^{-1} \in Ha^{-1}.$$

Иначе говоря, если

$$ab \in aH, \quad \text{то} \quad (ab)^{-1} \in Ha^{-1},$$

что и требовалось.

Определение 2.9.8. *Число различных смежных классов в разложении группы G по подгруппе H называется индексом подгруппы H в группе G. Он может быть конечным и бесконечным. Если число смежных классов конечно, то H называется подгруппой конечного индекса.*

Обозначив порядок (конечной!) группы G символом n, порядок и индекс подгруппы H соответственно j и i, получим, что справедлива

Теорема 2.9.9 (Лагранж).

$$n = ji.$$

Отсюда следует

Утверждение 2.9.10. *— порядок подгруппы конечной группы есть делитель порядка группы.*

— так как порядок элемента совпадает с порядком порождаемой им циклической подгруппы, то порядок элемента конечной группы есть делитель порядка группы.

— группа, порядок которой есть простое число, является циклической группой, так как порядок любого её элемента не может быть собственным делителем её (простого) порядка, и, значит, совпадает с порядком группы.

Заметим попутно, что если циклическая группа G порождается элементом a, и её подгруппа H порождается элементом a^l, то число l в формулировке теоремы 2.8.1 есть индекс подгруппы H в группе G.

Вспоминая, что вычеты приведенной системы вычетов по модулю m образуют мультипликативную группу, и что показатель, которому принадлежит вычет приведенной системы по модулю m, на языке теории групп называется порядком элемента, видим, что утверждение 1.11.4 есть простое следствие теоремы Лагранжа.

Утверждение 2.9.11. *Порядок любого элемента конечной абелевой группы есть делитель показателя группы.*

Доказательство. Пусть $\varepsilon-$ показатель группы, и $a^{\varepsilon} = 1$. Пусть $\nu-$ порядок элемента b, т.е. $b^{\nu} = 1$. Тогда порядок l элемента ab есть $l = m(\varepsilon, \nu) \geq \varepsilon$. Но ab есть элемент группы, и его порядок l не может быть больше показателя группы по определению последнего. Отсюда порядок $l \leq \varepsilon$. Значит, порядок $l = \varepsilon$. Поэтому и $m(\varepsilon, \nu) = \varepsilon$, и ν делит ε, что и требовалось.

2.10. Нормальные делители

Определение 2.10.1. *Подгруппа H называется нормальным делителем или инвариантной подгруппой, когда все левые смежные классы являются одновременно и правыми смежными классами, т.е., когда*

$$aH = Ha, \tag{2.10.5}$$

для всех a.

Важно подчеркнуть, что (2.10.5) вовсе не означает, что для всякого $h \in H \quad ah = ha$. Если для коммутативной группы это действительно так, то в некоммутативном случае это означает лишь, что для каждого $h_1 \in H$ найдется такое $h_2 \in H$, что $ah_1 = h_2a$. Если групповая операция не коммутативна, то соотношение (2.10.5) означает "перестановочность".

В абелевой группе любая подгруппа — нормальный делитель.

Утверждение 2.10.2. *Если H нормальный делитель, то произведение смежных классов есть снова смежный класс.*

Действительно, так как операция ассоциативна, то

$$aHbH = a(Hb)H = a(bH)H = abHH = abH,$$

что и требовалось.

Утверждение 2.10.3. *Если произведение любых двух смежных классов в разложении группы G по подгруппе H есть снова смежный класс, то подгруппа H есть нормальный делитель.*

Д о к а з а т е л ь с т в о. Пусть $aHbH = cH; \quad c \in G$. Тогда после умножения этого равенства на a^{-1} слева получим. $HbH = a^{-1}cH$, и $H(bH)$ есть левый смежный класс, который содержит смежный класс bH, а значит, совпадает с ним, т.е. $HbH = bH$. Но $(Hb)H$ содержит также и правый смежный класс Hb, и потому совпадает с ним.

Значит, $bH = Hb$, и H есть нормальный делитель.

Утверждение 2.10.4. *Если H есть нормальный делитель, то каждому смежному классу в разложении группы G по подгруппе H есть "обратный", т.е. такой xH (или Hx), что $aHxH = xHaH = H$.*

Действительно, если $ax = 1$, т.е. $x = a^{-1}$, то

$$aHxH = axHH = axH = H.$$

Или, что то же, если $xa = 1$, т.е. $x = a^{-1}$, то

$$HaHx = HHax = Hax = H.$$

Объединяя утверждения 2.10.2, 2.10.3 и 2.10.4, получим:

Утверждение 2.10.5. *Множество смежных классов по нормальному делителю H есть группа. Ее порядок равен индексу подгруппы H в группе G. Эта группа обозначается символом G/H и называется фактор-группой группы G по нормальному делителю H.*

2.11. Изоморфизм групп

Определение 2.11.1. *Изоморфизм двух групп G и $\overline{G}$ есть такое взаимно однозначное соответствие $a \Leftrightarrow \overline{a}$, $a \in G$, $\overline{a} \in \overline{G}$, при котором из того, что $b \Leftrightarrow \overline{b}, \quad c \Leftrightarrow \overline{c}$, и $bc = d, \quad \overline{b}\overline{c} = \overline{d}$, следует $d \Leftrightarrow \overline{d}$.*

Примеры изоморфизма групп:

1. Циклическая мультипликативная группа $1, a^{\pm 1}, a^{\pm 2}, \ldots$ изоморфна аддитивной группе всех целых чисел.

2. Все циклические группы $1, a, a^2, \ldots, a^n$ одного порядка изоморфны друг другу.

3. Аддитивная группа всех действительных чисел изоморфна мультипликативной группе всех (отличных от нуля) положительных вещественных чисел. Изоморфизм устанавливается соответствием $\log a \Leftrightarrow a$. Имеем:

$$0 < a \in \mathbb{R}, \quad 0 < b \in \mathbb{R}.$$

Пусть

$$\log a = \overline{a} \in \mathbb{R}, \quad \log b = \overline{b} \in \mathbb{R},$$

$$a \Leftrightarrow \overline{a}, \quad b \Leftrightarrow \overline{b}.$$

С одной стороны,

$$\log(ab) = \overline{ab} \in \mathbb{R}.$$

С другой —

$$\log(ab) = \log a + \log b = \overline{a} + \overline{b}.$$

Отсюда

$$ab \Leftrightarrow \overline{a} + \overline{b}.$$

Обозначение изоморфизма:

$$G \cong \overline{G}.$$

Теорема 2.11.2 (Кэли). *Всякая конечная группа порядка n изоморфна некоторой подгруппе симметрической группы степени n.*

Действительно, пусть группа G имеет порядок n, и пусть элементы этой группы записаны в определенном порядке:

$$a_1, a_2, \ldots, a_n.$$

Если b есть произвольный элемент группы G, то все произведения $a_i b = a_{\beta_i} \quad (i = 1, 2, \ldots, n.)$ различны между собой, т.е. система $a_{\beta_1}, a_{\beta_2}, \ldots a_{\beta_n}$, есть перестановка (сдвиг) элементов группы G. Это ставит в соответствие элементу b группы G перестановку

$$\begin{pmatrix} 1, 2, \ldots, n \\ \beta_1, \beta_2, \ldots, \beta_n \end{pmatrix}.$$

Значит, и каждому элементу группы G ставится в соответствие вполне определенная подстановка n-й степени.Двум различным элементам соответствуют различные подстановки, так как в противном случае из $a_1 b = a_1 b'$ следовало бы $b = b'$.

Найдем подстановку, отвечающую элементу bc, $\quad c \in G$.

Если элементу c отвечает подстановка

$$\begin{pmatrix} \beta_1, \beta_2, \ldots, \beta_n \\ \gamma_1, \gamma_2, \ldots, \gamma_n \end{pmatrix},$$

т.е. $a_{\beta_i} c = a_{\gamma_i}$, то из

$$a_i(bc) = a_{\beta_i} c = a_{\gamma_i}$$

следует, что элементу bc отвечает подстановка

$$\begin{pmatrix} 1, 2, \ldots, n \\ \gamma_1, \gamma_2, \ldots, \gamma_n \end{pmatrix} = \begin{pmatrix} 1, 2, \ldots, n \\ \beta_1, \beta_2, \ldots, \beta_n \end{pmatrix} \begin{pmatrix} \beta_1, \beta_2, \ldots, \beta_n \\ \gamma_1, \gamma_2, \ldots, \gamma_n \end{pmatrix}.$$

Из теоремы 2.11.2 и из очевидного утверждения, что конечная группа может обладать лишь конечным числом подгрупп, следует, что существует лишь конечное число неизоморфных конечных групп данного порядка n.

Следовательно, множество всех неизоморфных конечных групп, являясь суммой счетного множества конечных множеств, само счетно.

Если G и $\overline{G}$ совпадают, то это есть взаимно однозначное сопоставление элементам a элементов $\overline{a}$ той же группы, сохраняющее групповую операцию. Такое сопоставление называется *автоморфизмом.*

2.12. Гомоморфизм групп

Определение 2.12.1. *Пусть в двух множествах M и $\overline{M}$ определены некоторые соотношения между элементами, и пусть каждому элементу $a \in M$ поставлен в соответствие* **один и только один** *элемент $\overline{a} \in \overline{M}$ таким образом, что*

1) *Каждому элементу $\overline{a} \in \overline{M}$ отвечает* **по крайней мере один** *элемент $a \in M$,*

2) *Все соотношения между элементами множества M выполняются и для соответствующих элементов множества $\overline{M}$.*

Такое соответствие называется гомоморфизмом. Говорят также, что множество M гомоморфно отображается на множество $\overline{M}$.

Элемент $\overline{a}$ есть гомоморфный образ элемента a, и элемент a есть прообраз элемента $\overline{a}$. Гомоморфным будет соответствие между множеством $M = \mathbb{Z}$ целых чисел и множеством $\overline{M}$ классов вычетов по модулю m, если каждому числу поставить в соответствие класс вычетов, которому оно принадлежит.

Теорема 2.12.2. *Гомоморфный образ $\overline{G}$ группы G есть группа.*

Иначе говоря, если в множестве

$$\overline{G}$$

определены произведения

$$\overline{a}\overline{b} = \overline{c},$$

и группа

$$G$$

гомоморфно отображается на

$$\overline{G},$$

то

$$\overline{G}$$

также есть группа.

Д о к а з а т е л ь с т в о. Если по гомоморфизму

$$a \Rightarrow \overline{a}, b \Rightarrow \overline{b}, c \Rightarrow \overline{c},$$

то

1)

$$((ab)c = a(bc)) \Rightarrow ((\overline{a}\overline{b})\overline{c} = \overline{a}(\overline{b}\overline{c})),$$

т.е. в

$$\overline{G},$$

выполняется ассоциативность операции.

2) Замкнутость относительно операции следует из гомоморфизма.

3) Из того, что для всех $a \in G$ выполняется $a \cdot 1 = a$, следует, что для всех $\overline{a} \in \overline{G}$ выполняется $\overline{a} \cdot \overline{1} = \overline{a}$.

Иначе говоря, из существования единицы в

$$G$$

следует существование единицы в

$$\overline{G}.$$

4)

$$(ba = 1) \Rightarrow (\overline{b}\overline{a} = \overline{1})$$

Иначе говоря, из существования обратного элемента $b = a^{-1}$ в G следует существование обратного элемента $\overline{b} = \overline{a}^{-1}$ в $\overline{G}$.

Согласно утверждениям 2.10.2, 2.10.3 и 2.10.4 множество смежных классов группы по подгруппе замкнуто относительно групповой операции, и в нем выполняется обратная операция тогда и только тогда, когда подгруппа есть нормальный делитель. Это же утверждение получается из теоремы 2.12.2 о гомоморфизме, как её частный случай:

Множество смежных классов группы по подгруппе само будет группой тогда и только тогда, когда подгруппа есть нормальный делитель. Гомоморфизм устанавливается тем, что каждому элементу группы ставится в соответствие тот смежный класс, которому этот элемент принадлежит.

Вспомним, что в утверждении 2.10.5 группа смежных классов по нормальному делителю была названа фактор-группой.

Элементы группы, которые отображаются в единицу фактор-группы, называются ядром гомоморфизма. Все гомоморфизмы группы исчерпываются её фактор-группами, и, таким образом, ядрами гомоморфизмов являются нормальные делители. Любая группа, гомоморфная данной, изоморфна некоторой её фактор-группе.

2.13. Несколько замечаний

а) В абелевой группе каждая подгруппа — нормальный делитель.
б) Если групповая операция есть сложение, то группы и их подгруппы принято называть модулями.
в) Пусть G модуль и M его подмодуль. Смежные классы $a+M$, называются классами вычетов по модулю M, а фактор-группа

$\overline{G} = G/M$ называется фактор-модулем модуля G по подмодулю M.

г) Два элемента a, b лежат в одном смежном классе или классе вычетов, если их разность лежит в M. Такие два элемента называются сравнимыми по модулю M.

Это записывается следующим образом:

$$a \equiv b (\mathrm{mod}\, M). \tag{2.13.6}$$

Рассмотрим модуль классов вычетов, т.е. совокупность $a_1 + M$, $a_2 + M$, $\ldots$, $a_i + M, \ldots$ Пусть $\overline{a}$ и $\overline{b}$ принадлежат этому модулю и по гомоморфизму соответствуют элементам a, b. Тогда из (2.13.6) следует

$$\overline{a} = \overline{b}. \tag{2.13.7}$$

Наоборот, из (2.13.7) следует (2.13.6).

Из (2.13.6) и (2.13.7) получается, что при гомоморфизме сравнения переходят в равенства.

Сказанное выше – это чисто абстрактное теоретико-групповое рассуждение, безотносительно к какой бы то ни было конкретной интерпретации.

Выразительной и знакомой нам интерпретацией является множество $\mathbb{Z}$ целых чисел. Это аддитивная абелева группа, т.е. модуль. Множество всех чисел, кратных числа m, есть ее подмодуль. Обозначим его M. В главе 1 введена запись

$$a \equiv b (\mathrm{mod}\, m),$$

если $m|(a-b)$. Знакомое нам множество классов вычетов по модулю m является модулем, точнее фактор-модулем модуля $\mathbb{Z}$ всех целых чисел по подмодулю M. Классы вычетов по модулю m, т.е.элементы фактор-модуля, могут быть представлены числами $0, 1, \ldots, m-1$, и фактор-модуль есть циклическая группа порядка m.

2.14. Кольцо

Определение 2.14.1. *Кольцом называется такая система элементов с определёнными в ней сложением $a+b$ и умножением $a \cdot b$, что:*

1) *По сложению она — абелева группа.*

2) *Умножение ассоциативно.*

3) *Сложение и умножение связаны законами дистрибутивности:*

$$a(b+c) = ab + ac;\ (b+c)a = ba + ca.$$

Таким образом, в отличие от группы, кольцо — это множество с двумя операциями. Мы будем рассматривать только кольца, коммутативные по умножению. Поэтому достаточно только одного закона дистрибутивности.

Примером кольца служит множество целых чисел. Кольцом является множество классов вычетов по модулю m, так как это множество есть аддитивная группа, и произведение двух классов есть снова класс, т.е. умножение классов определено.

Закон дистрибутивности имеет место также и для вычитания, так как в кольце вычитание имеет место, как в группе по сложению. Действительно, выясним, чему равно $a(b-c)$? Вследствие дистрибутивности сложения

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

откуда получаем

$$a(b-c) = ab - ac.$$

Воспользуемся этим фактом для определения, что такое умножение на нуль, так как нуль имеется в аддитивной группе кольца, и умножение элементов постулировано. Поступим следующим образом:

$$a \cdot 0 = a(a-a) = aa - aa = 0.$$

Иначе говоря, если один из сомножителей равен нулю, то произведение равно нулю.

Может случиться, что обратное неверно, т.е. из того, что произведение равно нулю, необязательно следует равенство нулю хотя бы одного сомножителя. В качестве примера рассмотрим кольцо классов вычетов по составному модулю $m = ab$. То, что множество классов вычетов — аддитивная группа, отмечено выше. Перемножим два числа

$$(mt_1 + a)(mt_2 + b) = mT + ab = mT + m = mT',$$

и мы получили число, кратное модуля, т.е. число, принадлежащее нулевому классу вычетов. Таким образом, в кольце классов вычетов по составному модулю $m = ab$ два класса вычетов $\{a\}$ и $\{b\}$ являются делителями нуля: $\{a\}\{b\} = \{m\} = 0$.

Определение 2.14.2. *Кольцо без делителей нуля называется областью целостности.*

Кольцо не обязано иметь единицы, так как от него не требуется быть мультипликативной группой. Кольцо называется кольцом с единицей, если она есть, и — кольцом без единицы, если её нет. Если кольцо с единицей, то $\underbrace{1+1+\ldots+1}_{n \text{ слагаемых}} = n$, и, таким образом, n есть элемент кольца.

2.15. Поле

Определение 2.15.1. *Если отличные от нуля элементы коммутативного кольца образуют группу по умножению, то это кольцо называется полем. (В некоммутативном случае — телом.)*

Теорема 2.15.2. *Поле не имеет делителей нуля.*

Доказательство. Пусть $a, b \neq 0$, но $ab = 0$. В поле каждый ненулевой элемент имеет обратный. Для a это будет a^{-1}. Умножим на него обе части равенства $ab = 0$. Получим $a^{-1}ab = 0, \quad b = 0$, что противоречит условию.

Теорема 2.15.3. *Конечная область целостности есть поле.*

Доказательство. Пусть $a \neq 0$, и в произведении ax элемент x пробегает все ненулевые элементы кольца. Покажем, что различным x отвечают различные ax. Действительно, предположим противное, т.е. пусть при

$$x_1 \neq x_2 \tag{2.15.8}$$

имеет место

$$ax_1 = ax_2. \tag{2.15.9}$$

Тогда $a(x_1 - x_2) = 0$. Значит, так как $a \neq 0$, и делителей нуля нет, то $x_1 - x_2 = 0$, откуда $x_1 = x_2$, что противоречит условию (2.15.8). Значит, множество ненулевых элементов кольца отображается на себя. Но в силу конечности кольца такое отображение может быть только взаимно однозначным.

Иначе говоря произведение ax пробегает в точности по одному разу все ненулевые элементы кольца. Следовательно, каждый ненулевой элемент b может быть представлен в виде $b = ax$, что означает разрешимость уравнения (2.2.1). Остаётся вспомнить определение 2.3.1 группы. (Ср. с доказательствами теорем 1.6.2 и 1.7.2).

Требование конечности существенно: кольцо целых чисел не имеет делителей нуля, но полем не является.

Утверждение 2.15.4. *Кольцо классов вычетов по модулю m будет полем тогда и только тогда, когда m простое число.*

Доказательство. То, что кольцо по составному модулю содержит делители нуля и потому не может быть полем, доказано выше.

Так как модуль есть простое число, то кольцо классов вычетов по этому модулю не содержит делителей нуля и является конечной областью целостности, что и требовалось.

Поле классов вычетов по простому модулю p будем обозначать символом $GF(p)$. Иначе говоря, полем $GF(p)$ является полная система неотрицательных вычетов по модулю p, и операции сложения и умножения чисел $0, 1, 2, \ldots, p-1$ выполняются по модулю p. В связи с доказанными фактами снова уместно вспомнить теорему 1.7.2 о приведенной системе вычетов.

Примеры полей. Поле вещественных чисел $\mathbb{R}$, поле рациональных чисел $\mathbb{Q}$, поле комплексных чисел $\mathbb{C}$. В следующей главе подробно изучаются конечные поля — поля Галуа. Они играют важнейшую роль в теории помехоустойчивого кодирования.

2.16. Идеал

Непустое подмножество v кольца V само будет кольцом тогда и только тогда, когда

1. v — есть подгруппа аддитивной группы кольца, т.е. когда для любых

$$a, b \in v : a + b \in v, \quad a - b \in v,$$

2. Для любых

$$a, b \in v : ab \in v.$$

Среди подколец особую роль играют *идеалы*.

Определение 2.16.1. *Идеалом называется такое подкольцо v кольца V, что*

$$(a \in v \quad u \quad r \in V) \Rightarrow (ar \in v).$$

Примеры идеалов. Нулевой идеал (0), состоящий из одного элемента 0. Единичный идеал (e), состоящий из всего кольца V.

Идеал (a), порождённый элементом a, и состоящий из всевозможных выражений

$$ra + na, \quad r \in V, \quad n - \text{целое число}.$$

Рассмотрим существо этой конструкции. Положим, что элемент a принадлежит идеалу. Тогда по определению идеала элемент ra, $r \in V$, также должен принадлежать идеалу. Кроме того, так как идеал есть кольцо, то элемент a, повторенный слагаемым сколько угодно раз, также должен принадлежать идеалу. Значит, вместе с элементом a идеалу должны принадлежать все элементы вида

$$ra + na, \quad r \in V, \quad n - \text{целое число}.$$

Проверим, что построенная таким образом конструкция действительно есть идеал. То, что (a) есть кольцо, проверяется непосредственно, так как разность двух элементов такого вида есть снова элемент такого вида:

$$(r_1a + n_1a) - (r_2a + n_2a) = (r_1 - r_2)a + (n_1 - n_2)a = r_3a + n_3a.$$

То, что данное подкольцо удовлетворяет определению идеала, проверяется следующим образом:

Пусть $s \in V$. Тогда $s(ra + na) = (sr + ns)a$. т.е. имеет вид $r^{'}a = r^{'}a + 0 \cdot a$.

Вспомним, что означает na для колец без единицы и колец с единицей. Если кольцо с единицей, то $n \cdot 1 \in V$, т.е. n является элементом кольца. Тогда

$$ra + na = (r + n \cdot 1)a = r_1a,$$

и идеал состоит из обыкновенных кратных. Например, идеал (5) в кольце целых чисел состоит из всех чисел кратных 5.

Идеал, порождённый одним элементом, называется главным. Идеал (0) — всегда главный. Кольцо, в котором все идеалы главные, называется кольцом главных идеалов. Совокупность всех кратных ra элемента a есть идеал. Нетрудно показать что этот идеал может не совпадать с главным идеалом (a).

Для этого в качестве примера рассмотрим кольцо четных чисел $V = 2i, (i = \pm 1, \pm 2, ... \pm j...)$. В этом кольце все числа, кратные некоторого $a = 2i$, составляют идеал и имеют вид $2i2j = 4ij$, т.е. делятся на 4. С другой стороны, главный идеал

$$ra + na = 2i2 + n2; n = 1, 2, ...$$

содержит числа, которые при нечетном n не делятся на 4.

В дальнейшем будут рассматриваться главные идеалы как кратные.

Утверждение 2.16.2. *Пересечение*

$$v_0 = v_1 \cap v_2 \cap \ldots \cap v_n$$

идеалов

$$v_i,\ i = 1, 2, \ldots n$$

есть идеал.

Читатель легко справится с доказательством самостоятельно, используя метод доказательства утверждения 2.7.1.

Кроме нулевого и целого идеала, поле не содержит других идеалов. Действительно, если $a \neq 0$ принадлежит идеалу, то так как в поле имеется и a^{-1}, идеалу принадлежит и $aa^{-1} = 1$, а значит любой элемент, кратный единицы, т.е. все элементы поля.

2.17. Линейное векторное пространство

Определение 2.17.1. *Линейным векторным пространством над полем F называется множество V векторов, удовлетворяющее условиям:*

1) *Множество V является аддитивной абелевой группой.*
2) *Для любых $c \in F$ и $v \in V$ имеет место $cv \in V$.*
3) *Выполняются дистрибутивные законы, т.е.*

$$\text{если}\quad c \in F;\ u, v \in V, \quad \text{то}\quad c(u + v) = cu + cv,$$

$$\text{если}\quad c, d \in F;\ v \in V, \quad \text{то}\quad (c + d)v = cv + dv.$$

4) *Умножение ассоциативно, т.е.* $(cd)v = c(dv)$.

Элементы c, d поля F называются скалярами.

Подмножество векторов пространства V называется подпространством, если в нем выполняются условия определения 2.17.1 Пусть $A \subset V$ есть подпространство пространства V. Векторы $v_1, v_2, \ldots, v_k \in A$ называются линейно зависимыми над полем F, если найдутся такие не все равные нулю элементы $a_1, a_2, \ldots, a_k \in F$, что

$$a_1v_1 + a_2v_2 + \ldots + a_kv_k = 0.$$

В противном случае векторы $v_1, v_2, \ldots, v_k \in A$ называются линейно независимыми. Максимальное число k линейно независимых векторов подпространства A называется его размерностью, а сама совокупность этих векторов называется базисом подпространства.

2.18. Задачи к главе 2

2.1. Существуют две неизоморфные группы четвертого порядка. Доказать.
2.2. Каждая группа порядка, не превосходящего пяти, абелева. Доказать.
2.3. Если каждый элемент группы является обратным самому себе, то группа абелева. Доказать.
2.4. Если индекс подгруппы равен двум, то подгруппа есть нормальный делитель. Доказать.
2.5.Найти разложение циклической группы 10-го порядка по всем ее подгруппам.
2.6. Найти разложение бесконечной циклической группы, порожденной элементом x, по подгруппе, порожденной элементом x^3.
2.7. Если G – циклическая группа с порождающим элементом a, и ее подгруппа H порождена степенью a^l, то элементы 1, a, a^2, $\ldots$, a^{l-1} – представители различных смежных классов.
2.8. Пусть порядок элемента x некоторой группы равен pq, и $(p, q) = 1$. Доказать, что найдутся такие элементы u и v, для которых выполняются равенства: $x = uv = vu$, $u^p = 1$, $v^q = 1$.
2.9. Найти правое разложение симметрической группы S_3 по подгруппе, состоящей из двух элементов, e и транспозиции (1 2).
2.10. Пусть H_1 и H_2 две подгруппы конечной группы G порядков соответственно m_1 и m_2. Доказать, что множество H_1H_2

состоит из $m_1 m_2/d$ элементов, где d есть порядок пересечения подгрупп H_1 и H_2.
2.11. Что представляют собой разложения группы G по ее несобственным подгруппам.
2.12. Каждая циклическая группа порядка m изоморфна аддитивной группе классов вычетов по модулю m.
2.13. Найти все разложения группы 10-го порядка по всем ее подгруппам.
2.14. Сколько существует различных множеств представителей правого разложения группы 12-го порядка по ее подгруппе порядка 3?
2.15. Пусть H — произвольная подгруппа группы G, и N некоторый нормальный делитель группы G. Доказать, что HN является подгруппой группы G, причем $HN = NH$.
2.16. Доказать, что произведение конечного числа и пересечение произвольного множества нормальных делителей группы являются также ее нормальными делителями.
2.17. Построить поле, состоящее из трех элементов 0, +1, −1.
2.18. Совокупность всех кратных ra элемента a есть идеал. Показать (на примере кольца четных чисел), что этот идеал может не совпадать с главным идеалом (a).
2.19. Доказать теорему Вильсона: $(p-1)! \equiv -1 (\mathrm{mod} p)$.

Глава 3.

Конечные поля

3.1. Конечное поле как множество классов вычетов по модулю неприводимого многочлена

Оказалось плодотворным разбиение кольца целых чисел на классы вычетов по простому модулю p. Множество этих классов (см. теорему 1.7.2) образует поле.

Теперь рассмотрим множество $F[x]$ всех многочленов $f(x)$ всевозможных неотрицательных степеней с коэффициентами из поля $GF(p)$. В таком случае говорят о многочленах "над полем $GF(p)$". Введем в рассмотрение неприводимый над полем $GF(p)$ многочлен $p(x)$ степени m.

Определение 3.1.1. *Многочлен $p(x) = a_0 + a_1x + \ldots + a_mx^m$ называется неприводимым над полем $GF(p)$, если он не распадается на множители над этим полем.*

В множестве $F[x]$ неприводимый над полем $GF(p)$ многочлен $p(x)$ некоторой степени m играет роль, аналогичную той, которую играет некоторое простое число p в кольце целых чисел. Рассмотрим все остатки от деления многочленов из $F[x]$ на $p(x)$. Они имеют вид $b(x) = b_0+b_1x+\ldots+b_{m-1}x^{m-1}$, $b_i \in GF(p)$. Нетрудно посчитать, что всего таких остатков имеется в точности p^m.

Два многочлена из множества $F[x]$ называются сравнимыми по модулю многочлена $p(x)$, если при делении на $p(x)$ они дают одинаковый остаток. Таким образом, множество $F[x]$ распадается на непересекающиеся классы многочленов, сравнимых по модулю $p(x)$. Обозначим множество этих классов символом

$$F[x]_{/(p(x))}. \tag{3.1.1}$$

Очевидно, множество (3.1.1) замкнуто относительно сложения и вычитания над $GF(p)$, а выполняя умножение элементов из (3.1.1) по модулю неприводимого многочлена $p(x)$, легко убедиться, что оно замкнуто также относительно умножения. Следовательно, множество (3.1.1) есть кольцо.

Теорема 3.1.2. *Все остатки, которые получаются при делении на неприводимый многочлен $p(x)$, попарно несравнимы.*

Доказательство. Допустим противное, т.е. пусть для некоторых двух остатков $b_1(x) \neq b_2(x)$ выполняется соотношение $b_1(x) \equiv b_2(x)(\mathrm{mod} p(x))$.

Отсюда $b_1(x) - b_2(x) \equiv 0(\mathrm{mod} p(x))$. Так как степень разности в левой части этого сравнения не выше, чем $m-1$, оно возможно только тогда, когда $b_1(x) - b_2(x) = 0$, что противоречит условию $b_1(x) \neq b_2(x)$.

Теорема 3.1.3. *Кольцо* (3.1.1) *не имеет делителей нуля.*

Доказательство. Предположим противное, и пусть

$$(b_1(x) + p(x)d_1(x))(b_2(x) + p(x)d_2(x)) \equiv 0(\mathrm{mod} p(x)).$$

(Здесь в каждой из скобок в левой части стоит произвольный вычет класса, определяемого вычетом $b_1(x)$ в первой скобке и вычетом $b_2(x)$ – во второй).

Тогда $b_1(x)b_2(x) = p(x)D(x)$, а значит, $b_1(x)b_2(x) \equiv 0(\mathrm{mod}\ p(x))$, чего быть не может, так как благодаря неприводимости многочлена $p(x)$, он взаимно прост над $GF(p)$ с каждым из многочленов $b_1(x)$ и $b_2(x)$. Таким образом, кольцо (3.1.1) есть область целостности, а будучи конечным, оно есть поле.

В дальнейшем будем оперировать не целыми классами вычетов кольца (3.1.1), а их представителями, имеющими минимальную степень в своих классах. Разумеется, этими представителями являются сами остатки. Именно это множество остатков будем теперь обозначать символом (3.1.1). В силу всего сказанного множество (3.1.1) (в новом его понимании) также есть поле, в котором групповые операции выполняются по модулю многочлена $p(x)$.

Мультипликативную группу поля (3.1.1) обозначим символом

$$F[x]^*_{/(p(x))}. \tag{3.1.2}$$

Приведем более традиционное доказательство того, что элементы множества (3.1.2) образуют мультипликативную группу. Фиксируем произвольный многочлен $a(x) \in F[x]^*_{/(p(x))}$.

Теорема 3.1.4. *Если многочлен $b(x)$ пробегает множество* (3.1.2)*, то произведение $a(x)b(x)$ также пробегает множество* (3.1.2).

Д о к а з а т е л ь с т в о. Очевидно, что произведений $a(x)b(x) \neq 0$ столько же, сколько многочленов $b(x) \neq 0$, т.е. $p^m - 1$. Покажем, что если $b_1(x)$ и $b_2(x)$ несравнимы по модулю многочлена $p(x)$, то $a(x)b_1(x)$ и $a(x)b_2(x)$ также несравнимы по модулю $p(x)$. Предположим противное, т.е. пусть верно сравнение $a(x)b_1(x) \equiv a(x)b_2(x)(\mathrm{mod}p(x))$. Отсюда следует $a(x)(b_1(x) - b_2(x)) \equiv 0(\mathrm{mod}p(x))$.

Так как $(a(x), p(x)) = 1$, то $(b_1(x) - b_2(x)) \equiv 0(\mathrm{mod}p(x))$, и $b_1(x) \equiv b_2(x)(\mathrm{mod}p(x))$, что противоречит условию несравнимости $b_1(x)$ и $b_2(x)$.

Из теоремы 3.1.4 следует, что каков бы ни был многочлен $a(x) \in F[x]^*_{/(p(x))}$ найдется такой многочлен $b(x) \in F[x]^*_{/(p(x))}$, что $a(x)b(x) \equiv 1(\mathrm{mod}p(x))$.

Это означает, что для каждого $a(x) \in F[x]^*_{/(p(x))}$ найдется обратный ему многочлен $b(x) \in F[x]^*_{/(p(x))}$, и множество (3.1.2)действительно есть мультипликативная группа, каковой она и названа выше.

П р и м е р 3. 1.

Рассмотрим неприводимый над $GF(2)$ многочлен $p(x) = x^2 + x + 1$.

$F[x]_{/x^2+x+1} = \{0, 1, x, x+1\}$. $F[x]^*_{/x^2+x+1} = \{1,\ x,\ x+1\}$.

$x + 1 = x^{-1}$. $F[x]^*_{/x^2+x+1} = GF^*(2^2)$

П р и м е р 3. 2.

Рассмотрим неприводимые над $GF(2)$ многочлены

$$p(x) = x^3 + x + 1 \text{ и } p(x) = x^3 + x^2 + 1.$$

В обоих случаях, разумеется:

$$F[x]_{/p(x)} =$$

$$= \{0,\, 1,\, x,\, x+1, x^2,\, x^2+1,\, x^2+x,\, x^2+x+1\} = GF(2^3).$$

Но в первом случае

$$x^2+1 = x^{-1},\; x^2+x+1 = (x^2)^{-1},\; x^2+x = (x+1)^{-1},$$

а во втором

$$x^2+x = x^{-1},\; x^2 = (x+1)^{-1},\; x^2+x+1 = (x^2+1)^{-1}.$$

Таким образом, роль различных неприводимых многочленов, по модулям которых строится поле, состоит именно в том, что они по-разному "организуют" соотношения между элементами в мультипликативной группе: они по-разному создают пары взаимнообратных элементов. Само же "содержимое" поля зависит только от степени многочлена-модуля, так как она определяет максимальную степень остатка от деления. Легко проверить, что кроме рассмотренных неприводимых многочленов, других неприводимых многочленов над GF(2) второй и третьей степени нет.

3.2. Поле разложения многочлена $x^{p^m} - x$

Изложенная выше процедура называется расширением поля $GF(p)$. Если неприводимый многочлен $p(x)$ имеет степень m, то вместо $F[x]_{/(p(x))}$ пишут $GF(p^m)$, а вместо $F[x]^*_{/(p(x))}$ пишут $GF^*(p^m)$. Поле $GF(p^m)$ называется расширением степени m поля $GF(p)$. Группа $GF^*(p^m)$ называется мультипликативной группой поля $GF(p^m)$, и ее порядок равен $p^m - 1$. Так как порядок элемента α_i группы есть делитель порядка группы, то $\alpha_i^{p^m-1} = 1$. Это означает, что любой элемент группы $GF^*(p^m)$ является корнем уравнения $x^{p^m-1} - 1 = 0$, а все элементы поля $GF(p^m)$, включая $\alpha_i = 0$, являются корнями уравнения

$$x^{p^m} - x = 0. \tag{3.2.3}$$

Это означает, что

$$x^{p^m} - x = \prod_{i=0}^{p^m-1} (x - \alpha_i).$$

Говорят, что $GF(p^m)$ есть поле разложения двучлена в левой части уравнения (3.2.3).

Подчеркнем, что, каков бы ни был неприводимый многочлен $p(x)$, по модулю которого построено поле $GF(p^m)$, все элементы поля являются корнями одного и того же уравнения $x^{p^m} - x = 0$.

Исключив из рассмотрения $\alpha_0 = 0$, получим двучлен

$$x^{p^m-1} - 1 = \prod_{i=1}^{p^m-1} (x - \alpha_i), \tag{3.2.4}$$

корнями которого являются корни из единицы.

3.3. Цикличность мультипликативной группа поля

Определение 3.3.1. *Характеристикой поля называется такое наименьшее целое число n, что*

$$\underbrace{1 + 1 + \ldots + 1}_{n \text{ слагаемых}} = 0. \tag{3.3.5}$$

Если это число есть нуль, то поле называют полем характеристики нуль. В противном случае имеют дело с полем конечной характеристики. Нетрудно показать, что если характеристика не равна нулю, то она есть простое число p, так как в поле нет делителей нуля.

Покажем, что уравнение (3.2.3) не имеет кратных корней. Действительно, производная от его левой части равна [1]

$$p^m x^{p^m-1} - 1 = -1,$$

так как $p^m \equiv 0 (\mathrm{mod}\, p)$. Она не обращается в нуль, и потому не имеет общих корней с многочленом в (3.2.3).

Для дальнейшего рассмотрим

П р и м е р 3. 3.

Пусть поле $GF(2^3)$ построено по модулю многочлена $p(x) = x^3 + x + 1$. Возведем элемент $x \in GF(2^3)$ в последовательные

[1] Дабы избежать недоразумения, следует принять во внимание, что (3.3.5) относится только к элементам поля и не относится к показателям степеней при них.

степени, помня, что каждую степень x^i следует разделить на $p(x)$ и взять остаток от деления:

$$\begin{array}{l} x^0 = 1, \\ x^1 = x, \\ x^2 = x^2, \\ x^3 = 1 + x, \\ x^4 = x + x^2, \\ x^5 = 1 + x + x^2, \\ x^6 = 1 + x^2, \\ x^7 = x + x^3 = x + 1 + x = 1. \end{array} \tag{3.3.6}$$

На случай, когда модуль есть $p(x) = x^3 + x^2 + 1$, получим:

$$\begin{array}{l} x^0 = 1, \\ x^1 = x, \\ x^2 = x^2, \\ x^3 = 1 + x^2, \\ x^4 = x + x^3 = x^2 + x + 1, \\ x^5 = 1 + x + x^2 + x^3 = 1 + x, \\ x^6 = x + x^2, \\ x^7 = x^3 + x^2 = 1. \end{array} \tag{3.3.7}$$

Обе группы (3.3.6) и (3.3.7) оказались циклическими, и в каждой из них элемент x является порождающим.

В общем случае справедлива

Теорема 3.3.2. Группа $GF^*(p^m)$ циклична.

Д о к а з а т е л ь с т в о. Пусть $n = p^m - 1 = p_1^{\alpha_1} p_2^{\alpha_2} ... p_k^{\alpha_k}$

Рассмотрим уравнение $x^{n/p_i} = 1, \quad i = 1, 2, ..., k$. Оно имеет не более, чем n/p_i решений.

Значит, имеется не более, чем $n/p_i < n$ таких элементов b, что

$b^{n/p_i} = 1$.

Это значит, что в группе $GF^*(p^m)$ порядка $n = p^m - 1$ для каждого i существует элемент b_i с условием

$$b_i^{n/p_i} \neq 1. \tag{3.3.8}$$

Рассмотрим элемент $c_i = b_i^{n/p_i^{\alpha_i}}$ и покажем, что он имеет порядок $p_i^{\alpha_i}$.

Действительно,

$$(c_i)^{p_i^{\alpha_i}} = (b_i^{n/p_i^{\alpha_i}})^{p_i^{\alpha_i}} = b_i^n = 1. \tag{3.3.9}$$

С другой стороны,

$(c_i)^{p_i^{\alpha_i-1}} = (b_i^{n/p_i^{\alpha_i}})^{p_i^{\alpha_i-1}} = b_i^{n/p_i} \neq 1$ в силу (3.3.8).

Это означает, что порядок элемента c_i, будучи в силу (3.3.9) делителем числа $p_i^{\alpha_i}$, в то же время не равен ни одному из собственных его делителей (ибо все собственные делители числа $p_i^{\alpha_i}$ суть степени числа p_i).

Из этого следует, что порядок элемента c_i есть $p_i^{\alpha_i}$.

Это же справедливо для каждого $i = 1, 2, \ldots, k$. Отсюда следует, что произведение

$$c_1 c_2 \cdots c_k \tag{3.3.10}$$

имеет порядок, равный наименьшему общему кратному порядков сомножителей, т.е.

$p_1^{\alpha_1} p_2^{\alpha_2} ... p_k^{\alpha_k} = n$, так как они взаимно просты.

Таким образом, элемент (3.3.10) имеет порядок, равный порядку группы, а потому является образующим (первообразным) элементом, и группа $GF^*(p^m)$ циклическая. Этим завершается доказательство.

3.4. Задание поля посредством корня неприводимого многочлена

Дадим другой способ задания поля Галуа. Обозначим через α тот класс вычетов в кольце $F[x]_{/(p(x))}$, который содержит x. Тогда, так как $p(x)$ — это многочлен, по модулю которого построено поле $GF(p^m)$, то $p(\alpha) = 0$. Таким образом, α оказывается корнем уравнения $p(x) = 0$ в поле $GF(p^m)$. Благодаря этому $GF(p^m)$ состоит из многочленов от α степени, не превосходящей m над $GF(p)$. Говорится, что поле $GF(p^m)$ образовано из поля $GF(p)$ присоединением к нему корня α многочлена $p(x)$.

Читатель может вспомнить свои школьные годы и провести параллель между двумя событиями. Первое — это введение в обиход по необходимости "мнимой единицы" а второе – это и фактическое "назначение" элемента α корнем многочлена $p(x)$. В множестве действительных чисел не нашлось корня мнгочлена $x^2 + 1$. Тогда учитель почти волевым порядком

объявил решением уравнения $x^2 = -1$ число $i = \sqrt{-1}$. Так возникло поле комплексных чисел, как результат присоединения элемента i к полю действительных чисел. Поле комплексных чисел есть расширение поля действительных чисел. В нашем случае многочлен $p(x)$ неприводим, а значит, не имеет корней в поле $GF(p)$. "Назначив" его корнем элемент α, мы снова получили расширение исходного поля $GF(p)$.

П р и м е р 3. 4.

Пусть $p = 2$, $m = 4$. Построим $GF^*(2^4)$ по модулю многочлена $p(x) = x^4 + x + 1$, при условии $p(\alpha) = 0$, или, что то же $\alpha^4 = \alpha + 1$. Напомним, что в поле характеристики 2 выполняется равенство $-y = y$.

$$
\begin{array}{llllll}
\alpha^0 = & 1 & & & & = (1000) \\
\alpha^1 = & & \alpha & & & = (0100) \\
\alpha^2 = & & & \alpha^2 & & = (0010) \\
\alpha^3 = & & & & \alpha^3 & = (0001) \\
\alpha^4 = & 1 & +\alpha & & & = (1100) \\
\alpha^5 = & & \alpha & +\alpha^2 & & = (0110) \\
\alpha^6 = & & & \alpha^2 & +\alpha^3 & = (0011) \\
\alpha^7 = & 1 & +\alpha & & +\alpha^3 & = (1101) \\
\alpha^8 = & 1 & & \alpha^2 & & = (1010) \\
\alpha^9 = & & \alpha & & +\alpha^3 & = (0101) \\
\alpha^{10} = & 1 & +\alpha & +\alpha^2 & & = (1110) \\
\alpha^{11} = & & \alpha & +\alpha^2 & +\alpha^3 & = (0111) \\
\alpha^{12} = & 1 & +\alpha & +\alpha^2 & +\alpha^3 & = (1111) \\
\alpha^{13} = & 1 & & +\alpha^2 & +\alpha^3 & = (1011) \\
\alpha^{14} = & 1 & & & +\alpha^3 & = (1001) \\
\alpha^{15} = & 1 & & & & = (1000).
\end{array}
\tag{3.4.11}
$$

Если $p(x) = x^4 + x^3 + 1$ и $p(\beta) = 0$, а значит, $\beta^4 = \beta^3 + 1$, то

$GF^*(2^4)$ будет

$$\begin{array}{rccccl}
\beta^0 = & 1 & & & & = (1000)\\
\beta^1 = & & \beta & & & = (0100)\\
\beta^2 = & & & \beta^2 & & = (0010)\\
\beta^3 = & & & & \beta^3 & = (0001)\\
\beta^4 = & 1 & & & +\beta^3 & = (1001)\\
\beta^5 = & 1 & +\beta & & +\beta^3 & = (1101)\\
\beta^6 = & 1 & +\beta & +\beta^2 & +\beta^3 & = (1111)\\
\beta^7 = & 1 & \beta & +\beta^2 & & = (1110)\\
\beta^8 = & & \beta & +\beta^2 & +\beta^3 & = (0111)\\
\beta^9 = & 1 & & +\beta^2 & & = (1010)\\
\beta^{10} = & & \beta & & +\beta^3 & = (0101)\\
\beta^{11} = & 1 & & \beta^2 & +\beta^3 & = (1011)\\
\beta^{12} = & 1 & +\beta & & & = (1100)\\
\beta^{13} = & & \beta & +\beta^2 & & = (0110)\\
\beta^{14} = & & & \beta^2 & +\beta^3 & = (0011)\\
\beta^{15} = & 1 & & & & = (1000).
\end{array} \tag{3.4.12}$$

(Замена α на β непринципиальна и предпринята, чтобы избежать путаницы.)

В правой части каждой строки двоичный вектор длины $m = 4$ изображает набор коэффициентов при степенях α, соответственно, β. Бросается в глаза, что способ изображения элементов поля в (3.4.11) и (3.4.12) отличается от способа изображения в (3.3.6) и (3.3.7) только тем, что буква x заменена буквой α. Такая замена при обращении с полями Галуа позволит нам оперировать в терминах корней многочленов, что оказывается намного удобней. Разумеется, кроме таблиц (3.4.11) и (3.4.12), оба поля содержат ещё и нуленой элемент.Пример 3.4. демонстрирует связь между аддитивной и мультипликативной группами поля.

В самом деле,

П р и м е р 3. 5.

Пусть, имея дело с полем (3.4.11), надо перемножить два элемента поля, заданных векторами $(1110), (0011)$, и получить результат также в векторной форме. Сначала устанавливается, что в мультипликативной группе два этих вектора представляют собой элементы соответственно α^{10}, α^6.

Отсюда $\alpha^{10} \cdot \alpha^6 = \alpha^{16} = \alpha$. В векторной форме $\alpha = (0100)$; таким образом, $(1110)(0011) = (0100)$.

Обратно,

П р и м е р 3. 6.

Пусть, имея дело с полем (3.4.12), надо сложить два элемента поля β^3 и β^8 и получить результат в терминах степеней первообразного элемента поля. Сначала устанавливается, что в аддитивной группе поля два этих элемента представляют собой соответственно $(0001), (0111)$. Отсюда $(0001) + (0111) = (0110)$. В виде степени первообразного элемента $(0110) = \beta^{13}$. Таким образом, $\beta^3 + \beta^8 = \beta^{13}$.

Примеры 3.5 и 3.6 демонстрируют связь между таблицами сложения и умножения. Однако в поле возможно еще и деление, т.е. нахождение обратного элемента.

П р и м е р 3. 7.

Пусть, имея дело с полем (3.4.11), надо найти элемент, обратный к (0011), т.е. надо найти $(0011)^{-1}$. Последовательно имеем: $(0011) = \alpha^6$, $(\alpha^6)^{-1} = \alpha^9 = (0101)$. Иначе говоря, $(0011)^{-1} = (0101)$.

Для полноты приведем поле $GF(2^5)$, построенное по модулю многочлена $x^5 + x^2 + 1$

$$
\begin{array}{lll}
0 = (00000) & \alpha^{10} = (10001) & \alpha^{21} = (00011) \\
\alpha^0 = (10000) & \alpha^{11} = (11100) & \alpha^{22} = (10101) \\
\alpha^1 = (01000) & \alpha^{12} = (01110) & \alpha^{23} = (11110) \\
\alpha^2 = (00100) & \alpha^{13} = (00111) & \alpha^{24} = (01111) \\
\alpha^3 = (00010) & \alpha^{14} = (10111) & \alpha^{25} = (10011) \\
\alpha^4 = (00001) & \alpha^{15} = (11111) & \alpha^{26} = (11101) \\
\alpha^5 = (10100) & \alpha^{16} = (11011) & \alpha^{27} = (11010) \\
\alpha^6 = (01010) & \alpha^{17} = (11001) & \alpha^{28} = (01101) \\
\alpha^7 = (00101) & \alpha^{18} = (11000) & \alpha^{29} = (10010) \\
\alpha^8 = (10110) & \alpha^{19} = (01100) & \alpha^{30} = (01001) \\
\alpha^9 = (01011) & \alpha^{20} = (00110) & \alpha^{31} = (10000).
\end{array}
\tag{3.4.13}
$$

Заметим, что в некоторых руководствах принят обратный порядок следования компонент векторов, который, таким образом, является делом соглашения.

Специально обратим внимание на то, что, кроме двух рассмотренных выше многочленов 4-й степени, имеется, как будет показано в следующем разделе, еще один многочлен 4-й степени: $x^4 + x^3 + x^2 + x + 1$. Положив ξ корнем этого многочлена, получим, как и в случае двух других многочленов 4-й степени,

$$\xi^4 + \xi^3 + \xi^2 + \xi + 1 = 0,$$

а значит, и

$$\xi^4 = \xi^3 + \xi^2 + \xi + 1.$$

Станем, как и выше, возводить ξ в последовательные степени:

$$\begin{array}{lccccl}
\xi^0 = & 1 & & & & = (1000) \\
\xi = & & \xi & & & = (0100) \\
\xi^2 = & & & \xi^2 & & = (0010) \\
\xi^3 = & & & & \xi^3 & = (0001) \\
\xi^4 = & 1 & +\xi & +\xi^2 & +\xi^3 & = (1111) \\
\xi^5 = & 1 & & & & = (1000).
\end{array} \tag{3.4.14}$$

Таким образом, корень ξ порождает не всю мультипликативную группу поля $GF(2^4)$, а только ее подгруппу 5-го порядка. Таким же свойством, как читатель увидит в этой главе, обладают и все остальные корни многочлена $x^4 + x^3 + x^2 + x + 1$. Ниже порождающие свойства корней неприводимых многочленов обсуждаются в общем виде.

Покажем, что элемент $\xi + 1$ порождает все поле $GF(2^4)$. Действительно, выполняя все вычисления по правилу $\xi^4 = \xi^3 + \xi^2 + \xi + 1$, получим:

$$\begin{array}{lccccl}
(\xi+1)^0 = & 1 & & & & = (1000) \\
(\xi+1)^1 = & 1 & +\xi & & & = (1100) \\
(\xi+1)^2 = & 1 & & +\xi^2 & & = (1010) \\
(\xi+1)^3 = & 1 & +\xi & +\xi^2 & +\xi^3 & = (1111) \\
(\xi+1)^4 = & & \xi & +\xi^2 & +\xi^3 & = (0111) \\
((\xi+1)^5 = & 1 & & +\xi^2 & +\xi^3 & = (1011) \\
(\xi+1)^6 = & & & & \xi^3 & = (0001) \\
(\xi+1)^7 = & 1 & +\xi & +\xi^2 & & = (1110) \\
(\xi+1)^8 = & 1 & & & +\xi^3 & = (1001) \\
(\xi+1)^9 = & & & \xi^2 & & = (0010) \\
(\xi+1)^{10} = & & & \xi^2 & +\xi^3 & = (0011) \\
(\xi+1)^{11} = & 1 & +\xi & & +\xi^3 & = (1101) \\
(\xi+1)^{12} = & & \xi & & & = (0100) \\
(\xi+1)^{13} = & & \xi & +\xi^2 & & = (0110) \\
(\xi+1)^{14} = & & \xi & & +\xi^3 & = (0101) \\
(\xi+1)^{15} = & 1 & & & & = (1000).
\end{array} \tag{3.4.15}$$

Поле (3.4.15) построено по модулю неприводимого многочлена $x^4 + x^3 + x^2 + x + 1$, но в последовательные степени возводился не его корень ξ, а $\xi + 1$.

Добавление единицы к ξ ничем не мотивировано и на первый взгляд кажется случайным. Рассмотрим, однако, получившийся результат внимательней. Выясним, что собой представляет элемент $\xi + 1$, и можно ли вместо него воспользоваться другими элементами поля $GF(2^4)$? Обратившись к таблице (3.4.11), получим, что корнем многочлена $x^4 + x^3 + x^2 + x + 1$ является элемент α^3. Действительно, $\alpha^{12} + \alpha^9 + \alpha^6 + \alpha^3 + 1 = 0$. Значит, $\xi = \alpha^3$. Становится ясным, почему в (3.4.14) всего пять строчек: элемент α^3 в мультипликативной группе $GF^*(2^4)$ имеет порядок пять. В то же время элемент $\xi + 1 = \alpha^3 + 1 = \alpha^{14}$ в этой группе является элементом порядка 15, т.е. порождающим. Естественно поэтому искать теперь те элементы $\xi + x$, которые также являются порождающими. Для этого следует построить все 16 сумм $x = a_0 + a_1\xi + a_2\xi^2 + a_3\xi^3, a_i \in GF(2)$, и проверить, какие из них обращают сумму $\xi + x$ в порождающий элемент группы $GF^*(2^4)$. Проверим сначала простейшее значение $x = \xi^2$. Имеем $\xi + \xi^2 = \alpha^3 + \alpha^6 = \alpha^2$. Это порождающий элемент. Таким же образом $\xi + \xi^3 = \alpha^3 + \alpha^9 = \alpha$, и это снова порождающий элемент. А вот $\xi + \xi^4 = \alpha^3 + \alpha^{12} = \alpha^{10}$, и это элемент третьего порядка: $\alpha^{30} = 1$. Читатель без труда проверит, что вместе с $x = 1, \xi^2, \xi^3$, которые уже дали порождающие элементы $\alpha^{14}, \alpha^2, \alpha$, значения $x = \xi^2 + 1,\ \xi^2 + \xi + 1,\ \xi^3 + 1,\ \xi^3 + \xi + 1,\ \xi^3 + \xi^2$ доставляют соответственно следующие порождающие элементы: $\alpha^8, \alpha^{13}, \alpha^4, \alpha^7, \alpha^{11}$. Отметим, что получены все $\varphi(15) = 8$ порождающих элементов. Потому первоначально и взят порождающий элемент $\xi + 1 = \alpha^{14}$, что в (3.4.15) в последовательные степени достаточно возводить выражения, содержащие только первую степень ξ. Во всех остальных случаях при построении поля приходится иметь дело с более высокими степенями.

В самом общем случае, когда корень ξ некоторого многочлена, по модулю которого строится поле $GF(p^m)$, не является порождающим элементом мультипликативной группы поля, процесс вычислений совершенно аналогичен только-что рассмотренному. Действительно, рассмотрим уравнение $\xi + x = \alpha^i$, где α^i вместе с α есть порождающий элемент группы $GF^*(p^m)$. (Вспомним, что $(i, p^m) = 1$ и таких i в точности $\varphi(p^m - 1)$). Оно всегда имеет решение и при том единственное, так как поле всегда аддитивная группа. Последнее утверждение справедливо вообще для любого элемента α^i поля, а не только для порождающего. Но нас в данном случае интересуют только порожда-

ющие элементы. Как и в рассмотренном частном случае, испытываются все p^m значений $x = a_0 + a_1\xi + a_2\xi^2 + a_{m-1}\xi^{m-1}, a_i \in GF(p)$,

В качестве примера построим мультипликативную группу $GF^*(2^4)$, возводя в последовательные степени сумму $\xi^2+\xi$. Забегая вперёд предупредим читателя, что максимальная степень элемента ξ, которая встретится при вычислениях, будет ξ^5, но она, как известно равна единице. Все остальные вычисления выполняются по модулю многочлена $x^4+x^3+x^2+x+1$, т.е. при условии $\xi^4 = \xi^3 + \xi^2 + \xi + 1$. Например, в процессе построения мультипликативной группы, как увидит читатель, если он захочет вникнуть в подробности, получится $(\xi+\xi^2)^5 = 1+\xi^2+\xi^3$. Тогда $(\xi + \xi^2)^6 = (1 + \xi^2 + \xi^3)(\xi + \xi^2) = \xi + \xi^2 + \xi^3 + \xi^5 = 1 + \xi + \xi^2 + \xi^3$.

Итак, $GF^*(2^4)$ с порождающим элементом $(\xi + \xi^2)$:

$$
\begin{array}{lcccccl}
(\xi+\xi^2)^0 = & 1 & & & & & = (1000)\\
(\xi+\xi^2)^1 = & & \xi & +\xi^2 & & & = (0110)\\
(\xi+\xi^2)^2 = & 1 & +\xi & & +\xi^3 & & = (1101)\\
(\xi+\xi^2)^3 = & & & +\xi^2 & & & = (0010)\\
(\xi+\xi^2)^4 = & 1 & +\xi & +\xi^2 & & & = (1110)\\
(\xi+\xi^2)^5 = & 1 & & +\xi^2 & +\xi^3 & & = (1011)\\
(\xi+\xi^2)^6 = & 1 & +\xi & +\xi^2 & +\xi^3 & & = (1111)\\
(\xi+\xi^2)^7 = & 1 & +\xi & & & & = (1100)\\
(\xi+\xi^2)^8 = & & \xi & & +\xi^3 & & = (0101)\\
(\xi+\xi^2)^9 = & & \xi & & & & = (0100)\\
(\xi+\xi^2)^{10} = & & & \xi^2 & +\xi^3 & & = (0011)\\
(\xi+\xi^2)^{11} = & 1 & & & +\xi^3 & & = (1001)\\
(\xi+\xi^2)^{12} = & & & & \xi^3 & & = (0001)\\
(\xi+\xi^2)^{13} = & & \xi & +\xi^2 & +\xi^3 & & = (0111)\\
(\xi+\xi^2)^{14} = & 1 & & +\xi^2 & & & = (1010)\\
(\xi+\xi^2)^{15} = & 1 & & & & & = (1000).
\end{array}
$$

Легко сообразить, что во всех остальных случаях порождающих элементов группы $GF^*(2^4)$ максимальная степень элемента ξ, которая встретится при последовательном вычислении степеней, будет 6. Но $\xi^6 = \xi$.

В заключение этого раздела вернёмся к построению поля (3.4.15). Основной заботой там было построение мультипликативной группы поля в виде расположения её элементов по воз-

растающим степеням порождающего элемента. Однако, естественной последовательностью рассуждений, на наш взгляд, является следующее. Непреложный факт состоит в том, что по модулю любого неприводимого многочлена $p(x)$ степени m над полем $GF(p)$ всегда можно построить поле $GF(p^m)$. А это означает, что в первую очередь следует определить для каждого остатка $r(x)$, получающегося при делении на $p(x)$, тот остаток $r(x^{'})$, который удовлетворяет условию $r(x)r(x^{'}) \equiv 1(\mathrm{mod}p(x))$, ибо это и есть главный признак того, что множество всех отличных от нуля остатков есть мультипликативная группа. Выбор из этих остатков порождающих элементов группы, которая, как доказано выше, всегда циклическая, есть уже следующий шаг. Он и сделан при построении поля (3.4.15).

Легко проверить, что в (3.4.15) для всякого остатка под именем $(\xi+1)^i$ остаток под именем $(\xi+1)^{15-i}$ будет обратным по модулю многочлена $x^4+x^3+x^2+x+1$, что, конечно, усматривается непосредственно.

3.5. Строение конечных полей.

Теорема 3.5.1. *В поле характеристики p имеет место равенство $(a+b)^p = a^p + b^p$.*

Доказательство. После сокращения дробей в правой части разложения

$$(a+b)^p = \sum_{i=0}^{p} C_p^i a^i b^{p-i} = \sum_{i=0}^{p} \frac{p!}{i!(p-i)!} a^i b^{p-i}$$

множитель p сохранится в числителях коэффициентов всех слагаемых, где $i \neq 0$ и p, так как в знаменателях этих коэффициентов не содержится p, и p — простое. В силу определения характеристики поля все такие коэффициенты обращаются в нуль. При $i = 0$ и $i = p$ оба коэффициента равны 1.

После построения поля $GF(p^m)$ посредством расширения поля $GF(p)$, которое есть не что иное, как полная система вычетов по простому модулю p, не составит труда сделать следующий шаг, отнюдь не переходя на новый уровень абстракции: точно таким же способом строить поле $GF(q^m)$, как расширение поля $GF(q)$, $q = p^s$.

П р и м е р 3.8. Единственный неприводимый многочлен второй степени над $GF(2)$ есть $1 + x + x^2$. Поле $GF(2^2)$, построенное по модулю этого многочлена есть $\{0 = (00), 1 = (10), \alpha = (01), \alpha^2 = 1 + \alpha = (11)\}$. Расширение второй степени теперь уже этого поля будет $GF((2^2)^2)$. Оно строится по модулю неприводимого над $GF(2^2)$ многочлена $x^2 + x + \alpha$. Легко проверить, что ни один из элементов поля $GF(2^2)$ не является корнем этого многочлена. Положив его корнем элемент γ, получим $\gamma^2 = \gamma + \alpha$. Отсюда

$$
\begin{array}{llll}
0 = & = (0,0) & \gamma^7 = \alpha^2 + \alpha\gamma & = (\alpha^2, \alpha) \\
\gamma^0 = 1 & = (1,0) & \gamma^8 = \alpha^2 + \gamma & = (\alpha^2, 1) \\
\gamma^1 = \gamma & = (0,1) & \gamma^9 = \alpha + \alpha\gamma & = (\alpha, \alpha) \\
\gamma^2 = \alpha + \gamma & = (\alpha, 1) & \gamma^{10} = \alpha^2 & = (\alpha^2, 0) \\
\gamma^3 = \alpha + \alpha^2\gamma & = (\alpha, \alpha^2) & \gamma^{11} = \alpha^2\gamma & = (0, \alpha^2) \\
\gamma^4 = 1 + \gamma & = (1,1) & \gamma^{12} = 1 + \alpha^2\gamma & = (1, \alpha^2) \\
\gamma^5 = \alpha & = (\alpha, 0) & \gamma^{13} = 1 + \alpha\gamma & = (1, \alpha) \\
\gamma^6 = \alpha\gamma & = (0, \alpha) & \gamma^{14} = \alpha^2 + \alpha^2\gamma & = (\alpha^2, \alpha^2)
\end{array}
$$

Теперь должно и можно доказать

Следствие 3.5.2. *В поле характеристики p*

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

Действительно, если

$$(a + b)^{p^{s-1}} = a^{p^{s-1}} + b^{p^{s-1}},$$

то согласно теореме 3.5.1

$$(a + b)^{p^s} = ((a + b)^{p^{s-1}})^p = (a^{p^{s-1}})^p + (b^{p^{s-1}})^p = a^{p^s} + b^{p^s}.$$

Этим выполнен индукционный шаг, и та же самая теорема доставляет основание индукции.

Смотря по случаю, далее будем полагать $q = p$ или $q = p^s$.

Определение 3.5.3. *Минимальной функцией (минимальным многочленом) для элемента $\beta \in GF(q^m)$ называется такой нормированный многочлен $m(x)$ над $GF(q)$ минимальной степени, что $m(\beta) = 0$.*

Теорема 3.5.4. *Минимальная функция для β — неприводимый многочлен над $GF(q)$.*

Д о к а з а т е л ь с т в о. Предположим противное, и пусть

$$m(x) = m_1(x)m_2(x).$$

Тогда $= m_1(\beta)m_2(\beta) = 0$, и по крайней мере один из многочленов степени, меньшей чем степень многочлена $m(x)$, имеет своим корнем β, что противоречит тому, что по условию $m(x)$ есть минимальный многочлен для β.

Теорема 3.5.5. *Если многочлен $f(x)$ таков, что $f(\beta) = 0$, то $m(x)|f(x)$, где $m(x)$ — минимальная функция для β.*

Действительно, если

$$f(x) = m(x)q(x) + r(x),$$

то

$$f(\beta) = m(\beta)q(\beta) + r(\beta) = 0.$$

Отсюда $r(\beta) = 0$, и $r(x) = 0$ тождественно по x, так как в противном случае, имея степень меньшую, чем степень $m(x)$, многочлен $r(x)$ был бы минимальной функцией для β, вопреки условию.

Следствие 3.5.6. *Минимальная функция для β единственна.*

В самом деле, пусть $m_1(x)$ и $m_2(x)$ минимальные функции для β. Их степени совпадают, и на основании теоремы 3.5.5 они делят друг друга, а потому могут отличаться только на постоянный множитель. Но так как они нормированы, то совпадают.

Следствие 3.5.7. *Всякий нормированный неприводимый многочлен $p(x)$, такой, что $p(\beta) = 0$, является минимальной функцией для β.*

Действительно, на основании теоремы 3.5.5 многочлен $p(x)$ делится на $m(x)$. Но он неприводим и потому совпадает с $m(x)$.

Следствие 3.5.8. *Каждый неприводимый над $GF(q)$ многочлен $p(x)$ степени m является делителем двучлена $x^{q^m} - x$.*

Д о к а з а т е л ь с т в о. Если $p(x) = x$, то утверждение тривиально. Пусть $p(x) \neq x$. Построим поле $GF(q^m)$ по модулю многочлена $p(x)$, и пусть элемент $\beta \in GF(q^m)$ таков, что $p(\beta) = 0$. Его порядок делит порядок $q^m - 1$ мультипликативной группы поля $GF(q^m)$. Поэтому β является корнем двучлена $x^{q^m-1} - 1$. Остается применить теорему 3.5.5.

Теорема 3.5.9. *Для каждого элемента $\beta \in GF(q^m)$ существует минимальная функция.*

Д о к а з а т е л ь с т в о. Поле $GF(q^m)$ есть линейное векторное пространство, так как оно является аддитивной абелевой группой, и определено умножение вектора на скаляры. (Каковыми являются элементы поля $GF(q)$). Размерность этого пространства равна m. Поэтому $m+1$ векторов $1, \beta, \beta^2, \ldots, \beta^m$ заведомо линейно зависимы над $GF(q)$. Это значит, что найдутся такие не все равные нулю элементы $b_i \in GF(q)$, $i = 0, 1, \ldots, m$, что

$$b_0 + b_1\beta + b_2\beta^2 + \ldots + b_m\beta^m = 0.$$

Это значит, что существует многочлен степени не выше, чем m, для которого элемент β является корнем. Остается поделить этот многочлен на его старший коэффициент, что возможно, так как $GF(q)$ — поле. Может оказаться, что сам многочлен с коэффициентами b_i, $i = 0, 1, \ldots, m$, приводим. Тогда минимальной функцией будет некоторый его делитель.

Теорема 3.5.10. *Многочлен $x^n - 1$ делит многочлен $x^m - 1$ тогда и только тогда, когда m кратно n.*

Д о к а з а т е л ь с т в о. *Достаточность.* Если $m = nd$, то

$$x^m - 1 = x^{nd} - 1 = (x^n - 1)(x^{n(d-1)} + x^{n(d-2)} + \ldots + 1),$$

как сумма членов геометрической прогрессии со знаменателем x^n.

Необходимость. Пусть $m = nd + r$, $r < n$. Тогда

$$x^m - 1 = x^{nd+r} - 1 = x^r(x^{nd} - 1) + x^r - 1.$$

Если $x^m - 1 \equiv 0(\mathrm{mod}(x^n - 1))$, то

$$x^m - 1 = x^r(x^{nd} - 1) + x^r - 1 \equiv 0(\mathrm{mod}(x^n - 1)),$$

где $r < n$. Но по доказанному выше $x^{nd} - 1 \equiv 0(\mathrm{mod}(x^n - 1)$. Поэтому $x^r - 1 \equiv 0(\mathrm{mod}(x^n - 1)$, что возможно только при $r = 0$.

Следствие 3.5.11. *$x^{q^m-1} - 1$ делится на $x^{q^n-1} - 1$ тогда и только тогда, когда m делится на n.*

Для доказательства достаточно заменить $q^m - 1$ на M, и $q^n - 1$ на N и применить теорему сначала к $x^M - 1$ и $x^N - 1$, а затем к M и N.

Теорема 3.5.12. *Если $f(x)$ многочлен над $GF(q)$, $\beta \in GF(q^m)$ и $f(\beta) = 0$, то $f(\beta^q) = 0$.*

Иначе говоря, если β корень многочлена, то и β^q — также его корень.

Д о к а з а т е л ь с т в о. Пусть $f(x) = a_0 + a_1x + \ldots + a_nx^n$. Согласно теореме 3.5.1

$$(f(x))^q = (a_0)^q + (a_1)^q x^q + \ldots + (a_n)^q (x^n)^q =$$

$$= a_0 + a_1x^q + \ldots + a_nx^{nq} = f(x^q),$$

так как $a \in GF(q)$, а потому $a^{q-1} = 1$, и $a^q = a$.

Определение 3.5.13. *Если подмножество K элементов поля F само есть поле относительно операций в поле F, то оно составляет подполе поля F : $K \subset F$.*

Это означает, что аддитивная и мультипликативная группы поля K являются подгруппами соответственно аддитивной и мультипликативной группы поля F. Поле, не имеющее подполей, называется *простым* полем.

Теорема 3.5.14. *Если все корни многочлена $x^{q^n} - x$ содержатся в поле $GF(q^m)$, то они составляют его подполе $GF(q^n) \subset$*

$\subset GF(q^m)$.

Д о к а з а т е л ь с т в о. Покажем, что все корни многочлена $x^{q^n} - x$ образуют аддитивную группу, т.е подгруппу аддитивной группы поля $GF(q^m)$. Покажем также, что все ненулевые корни этого многочлена образуют мультипликативную группу, т.е. подгруппу мультипликативной группы поля $GF(q^m)$.

Пусть $a^{q^n} - a = 0, \quad b^{q^n} - b = 0$. Тогда

1. Согласно теореме 3.5.1, $(a+b)^{q^n} = a^{q^n} + b^{q^n}$.

Но $a^{q^n} = a, \quad b^{q^n} = b$, откуда $(a+b)^{q^n} = a+b$. Таким образом, множество корней многочлена $x^{q^n} - x$ замкнуто относительно сложения. Кроме того $(-a)^{q^n} = (-1)^{q^n} a^{q^n} = -a$. Последнее имеет место потому, что при q нечётном $(-1)^{q^n=-1}$, а при q чётном $-a = a$.

2. $(ab)^{q^n} = a^{q^n} b^{q^n} = ab$, т.е. произведение корней снова корень. Таким образом, множество ненулевых корней многочлена $x^{q^n} - x$ замкнуто относительно умножения.

Как известно (см. раздел 2.6 "Подгруппы"), наличие противоположного элемента в конечной аддитивной группе и обратного — в конечной мультипликативной группе следует из замкнутости. Этим завершается доказательство.

Из следствия 3.5.11 и теоремы 3.5.14 получается

Теорема 3.5.15. *Для каждого делителя n числа m существует одно и только одно подполе $GF(q^n)$ поля $GF(q^m)$.*

Действительно, пусть поле $GF(q^m)$ есть расширение степени m поля $GF(q)$. Порядок подгруппы аддитивной группы поля всегда есть степень q^n, а порядок $q^n - 1$ подгруппы $GF^*(q^n)$ мультипликативной группы $GF^*(q^m)$ поля $GF(q^m)$ есть делитель порядка $q^m - 1$ этой группы. Но $q^n - 1$ есть делитель числа $q^m - 1$ тогда и только тогда, когда m делится на n, и для каждого делителя l порядка циклической группы существует циклическая же подгруппа порядка l, и она единственна.

Следует помнить, что у мультипликативной группы поля могут быть и другие подгруппы, так как мультипликативная группа поля циклична, и для каждого делителя ее порядка, как отмечено только-что, есть соответствующая ему циклическая подгруппа. Но мультипликативными группами *подполей* будут только группы порядка вида $q^n - 1$, где n делит m.

Теорема 3.5.16. *Неприводимые над $GF(q)$ многочлены $p(x)$, степени n которых делят m, и только они, являются делителями многочлена $x^{q^m} - x$.*

Д о к а з а т е л ь с т в о. Построим поле $GF(q^n)$, по модулю многочлена $p(x)$ степени n, которая делит m. Согласно теореме 3.5.15 оно является подполем поля $GF(q^m)$. Если α корень многочлена $p(x)$, то, будучи элементом поля $GF(q^n)$, он

принадлежит также полю $GF(q^m)$. А потому согласно теореме 3.5.5 $p(x)$ делит $x^{q^m} - x$.

Если же n, не делит m, то поле $GF(q^n)$ по той же теореме не является подполем поля $GF(q^m)$, а потому $p(x)$ не делит $x^{q^m} - x$, так как все делители многочлена $x^{q^m} - x$ исчерпываются минимальными функциями элементов поля $GF(q^m)$, в том числе и элементов его подполей.

П р и м е р 3. 9.

Пусть $p = 2$. При $m = 1$ имеем $GF(2) = \{0, 1\}$.

При $m = 2$ (это значение m рассматривается для полноты изложения, хотя в примере 3.8 ему уже отводилось место) имеем $GF(2^2) = \{0 = (00), 1 = (10), \alpha = (01), \alpha^2 = (11)\}$.

$$x^{2^2} - x = x(x-1)(x^2 + x + 1).\ GF^*(2^2) = \{1, \alpha, \alpha^2\}.$$

Корень одночлена x есть 0, корень двучлена $(x - 1)$ есть 1, корнями многочлена $(x^2 + x + 1)$ являются α, α^2. Элементы 0 и 1 — это корни многочлена $x^{q^m} - x$ при любом m.

При $m = 3$ $\quad x^{2^3} - x = x(x-1)(x^3 + x + 1)(x^3 + x^2 + 1)$. Заменив (по аналогии с (3.4.11) и (3.4.12)) в (3.3.6) и (3.3.7) x соответственно на α и β, получим, что в поле (3.3.6) корнями многочлена $m_1(x) = (x^3 + x + 1)$ будут α, α^2, α^4, а корнями многочлена $m_3(x) = (x^3 + x^2 + 1)$ будут α^3, α^6, α^5. Многочлены $m_1(x)$ и $m_3(x)$ есть минимальные функции своих корней.

Здесь и далее номер минимальной функции выбран равным минимальному показателю степени ее корней. Ясно, что эта нумерация может меняться в зависимости от неприводимого многочлена, по модулю которого строится поле. Легко видеть, что первый номер присваивается той функции, именно по модулю которой построено поле. На самом деле нумерация принципиального значения не имеет и вводится по соглашению, только из соображений наглядности и удобства.

В поле (3.3.7) корнями этих мнгочленов будут соответственно β^3, β^6, β^5 и β, β^2, β^4.

Так как степень $m = 3$ расширения есть простое число, то нетривиальных (кроме $GF(2)$) подполей нет. А так как порядок 7 группы $GF^*(2^3)$ также простое число, то нет и мультипликативных подгрупп.

При $m = 4$ $\quad x^{2^4} - x = x(x-1)m_5(x)m_1(x)m_7(x)m_3(x)$, где

$$m_5(x) = x^2 + x + 1,\ m_1(x) = x^4 + x + 1,$$

$$m_7(x) = x^4 + x^3 + 1,\ m_3(x) = x^4 + x^3 + x^2 + x + 1.$$

Корнями минимальных функций $m_5(x)$, $m_1(x)$, $m_7(x)$, $m_3(x)$ будут соответственно

в поле (3.4.11): (α^5, α^{10}); (α, α^2, α^4, α^8); (α^7, α^{14}, α^{13}, α^{11}); (α^3, α^6, α^{12}, α^9);

в поле (3.4.12): (β^5, β^{10}); (β^7, β^{14}, β^{13}, β^{11}); (β, β^2, β^4, β^8); (β^3, β^6, β^{12}, β^9);

$GF(2^4)$ содержит одно нетривиальное подполе $GF(2^2)$: в поле (3.4.11) это 0, 1, α^5, α^{10}, в поле (3.4.12) это 0, 1, β^5, β^{10}.

Мультипликативная группа $GF^*(2^4)$ содержит две собственные подгруппы.

Одна — это мультипликативная группа подполя $GF(2^2)$, другая — подгруппа пятого порядка 1, α^3, α^6, α^9, α^{12} — в поле (3.4.11), и 1, β^3, β^6, β^9, β^{12} — в поле (3.4.12).

При $m = 5$

$$x^{2^5} - x = x(x-1)m_1(x)m_3(x)m_5(x)m_{15}(x)m_7(x)m_{11}(x),$$

где

$$m_1(x) = x^5 + x^2 + 1,\ m_3(x) = x^5 + x^4 + x^3 + x^2 + 1,$$
$$m_5(x) = x^5 + x^4 + x^2 + x + 1,\ m_{15}(x) = x^5 + x^3 + 1,$$
$$m_7(x) = x^5 + x^3 + x^2 + x + 1,\ m_{11}(x) = x^5 + x^4 + x^3 + x + 1.$$

Корнями минимальных функций

$$m_1(x), \quad m_2(x), \quad m_3(x), \quad m_4(x), \quad m_5(x), \quad m_6(x)$$

в поле, построенном по модулю многочлена $(x^5 + x^2 + 1)$, будут соответственно

$$(\alpha, \alpha^2, \alpha^4, \alpha^8, \alpha^{16}), (\alpha^3, \alpha^6, \alpha^{12}, \alpha^{24}, \alpha^{17}), (\alpha^5, \alpha^{10}, \alpha^{20}, \alpha^9, \alpha^{18}),$$

$$(\alpha^{15}, \alpha^{23}, \alpha^{27}, \alpha^{29}, \alpha^{30}), (\alpha^7, \alpha^{14}, \alpha^{28}, \alpha^{25}, \alpha^{19}),$$
$$(\alpha^{11}, \alpha^{22}, \alpha^{13}, \alpha^{26}, \alpha^{21}).$$

Как и в случае $m = 3$, здесь и степень $m = 5$ расширения поля, и порядок 31 мультипликативной группы поля — суть

простые числа, а потому ни подполей (кроме простого подполя) ни мультипликативных подгрупп нет. К этому примеру еще будет повод вернуться.

Кстати, принадлежность корней к их минимальным функциям проверяется непосредственно. Например, легко проверить, что элемент β^5 обращает в нуль многочлен $m_5(x) = x^5 + x^4 + x^2 + x + 1$. Действительно, $m_5(\alpha^5) = \alpha^{25} + \alpha^{20} + \alpha^{10} + \alpha + 1 = (10011) + (00110) + (10001) + (10100) + (10000) = (00000)$ согласно правилам поразрядного сложения векторов по модулю 2. Векторные представления степеней α взяты из таблицы (3.4.13) поля $GF(2^5)$.

Несколько более трудоемкой выглядит задача построения многочлена по его корням.

Положим, что известны только корни α^5, α^{10}, α^{20}, α^9, α^{18} многочлена $m_5(x)$, а сам многочлен неизвестен. Очевидно,

$$m_5(x) = (x - \alpha^5)(x - \alpha^{10})(x - \alpha^{20})(x - \alpha^9)(x - \alpha^{18}).$$

Раскрывая скобки и приводя подобные члены посредством таблицы (3.4.13) поля $GF(2^5)$, получим последовательно: $m_5(x) = (x^2 + \alpha^7 x + \alpha^{15})(x^2 + \alpha^{28} x + \alpha^{29})(x + \alpha^{18}) = (x^4 + (\alpha^{28} + \alpha^7)x^3 + (\alpha^{15} + \alpha^{29} + \alpha^4)x^2 + (\alpha^5 + \alpha^{12})x + \alpha^{13})(x + \alpha^{18}) = x^5 + (\alpha + \alpha^{18})x^4 + (\alpha^{19} + \alpha^{19})x^3 + (\alpha^{27} + \alpha^6)x^2 + (\alpha^{13} + \alpha^{14})x + 1 = x^5 + x^4 + x^2 + x + 1$.

Теорема 3.5.17. *Если минимальная функция $m(x)$ элемента $\beta \in GF(q^m)$ имеет степень m, то минимальное число вида $q^k - 1$, которое делится на порядок e элемента β, равно $q^m - 1$.*

Доказательство. Так как $\beta \in GF(q^m)$, то $\beta^{q^m - 1} - 1 = 0$, а потому $m(x)$ делит $x^{q^m - 1} - 1$. Если e делит $q^k - 1$ при $k < m$, то из $\beta^e = 1$ следует, что подавно $\beta^{q^k - 1} = 1$, т.е. $\beta^{q^k - 1} - 1 = 0$, и $m(x)$ делит $x^{q^k - 1} - 1$, чего быть не может по теореме 3.5.16

Теорема 3.5.18. *Все корни $\beta, \beta^q, \ldots, \beta^{q^{m-1}} \in GF(q^m)$ неприводимого над $GF(q)$ многочлена $p(x)$ степени m различны.*

Доказательство. Все элементы $\beta, \beta^q, \ldots, \beta^{q^{m-1}} \in GF(q^m)$ являются корнями многочлена $p(x)$ по теореме 3.5.12. Предположим, что для двух каких-нибудь корней выполняется

равенство $\beta^{q^i} = \beta^{q^j}$, $i > j$. Тогда $\beta^{q^j(q^{i-j}-1)} = 1$, и порядок e элемента β делит число $q^j(q^{i-j}-1)$. Но порядок e, будучи делителем порядка мультипликативной группы поля Галуа, т.е. чисел вида $q^k - 1$, взаимно прост с числами вида q^k, а потому делит число $q^{i-j} - 1$. Однако по теореме 3.5.17 и это невозможно, так как $i - j < m$. Противоречие доказывает теорему.

Впрочем, можно рассуждать иначе: $x^{q^m} - x$ не имеет кратных корней, и многочлен степени m имеет m корней.

Определение 3.5.19. *Элементы поля, являющиеся корнями одного и того же неприводимого многочлена, называются сопряженными элементами поля.*

Теорема 3.5.20. *Все корни одного и того же неприводимого многочлена имеют одинаковый порядок.*

Другими словами, сопряженные элементы поля Галуа имеют одинаковый порядок.

Доказательство. Пусть элемент β порождает циклическую подгруппу порядка n. Как утверждается в разделе 2.8, вместе с элементом β порождающими элементами подгруппы будут все такие степени β^m, что $(m, n) = 1$. Но показатели q^i все взаимно просты со всеми делителями чисел вида $q^j - 1$. Эти делители и только они являются порядками подгрупп мультипликативной группы поля Галуа.

Определение 3.5.21. *Порядок корней неприводимого многочлена называется показателем, которому этот многочлен принадлежит. Если корни неприводимого многочлена являются порождающими (образующими) элементами мультипликативной группы поля, то корни называются примитивными, а сам неприводимый многочлен — примитивным.*

Принципиальное различие между примитивным и неприводимым многочленами продемонстрировано мультипликативными группами (3.4.11), (3.4.12), (3.4.14), (3.4.15). Первые две построены по модулям примитивных многочленов, соответственно,

$$x^4 + x + 1, \quad x^4 + x^3 + 1,$$

и возведение в последовательные степени их корней порождает целиком мультипликатиные группы полей.

Обе группы (3.4.14) и (3.4.15) построены по модулю неприводимого многочлена $x^4 + x^3 + x^2 + x + 1$, который не является

примитивным. Поэтому возведение в последовательные степени его корня ξ не дает всей мультипликативной группы поля $GF(2^4)$, а дает только ее подгруппу 5-го порядка. Чтобы выйти из положения и все-таки построить поле по модулю неприводимого многочлена $x^4 + x^3 + x^2 + x + 1$ (а вследствие общей теории такое построение выполнимо), пришлось возводить в последовательные степени не корень ξ, а элемент $\xi + 1$.

Для сопряженных элементов поля, разумеется, результаты совпадают.

Ясно, что все утверждения раздела 2.8 справедливы для мультипликативной группы поля Галуа.

П р и м е р 3. 10.

Рассмотрим поле $GF(2^6)$ и ее мультипликативную группу $GF^*(2^6)$. Эта группа изучалась в примере 2.3. Два делителя 2 и 3 степени расширения поля дают два подполя $GF(2^2)$, $GF(2^3) \subset GF(2^6)$, простым подполем всех трех полей является поле $GF(2)$. Их мультипликатиные группы $GF^*(2^2)$, $GF^*(2^3)$ являются подгруппами группы $GF^*(2^6)$. В примере 2.3. они обозначены, соответственно H_{21}, H_9. Кроме того, $H_{21} = GF^*(2^2) \subset H_7$, и H_7, равно как и H_3, не является мультипликативной группой никакого подполя. Легко заметить, что сказанное непосредственно связано с тем, что $x^{63} - 1$ делится на $x - 1$, $x^3 - 1$, $x^7 - 1$, $x^9 - 1, x^{21} - 1$. Многочлен $x^3 - 1$ делит многочлены $x^9 - 1$, $x^{21} - 1$, а на многочлен $x - 1$ делятся все перечисленные. Отвлекаясь от конкретного примера, рассмотрим канонический и наиболее распространенный метод описания конструкции конечного поля и его мультипликативной группы. Он использует так называемые круговые многочлены. Название "круговые многочлены" восходит к следующей задаче. Пусть n есть натуральное число. Корни двучлена $x^n - 1$ в произвольном поле называют корнями n-й степени из единицы (ср. с (3.2.4)). Для произвольного корня n-й степени из единицы β получается $\beta^n = 1$. Если K есть поле комплексных чисел, то эти корни можно представить как точки на единичном круге:

$$\beta = e^{i\gamma} = \cos\gamma + i\sin\gamma. \tag{3.5.16}$$

Угол γ удовлетворяет условию $n\gamma = k \cdot 2\pi$; $k = 0, 1, \ldots, n - 1$. Полученные n точек на круге делят его на n равных дуг.

Все корни из единицы образуют циклическую группу. Пусть μ есть ее примитивный элемент. Для каждого натурального делителя d числа n назначим все элементы порядка d корнями многочлена $Q_d(x)$. Он называется круговым многочленом.

Единица группы будет корнем только одного кругового многочлена $Q_1 = x-1$, и все степени μ^t, где $(t,n)=1$, будут корнями многочлена $Q_n(x)$. Его степень равна числу примитивных элементов группы корней $n-$й степени из единицы, т.е. $\varphi(n)$.

Имеем

$$x^n - 1 = \prod_{i=d|n} Q_i(x). \tag{3.5.17}$$

Пусть α есть примитивный элемент поля $GF(q^m)$. Вообразим теперь, что мультипликативная группа поля $GF(q^m)$ расположена в виде замкнутого круга, т.е. за элементом α^{q^m-2} следует $\alpha^{q^m-1} = \alpha^0$. Тогда $n = q^m - 1$, и порядки d всех элементов мультипликативной группы суть делители числа $q^m - 1$. Если многочлен $\psi_d(x)$ имеет своими корнями все элементы одного порядка d, то

$$x^{q^m-1} - 1 = \prod_{d|q^m-1} \psi_d(x).$$

Произведение распространено на все натуральные делители d числа $q^m - 1$. Имеет место полная аналогия с задачей деления круга. Каждый нормированный примитивный многочлен, который делит $x^{q^m-1} - 1$, имеет степень m. Поэтому круговой многочлен $\psi_{q^m-1}(x)$ распадается в произведение $\varphi(q^m - 1)/m$ нормированных примитивных многочленов степени m.

Продолжим пример. Пусть поле $GF(2^6)$ построено по модулю неприводимого многочлена $m_1(x) = x^6+x+1$, и $\alpha^6+\alpha+1 = 0$. Тогда, в соответствии со сказанным выше, и зная, что порядками элементов мультипликативной группы поля $GF(2^6)$ являются делители числá 63, т.е. 1, 3, 7, 9, 21, 63 получим:

$$\psi_1(x) = x - 1,$$

$$\psi_3(x) = \frac{x^3-1}{\psi_1(x)} = m_{21}(x) = x^2 + x + 1,$$

$$\psi_7(x) = \frac{x^7-1}{\psi_1(x)} = m_9(x)m_{27}(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 =$$

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

$$\psi_9(x) = \frac{x^9-1}{\psi_1(x)\psi_3(x)} = m_7(x) = x^6 + x^3 + 1,$$

$$\psi_{21}(x) = \frac{x^{21}-1}{\psi_1(x)\psi_3(x)\psi_7(x)} = m_3(x)m_{15}(x) =$$

$$= (x^6+x^4+x^2+x+1)(x^6+x^5+x^4+x^2+1),$$

$$\psi_{63}(x) = \frac{x^{63}-1}{\psi_1(x)\psi_3(x)\psi_7(x)\psi_9(x)\psi_{21}(x)} =$$

$$= m_1(x)m_5(x)m_{11}(x)m_{13}(x)m_{23}(x)m_{31}(x) =$$

$$= (x^6+x+1)(x^6+x^5+x^2+x+1)(x^6+x^5+x^3+x^2+1)\times$$

$$\times(x^6+x^4+x^3+x+1)(x^6+x^5+x^4+x+1)(x^6+x^5+1),$$

$$x^{63}-1 = \psi_1(x)\psi_3(x)\psi_7(x)\psi_9(x)\psi_{21}(x)\psi_{63}(x) =$$

$$= (x-1)m_{21}(x)m_9(x)m_{27}m_7(x)m_3(x)m_{15}\times$$

$$\times(x)m_1(x)m_5(x)m_{11}(x)m_{13}(x)m_{23}(x)m_{31}(x). \qquad (3.5.18)$$

Заметим, что все сопряженные элементы поля — одного порядка, но не все элементы одного порядка являются сопряженными. Они распадаются на комплекты сопряженных. Ниже комплекты сопряженных элементов поля $GF(2^6)$ расположены в той же последовательности, что и отвечающие им минимальные функции в произведении (3.5.18):

$$(\alpha^0=1),\ (\alpha^{21},\ \alpha^{42}),\ (\alpha^9,\ \alpha^{18},\ \alpha^{36}),\ (\alpha^{27},\ \alpha^{54},\ \alpha^{45}),$$

$$(\alpha^7,\ \alpha^{14},\ \alpha^{28},\ \alpha^{56},\ \alpha^{49},\ \alpha^{35}),\ (\alpha^3,\ \alpha^6,\ \alpha^{12},\ \alpha^{24},\ \alpha^{48},\ \alpha^{33}),$$

$$(\alpha^{15},\ \alpha^{30},\ \alpha^{60},\ \alpha^{57},\ \alpha^{51},\ \alpha^{39}),\ (\alpha^1,\ \alpha^2,\ \alpha^4,\ \alpha^8,\ \alpha^{16},\ \alpha^{32}),$$

$$(\alpha^5,\ \alpha^{10},\ \alpha^{20},\ \alpha^{40},\ \alpha^{17},\ \alpha^{34}),\ (\alpha^{11},\ \alpha^{22},\ \alpha^{44},\ \alpha^{25},\ \alpha^{50},\ \alpha^{37}),$$

$$(\alpha^{13},\ \alpha^{26},\ \alpha^{52},\ \alpha^{41},\ \alpha^{19},\ \alpha^{38}),(\alpha^{23},\ \alpha^{46},\ \alpha^{29},\ \alpha^{58},\ \alpha^{53},\ \alpha^{43}),$$

$$(\alpha^{31},\ \alpha^{62},\ \alpha^{61},\ \alpha^{59},\ \alpha^{55},\ \alpha^{47}).$$

В таком же порядке перечислены порядки корней минимальных функций: 1, 3, 7, 7, 9, 21, 21, 63, 63, 63, 63, 63, 63.

Определение 3.5.22. *Комплект показателей степеней в комплекте сопряженных элементов называется циклотомическим классом. Если i минимальный показатель в этом комплекте, то в соответствии с теоремами* 3.5.12 *и* 3.5.18 *циклотомическим классом будет $i, iq, iq^2, \ldots, iq^{t-1}$, где t — степень неприводимого многочлена, корнями которого являются упомянутые сопряженные элементы. В случае поля $GF(q^m)$ указанные показатели степеней приводятся по модулю q^m-1.*

В примерах 3.9. и 3.10. циклотомические классы усматриваются непосредственно из комплектов сопряженных элементов полей $GF(2^m)$ на случай $m = 2, 3, 4, 5, 6$. Разумеется, нуль — всегда циклотомический класс, так как неприводимый многочлен $m(x) = x - 1$ имеет своим корнем $\alpha^0 = 1$.

На случай $q = 3$ при $m = 2$ и $m = 3$ циклотомическими классами соответственно будут: (0), (1, 3), (2, 6), (4), (5, 7); (0), (1, 3, 9), (2, 6, 18), (4, 12, 10), (5, 15, 19), (7,21,11), (8, 24, 20), (13), (14, 16, 22), (17, 25, 23).

В заключение этого раздела проведём следующее рассуждение. Все построения конечных полей, предпринятые в этой главе, приводили к полям, порядок которых оказывался степенью простого числа. Справедлива

Теорема 3.5.23. *Прядок конечного поля всегда есть степень простого числа.*

Д о к а з а т е л ь с т в о. Конечное поле F есть конечное векторное пространство над своим подполем K порядка q. Пусть размерность этого пространства есть m, и пусть $b_1, b_2, \ldots, b_m$ есть базис этого пространства над подполем K. Тогда каждый элемент поля F однозначно представим в виде линейной комбинации $a_1b_1 + a_2b_2 + \ldots, +a_mb_m$, где $a_1, a_2, \ldots, a_m \in K$. Каждый коэффициент a_i может принимать q значений, и потому порядок поля F над K есть q^m. Далее, если поле K конечно, то его характеристика p есть простое число. Это означает, что его простое подполе есть $GF(p) = \{0, 1, 2, \ldots p-1\}$. Последнее следует из того, что включая в себя с необходимостью элементы 0 и 1, оно, благодаря замкнутости относительно сложения, содержит и все элементы $GF(p)$. Остаётся применить к полю K, как к пространству над $GF(p)$, первую часть доказательства и получить равенство $q = p^n$.

3.6. Изоморфизм полей Галуа

Уже отмечалось, что каков бы ни был неприводимый над $GF(q)$ многочлен $p(x)$ степени m, по модулю которого построено поле $GF(q^m)$, все элементы поля являются корнями многочлена $x^{q^m} - x$, и в этом смысле есть только одно поле данного порядка. Однако в зависимости от многочлена $p(x)$ поля могут обладать различными частными свойствами. Например, два поля (3.4.11) и (3.4.12) обладают тем свойством, что каждый из корней многочленов, по модулям которых построены оба поля, является примитивным элементом соответствующих мультипликативных групп полей. Но есть еще один многочлен $x^4 + x^3 + x^2 + x + 1$, корень γ которого не примитивный элемент поля. Действительно, так как $\gamma^4 + \gamma^3 + \gamma^2 + \gamma + 1 = 0$, то $\gamma^4 = \gamma^3 + \gamma^2 + \gamma + 1$. отсюда получаем, что $\gamma^5 = 1$, т.е. порядок элемента γ равен 5, а не 15. Таким образом, поля Галуа одного порядка, построенные по модулям разных неприводимых многочленов, не совпадают. Зато они *изоморфны.*

Определение 3.6.1. *Два поля A и B изоморфны, если между их элементами $\alpha \in A$ и $\beta \in B$ существует такое взаимнооднозначное соответствие, которое сохраняет операции сложения и умножения.*

Теорема 3.6.2. *Все конечные поля одного порядка изоморфны друг другу.*

Доказательство. Пусть α есть примитивный элемент поля A, и $m(x)$ его минимальный многочлен: $m(\alpha) = 0$. Многочлен $m(x)$ примитивный. Согласно теореме 3.5.5 $m(x)$ делит двучлен $x^{q^m} - x$. Пусть β есть примитивный элемент поля B. Так как каждый элемент поля B есть корень двучлена $x^{q^m} - x$, то в поле B найдется такой элемент β^i, что $m(\beta^i) = 0$, причем, будучи корнем примитивного многочлена, элемент β^i сам примитивный. Поле A есть множество многочленов от α степени не выше, чем $m - 1$, а поле B есть множество многочленов от β степени не выше, чем $m - 1$. Соответствие $\alpha \Leftrightarrow \beta^i$ как раз задает изоморфизм полей A и B.

В соответствии с определением изоморфизма и доказанной теоремой, сохранение операций сложения и умножения при установлении соответствия $\alpha \Leftrightarrow \beta^i$ означает, что каковы бы ни

были u_1 и u_2, при условии $\alpha^{u_1} \Leftrightarrow (\beta^i)^{u_1}$ и $\alpha^{u_2} \Leftrightarrow (\beta^i)^{u_2}$, из соотношений $\alpha^{u_1} + \alpha^{u_2} = \alpha^{u_3}$ и $(\beta^i)^{u_1} + (\beta^i)^{u_2} = (\beta^i)^{u_3}$, согласно тому же взаимно однозначному соответствию, с необходимостью следует $\alpha^{u_3} \Leftrightarrow (\beta^i)^{u_3}$.

Что касается операции умножения, то при тех же условиях должно быть $\alpha^{u_1+u_2} \Leftrightarrow (\beta^i)^{(u_1+u_2)}$.

П р и м е р 3. 11.

Пусть поле A есть поле 3.4.11, и поле B — поле 3.4.12. Вспомним случай $m = 4$ в примере 3.9. Многочлен x^4+x+1 является минимальным многочленом элемента α в поле 3.4.11 и минимальным многочленом элемента β^7 в поле 3.4.12. Изоморфизм двух этих полей устанавливается соответствием $\alpha \Leftrightarrow \beta^7$. Т.е. $i = 7$. Положим $u_1 = 2, u_2 = 6$. Тогда $\alpha^2 \Leftrightarrow \beta^{14}$, $\alpha^6 \Leftrightarrow \beta^{42} = \beta^{12}$, $\alpha^2 + \alpha^6 \Leftrightarrow \beta^{14} + \beta^{12}$. При этом в поле 3.4.11 $\alpha^2 + \alpha^6 = \alpha^3$, а в поле 3.4.12 $\beta^{14} + \beta^{12} = \beta^6 = (\beta^7)^3$, так как $i_3 = 3$, $21 \equiv 6(\mathrm{mod}15)$), что и требовалось.

Демонстрация сохранения операции умножения выглядит тривиальной, и читатель справится с нею самостоятельно.

Обратим внимание на обоснованность выбора элемента β^7.

Действительно, вместе с корнем β^7 корнями минимального многочлена $x^4 + x + 1$ в поле 3.4.12 будут также сопряженные элементы β^{14}, β^{13}, β^{11}.

Легко проверить, что, употребив в предыдущих выкладках любой из них, получим аналогичный результат.

В общем виде это утверждение выглядит следующим образом:

$$\alpha^{u_3} = \alpha^{u_1} + \alpha^{u_2} \Leftrightarrow (\beta^{iq^j})^{u_1} + (\beta^{iq^j})^{u_2} = (\beta^{iu_1} + \beta^{iu_2})^{q^j} = (\beta^{iq^j})^{u_3}, \tag{3.6.19}$$

где iq^j есть член циклотомического класса. Пусть, как и выше $u_1 = 2, u_2 = 6$, $i = 7$, $q = 2$. Пусть при этом $j = 2$. Изоморфизм устанавливается соотношением $\alpha \Leftrightarrow \beta^{13}$. Подставляя в (3.6.19) численные значения, по правилам вычисления в полях (3.4.11) и (3.4.12) получим $u_3 = 3$, $\alpha^3 \Leftrightarrow (\beta^{13})^3 = \beta^9$, что и доставляется суммой $\beta^{11} + \beta^3 = \beta^9$.

3.7. Автоморфизм поля Галуа

Определение 3.7.1. *Автоморфизмом σ поля Галуа $GF(q^m)$ над полем $GF(q)$ называется такое отображение поля $GF(q^m)$*

на себя, которое оставляет неподвижными элементы поля $GF(q)$.

Отображение $\sigma(\alpha) = \alpha^q$ является автоморфизмом. Действительно, если $\alpha \in GF(q)$, то $\alpha^q = \alpha$, так как порядок элемента α есть делитель порядка $q-1$ мультипликативной группы поля $GF(q)$, и, значит, $\alpha^{q-1} = 1$. Если же $\alpha \in GF(q^m)$, то отображение σ однозначно, так как различные элементы поля отображаются в различные. В самом деле, предположим противное, т.е. пусть при $\alpha \neq \beta \quad \alpha^q = \beta^q$. Тогда $\alpha^q - \beta^q = 0$, но $\alpha^q - \beta^q = (\alpha - \beta)^q = 0$, и потому $\alpha = \beta$, вопреки условию. Далее,

$$\sigma(\alpha + \beta) = (\alpha + \beta)^q = \alpha^q + \beta^q = \sigma(\alpha) + \sigma(\beta),$$

и

$$\sigma(\alpha\beta) = (\alpha\beta)^q = \alpha^q\beta^q = \sigma(\alpha)\sigma(\beta).$$

Обозначив $\sigma_j(\alpha) = \alpha^{q^j}$, $j = 0, 1, \ldots, m-1$, получим m автоморфизмов, которые переводят сопряженные элементы поля друг в друга. Других нет, так как максимальная степень неприводимого многочлена над полем $GF(q)$, корни которого принадлежат полю $GF(q^m)$, равна m, и потому любой комплект сопряженных элементов поля содержит не более m штук.

Множество всех m автоморфизмов σ_j замкнуто относительно операции последовательного выполнения двух автоморфизмов. Действительно,

$$\sigma_{j_1}\sigma_{j_2} = \sigma_{j_2}(\sigma_{j_1}(\alpha)) = \sigma_{j_2}(\alpha^{q^{j_1}}) = (\alpha^{q^{j_1}})^{q^{j_2}} = \alpha^{q^{j_1+j_2}} = \sigma_{j_1+j_2}.$$

Ассоциативность операции последовательных автоморфизмов проверяется непосредственно. Любой автоморфизм $\sigma_j = (\sigma_1)^j$. Таким образом, автоморфизмы образуют группу порядка m, и она циклическая с $\varphi(m)$ порождающими элементами σ_1 и σ_j, где $(j, m) = 1$. Кстати, вариативность выбора изоморфного соответствия $\alpha \Leftrightarrow \beta^{iq^j}$ в зависимости от $j = 0, 1, \ldots, m-1$ в предыдущем разделе как раз следует из наличия m автоморфизмов поля Галуа $GF(q^m)$.

Легко установить соответствие между подполями поля Галуа и подгруппами групы его автоморфизмов, стоит лишь разложить на множители степень m расширения поля.

3.8. Представление поля Галуа матрицами

Определение 3.8.1. *Сопровождающей матрицей нормированного многочлена*

$$p(x) = a_0 + a_1 x + \ldots + a_{m-1}x^{m-1} + x^m \tag{3.8.20}$$

называется $(m \times m)-$*матрица*

$$B = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{m-2} & -a_{m-1} \end{bmatrix}. \tag{3.8.21}$$

В этом разделе ограничимся многочленами над полем $GF(2)$, а потому опустим знаки минус.

Заметим, что, если многочлен (3.8.20) примитивный, то m строк матрицы (3.8.21) есть векторные представления элементов α, $\alpha^2, \ldots \alpha^m$ поля $GF(2^m)$, построенного по модулю многочлена (3.8.20), корнем которого является α.

Матрицу B можно возводить в степень. Помня правило умножения матриц "строка на столбец", запишем элемент c_{vu} матрицы B^2, стоящий на пересечении ее $v-$й строки и $u-$го столбца, пронумерованных, соответственно, сверху вниз и слева направо числами $0, 1, \ldots m-1$:

$$c_{vu} = \sum_{z=0}^{m-1} b_{vz} b_{zu}, \tag{3.8.22}$$

где b_{vz}, b_{zu} есть элементы $v-$й строки и $u-$го столбца.

Выполняя фактически операцию (3.8.22) для матрицы (3.8.21), получим: $B^2 =$

$$\begin{bmatrix} 0 & 0 & 1 & \ldots & 0 \\ . & . & . & \ldots & . \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 1 \\ a_0 & a_1 & a_2 & \ldots & a_{m-1} \\ a_{m-1}a_0 & a_{m-1}a_1 + a_0 & a_{m-1}a_2 + a_1 & \ldots & a_{m-1}a_{m-1} + a_{m-2} \end{bmatrix}. \tag{3.8.23}$$

Первые $m-1$ строк этой матрицы есть не что иное, как векторные представления элементов α^2, $\alpha^3, \ldots \alpha^m$ поля $GF(2^m)$,

построенного по модулю многочлена (3.8.20), корнем которого является α.

Последняя строка — это векторное представление элемента α^{m+1}. Действительно, если $a_{m-1} = 0$, то вектор элемента α^{m+1} получается простым сдвигом вправо вектора элемента α^m. При этом получится строка

$$0\, a_0\, a_1\, \ldots\, a_{m-2}. \tag{3.8.24}$$

Если же $a_{m-1} = 1$, то правило построения элементов поля по модулю многочлена (3.8.20) требует сложить вектор (3.8.24) с вектором

$$(a_0,\, a_1,\, \ldots\, , a_{m-1}). \tag{3.8.25}$$

Получится строка $0 + a_0,\, a_0 + a_1,\, \ldots\, , a_{m-2} + a_{m-1}$.

Теперь, используя доказанный факт как основание индукции, читатель самостоятельно сделает индукционный шаг, чтобы убедиться, что верна

Теорема 3.8.2. *Строками матрицы* $B^j,\ j = 0,\, 1,\, \ldots\, , 2^m - 2$ (3.8.21) *являются векторы длины* m *над* $GF(2)$, *изображающие элементы* $\alpha^j,\, \alpha^{j+1},\, \ldots\, , \alpha^{j+m-1}$, *где* α *есть корень примитивного многочлена* (3.8.20)

Так как произведение $B^{j_1} B^{j_2} = B^{j_1 + j_2}$ есть снова степень матрицы, то все степени сопровождающей матрицы образуют мультипликативную (циклическую) группу.

Нетрудно убедиться, что матрицы $B^j,\ j = 0,\, 1,\, \ldots\, , 2^m - 2$ можно поэлементно складывать. Присоединив к ним нулевую матрицу, получим поле 2^m матриц, изоморфное полю $GF(2^m)$.

П р и м е р 3.12.

Согласно таблице (3.4.11),

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

$$B^6 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \tag{3.8.26}$$

$$B^{14} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{3.8.27}$$

Согласно таблице (3.4.12),

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

$$B^6 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix},$$

$$B^{14} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

(Читатель не будет требовать замены α на β в формулировке теоремы 3.8.2 на случай поля (3.4.12)).

3.9. Задачи к главе 3

3.1. Построить поля $GF(2^5), GF(2^6), GF(3^2)$ по модулям многочленов, соответственно, $x^5+x^2+1, \quad x^6+x+1, \quad x^2+x+2$.
3.2. Найти все делители вида x^k-1 двучленов $x^{63}-1, x^{127}-$ $-1, x^{255}-1$. Найти все подполя полей $GF(2^i)$, где $i = 4, 5, 6, 7, 8$.
3.3. Какие из чисел 29, 30, 31, 32. являются порядками конечных полей?
3.4. Зная, что x^5+x^2+1 неприводим над $GF(2)$, найти все неприводимые над $GF(2)$ многочлены пятой степени.
3.5. Показать, что двучлены $x^2+1, \ x^2+x+4$ неприводимы над полем $GF(11)$.
3.6. Пусть $m = 2k+1, a, b \in GF(2^m)$. Показать, что если $a^2 + ab + b^2 = 0$, то $a = b = 0$.
3.7. Найти все примитивные элементы поля $GF(7)$.
3.8. Найти все примитивные элементы поля $GF(17)$.
3.9. Найти все примитивные элементы поля $GF(9)$.
3.10. Сколько примитивных элементов содержит поле $GF(q^m)$?
3.11. Показать, что любой квадратный многочлен над $GF(q)$ разлагается над $GF(q^2)$ на линейные множители.
3.12. Доказать, что многочлен $x^8+x^7+x^3+x+1$ неприводим над $GF(2)$, и найти его показатель.

3.13. Многочлен $f^*(x)$ называется двойственным многочлену $f(x)$ степени m, если $f^*(x) = x^m f(1/x)$. Показать, что
a) Оба многочлена одновременно неприводимы или нет.
b) Оба многочлена принадлежат одному показателю.
c) Если $f(x) = g(x)h(x)$, $g(x), h(x)$ неприводимые многочлены, и $f^*(x) = f(x)$, то либо $h(x) = g^*(x)$, либо $g^*(x) = g(x)$ и $h(x) = h^*(x)$.
3.14. Если $f^*(x) = f(x)$, то многочлен $f(x)$ называется самодвойственным. Какие многочлены с таким свойством являются примитивными и над какими полями?
3.15. Доказать,что если $f(x)$ неприводимый самодвойственный многочлен над $GF(q)$ степени $m > 1$ с показателем e, то каждый неприводимый многочлен над $GF(q)$ степени $d > 1$ с показателем $e^{'}|e$ самодвойственный.
3.16.Показать, что многочлен $x^6 + x^5 + x^2 + x + 1$ примитивен над $GF(2)$.
3.17. Показать, что многочлен $x^8 + x^6 + x^5 + x + 1$ примитивен над $GF(2)$.
3.18. Показать, что многочлен $x^5 - x + 1$ примитивен над $GF(3)$.
3.19. Найти число примитивных многочленов степени m над $GF(q)$.
3.20. Пусть m натуральное составное число. Доказать, что среди нормированных неприводимых многочленов степени m над $GF(q)$ найдется непримитивный.
3.21. Пусть m простое число. Доказать, что все нормированные неприводимые многочлены степени m над $GF(q)$ примитивны в том и только в том случае, когда $q = 2$ и $2^m - 1$ простое.
3.22. Показать, что если $GF(q)$ – конечное поле нечетной характеристики, то элемент $a \in GF^*(q)$ имеет в поле $GF(q)$ квадратный корень тогда и только тогда, когда $a^{(q-1)/2} = 1$.
3.23. Доказать, что для данного натурального числа k элемент $a \in GF^*(q)$ является k-й степенью некоторого элемента поля $GF(q)$ в том и только в том случае, если $a^{(q-1)/d} = 1$, где $d = ((q-1), k)$.
3.24. Доказать, что всякий самодвойственный неприводимый многочлен, отличный от $x + 1$, имеет четную степень.
3.25. Доказать, что всякий самодвойственный неприводимый многочлен степени m является делителем многочлена

$$x^{p^{m/2}+1} - 1 = H_m(x).$$

3.26. Степень n любого неприводимого делителя многочлена $H_m(x)$ есть делитель числа m.
3.27. Разложить на множители двучлен $x^8 - 1$ над $GF(3)$.
3.28.Доказать теорему Вильсона: $(p-1)! \equiv -1 (\mathrm{mod}\ p)$, где p — простое.
3.29. В некоторых руководствах порядок следования компонент векторов, изображающих элементы поля $GF(2^m)$, заменен на обратный по сравнению с принятым здесь. Сохранится ли в этом случае формулировка теоремы 3.8.2, если умножение матриц выполнять по правилу (3.8.22)? Если нет, то как оно должно быть изменено?
3.30. Показать, что все круговые многочлены поля $GF(q^m)$ являются многочленами над полем $GF(q)$.

Глава 4.

Линейные коды

4.1. Код как линейное векторное подпространство.

Теперь можно перейти к задачам построения кодов длины n с заданным кодовым расстоянием d между любыми двумя векторами.

В самом общем случае рассматривается код A, являющийся подпространством линейного векторного пространства (см. раздел 2.17). Такой код называют линейным кодом.

Всюду ниже и компоненты вектора, и скаляры принадлежат одному и тому же полю.

В качестве поля F ниже выступает поле Галуа $GF(q)$, где $q = p^m$, и p простое число. Если длина вектора есть n, то пространство V содержит q^n векторов. Иногда пространство V будет обозначаться символом $E_q^{(n)}$.

Интересен случай $m = 1$. Тогда $q = p$, и пространство V будет обозначаться символом $E_p^{(n)}$. Оно представляет собой совокупность всех p^n векторов длины n над полем $GF(p)$.

Вспомним, что множество всех этих p^n векторов образует группу с операцией поразрядного сложения векторов по модулю p. Кроме того, если $v \in E_p^{(n)}$, то $av \in E_p^{(n)}$, где $a \in GF(p)$, и av означает умножение всех компонент вектора v на a по модулю p.

На самом деле, при $m = 1$ условие 2) в определении 2.17.1 пространства можно опустить, т.е. достаточно потребовать, чтобы множество V было группой. Действительно, выполнимость умножения вектора на скаляр $a \in GF(p)$ есть не что иное, как сложение вектора с самим собой a раз. Если же $m > 1$, то это не так, ибо в таком случае умножение элементов $\alpha\beta$ по-

ля $GF(p^m)$ не сводится к сложению вектора α с самим собой β "раз". Это лишено смысла. Таким образом, код A является либо подпространством линейного пространства и называется линейным кодом, если $m > 1$, либо подгруппой группы $E_p^{(n)}$ и называется групповым кодом, если $m = 1$.

Всякое пространство является аддитивной группой. Обратное неверно.

Так как понятие "линейный код" включает в себя понятие "групповой код" , то не будет ошибкой всегда пользоваться первым, если нет специальной нужды подчеркивать, что код есть именно группа.

Всё пространство V не может быть помехоустойчивым кодом, так как минимальное кодовое расстояние между векторами равно 1, и ни исправление, ни обнаружение ошибок невозможно.

Действительно, любая ошибка в векторе переведёт его в другой вектор того же кода.

Таким образом, для того чтобы код A был помехоустойчивым, необходимо выполнение условия $A \subset V$.

Перейдя в этой главе к кодам над $GF(q)$, надлежит перейти от двоичного симметричного канала к *q-ичному симметричному каналу*. Пусть передаваемый символ сохраняет свое значение с вероятностью $1 - \tau$, и искажается под действием помехи с вероятностью τ. Если при этом все $q - 1$ ошибочные значения возникают с одинаковой вероятностью $\tau/(q - 1)$, то такой канал и называется *q-ичным симметричным каналом.* Троичный симметричный канал схематически представлен на рис. 4.1.

4.2. Порождающая матрица кода

Теорема 4.2.1. *Размерность пространства V всех q^n векторов длины n над полем F из q элементов равна n.*

Действительно, n векторов

$$v_1 = (1, 0, \ldots, 0), v_2 = (0, 1, 0, \ldots, 0), \ldots, v_n = (0, 0, \ldots, 0, 1),$$

линейно независимы, и любой вектор $v = (a_1, a_2, \ldots, a_n) \in V$ может быть представлен в виде суммы

$$v = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n.$$

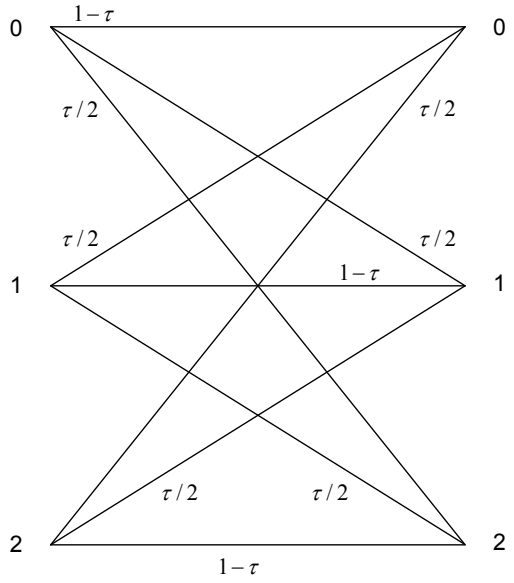

Рис. 4.1. Троичный симметричный канал

Если же в пространстве V найдется система $v_1^{'}$, $v_2^{'}$, $\ldots v_{n+1}^{'}$ из $n+1$ линейно независимых векторов, то линейных комбинаций этих векторов будет в точности q^{n+1}. Однако всего в пространстве V имеется только q^n векторов. Следовательно, найдутся две одинаковые линейные комбинации

$$a_1 v_1^{'} + a_2 v_2^{'} + \ldots + a_1 v_{n+1}^{'} = b_1 v_1^{'} + b_2 v_2^{'} + \ldots + b_{n+1} v_{n+1}^{'},$$

где по крайней мере для одного i будет $a_i \neq b_i$. Это означает, что в левой части равенства

$$(a_1 - b_1) v_1^{'} + (a_2 - b_2) v_2^{'} + \ldots + (a_i - b_i) v_i^{'} \ldots + (a_{n+1} - b_{n+1}) v_{n+1}^{'} = 0$$

по крайней мере в одном слагаемом $(a_i - b_i) v_i^{'}$ коэффициент $(a_i - b_i)$ отличен от нуля. Отсюда получается, что векторы $v_1^{'}$, $v_2^{'}$, $\ldots v_{n+1}^{'}$ линейно зависимы вопреки предположению. Этим завершается доказательство теоремы.

Переходя к линейному коду A и условившись, что он должен быть подпространством, т.е. подгруппой по определению, найдем число векторов кода, т.е. порядок кода.

Так как порядок подгруппы есть делитель порядка группы и так как порядок пространства V равен q^n, то порядок подпространства A всегда есть степень числа q.

Положим, порядок кода A будет q^k. Тогда базис кода как подпространства содержит k линейно независимых векторов.

Пусть

$$v_1 = (a_{11}, \ldots, a_{1n}), v_2 = (a_{21}, \ldots, a_{2n}), \ldots, v_k = (a_{k1}, \ldots, a_{kn})$$

суть k выбранных векторов базиса. Расположим их в виде строк матрицы:

$$G = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \ldots & a_{kn} \end{bmatrix}.$$

Эта матрица называется *порождающей матрицей* кода A. Если бы при передаче по каналу связи не было опасности возникновения ошибок, то каждое из q^k сообщений можно было бы передавать посредством вектора $u = (u_1, u_2, \ldots, u_k)$ длины k. Однако ради создания нужного нам расстояния между кодовыми векторами вектор u должен быть подвергнут некоторому преобразованию. Оно состоит в том, что вектор u задаёт линейную комбинацию строк матрицы G.

Именно, вектор $v \in A$, соответствующий вектору u, который в отсутствие помех должен был бы передаваться по каналу связи, получается следующим образом:

$$v = uG. \tag{4.2.1}$$

Иначе говоря, строка $(a_{i1}\, a_{i2} \ldots a_{in})$, $(i = 1, 2, \ldots, k)$ матрицы G умножается на u_i, и все k произведений $u_i a_{i1}$, $u_i a_{i2}$, $\ldots$, $u_i a_{in}$ складываются поразрядно по правилам сложения элементов поля $GF(q)$. Получается кодовый вектор

$$v = (\sum_{i=1}^{k} u_i a_{i1}, \sum_{i=1}^{k} u_i a_{i2}, \ \ldots, \sum_{i=1}^{k} u_i a_{in},). \tag{4.2.2}$$

Эта операция называется кодированием вектора u в вектор v кода A. Остаётся наделить код A, а значит, матрицу G, требуемыми метрическими свойствами, т.е. свойствами помехоустойчивости. Попутно заметим, что одно и то же подпространство

A обладает множеством базисов. Подсчитаем их количество. Первый вектор v_1 базиса может быть выбран $q^n - 1$ способами из всех ненулевых векторов подпростриства. Второй вектор v_2 базиса подпространства размерности k, т.е. подпространства порядка q^k, может быть выбран $q^n - q$ способами из всех векторов, оставшихся после выбора первого и всех его кратных. Третий вектор v_3 может быть выбран $q^n - q^2$ способами из векторов, не являющихся линейными комбинациями уже выбранных двух. Продолжая это процесс, выберем вектор v_{k-1} и построим все q^{k-1} линейные комбинации уже выбранных векторов $v_1, v_2, \ldots, v_{k-1}$. Последний вектор v_k может быть выбран $q^n - q^{k-1}$ способами из векторов , не являющихся линейными комбинациями уже выбранных. Итак, выбраны

$$(q^n - 1)(q^n - q) \ldots (q^n - q^{k-1})$$

комплектов базисных векторов.

Любой из базисных векторов может быть выбран на любом из k шагов описанной выше процедуры. Поэтому среди выбранных базисов есть все, которые отличаются друг от друга только нумерацией базисных векторов. Значит, точное число базисов подпространства A, которые не могут быть получены друг из друга перестановкой строк порождающей матрицы, равно

$$\frac{(q^n - 1)(q^n - q) \ldots (q^n - q^{k-1})}{k!}.$$

4.3. Проверочная матрица кода

Наряду с подпространством $A \subset V$ рассмотрим такое подпространство $B \subset V$, что B содержит в точности q^{n-k} векторов, и для любой пары векторов $v \in A$, $h \in B$ скалярное произведение $< v, h >= 0$. Подпространства $A \subset V$ и $B \subset V$ называются взаимно ортогональными.

Порождающую матрицу подпространства B будем изображать следующим образом:

$$H = \begin{bmatrix} h_{11} & h_{12} & \ldots & h_{1n} \\ h_{21} & h_{22} & \ldots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{(n-k)1} & h_{(n-k)2} & \ldots & h_{(n-k)n} \end{bmatrix}$$

Из ортогональности подпространств A и B следует, что, каков бы ни был вектор $v \in A$, всегда $vH^T = 0$, где H^T означают транспонированную матрицу H. На деле это означает, что скалярные произведения вектора v на каждую из строк матрицы H равны нулю.

Соотношение $vH^T = 0$ *проверяет* принадлежность вектора v к коду A, и потому матрицу H называют *проверочной матрицей* кода A. Так как кодовый вектор v есть линейная комбинация строк матрицы G, то соотношение

$$GH^T = [0], \tag{4.3.3}$$

где $[0]$ нулевая матрица размера $k \times (n-k)$, является необходимым и достаточным условием ортогональности подпространств A и B размерностей, соответственно k и $n-k$.

Иногда ради краткости будем называть обе рассмотренные матрицы базисными матрицами кода A и ортогонального ему кода B.

4.4. Каноническая форма базисных матриц

Обозначим для удобства порождающую матрицу и порождаемый ею код соответственно символами G' и A'. Некоторые операции над строками порождающей матрицы G' не изменяют всего подпространства A'. Таковыми, очевидно, являются: любая перестановка строк, умножение любой строки на произвольный, отличный от нуля, элемент поля, сложение любой линейной комбинации некоторых строк с произвольной строкой. Эти операции называют элементарными.

Посредством элементарных операций преобразуем матрицу G' к матрице G'' :

1. В i-й строке $(i = 1, 2, \ldots, k)$ матрицы найдется по крайней мере одна ненулевая компонента, так как базис подпростриства не может содержать нулевой строки. Пусть первая отличная от нуля компонента этой строки находится в j-м столбце. Разделим каждую компоненту строки на a_{ij}. В результате получится новая компонента $a_{ij}^{'}$ матрицы, равная единице.

2. К каждой z-й строке $(z \neq i)$ прибавим i-ю строку, умноженную на $-a_{zj}$. В результате в j-м столбце i-я строка будет содержать единицу, а все остальные строки — нули.

Заметим, что как только какой-нибудь столбец будет содержать единицу в одной из строк и нули в остальных строках, то ни первая, ни вторая операции над строками уже не

смогут изменить этот столбец. Таким образом, после того как эти операции будут проделаны над каждой строкой, получится матрица G'', содержащая k столбцов, каждый из которых содержит единицу и $k-1$ нулей, причем единица появится в каждой строке.

К элементарным операциям можно добавить ещё одну — перестановку столбцов матрицы. В результате получится матрица G. Разумеется, эта операция изменит не только базис, но и порождаемое им подпространство A'. Однако новое подпространство A, порождаемое матрицей G, будет обладать теми же метрическими свойствами, что и подпространство A' : все попарные расстояния между его векторами останутся прежними.

Подходящей комбинацией перечисленных выше операций всегда можно преобразовать матрицу G' к виду

$$G = [I_{k\times k}, P_{k\times(n-k)}], \tag{4.4.4}$$

где $I_{k\times k}$ — единичная матрица. Это и есть каноническая форма порождающей матрицы. Иногда её называют приведённо-ступенчатой формой.

П р и м е р 4. 1.

Пусть

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

Прибавим к первой строке вторую, а к третьей — четвёртую. Новая матрица будет:

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

Матрицы G' и G'' порождают одно и то же подпространство, так как одна из другой получены элементарными операциями. Если теперь в G'' поменять местами четвёртый и седьмой столбцы, получим матрицу

$$G = \left[\begin{array}{cccc:ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right]. \tag{4.4.5}$$

Матрица G имеет каноническую форму. Матрицы G' и G порождают различные подпространства, т.е. групповые коды, однако их корректирующие способности совпадают. Легко заметить, что кодирование посредством матрицы G в канонической форме сохраняет все символы вектора $u = (u_1, u_2, \ldots, u_k)$ длины k в качестве первых k символов кодового вектора. Эти символы называют *информационными* символами. Остальные $n-k$ символов называются *проверочными* и (или) *избыточными*. Код длины n с k информационными символами называют линейным (групповым) (n,k)-кодом. Иногда употребляют обозначение (n,k,d)-код, имея в виду ещё и его кодовое расстояние. Вообще говоря, информационные символы кодового вектора не обязаны предшествовать проверочным. Во всяком случае, если информационные символы отделены от проверочных, код называется систематическим. Точнее, код называется систематическим, если среди n номеров (т.е. нижних индексов) компонент $a_1, a_2, \ldots, a_n$ кодового вектора можно указать такие k номеров $i_1, i_2, \ldots, i_k$, одинаковые для всех векторов кода, что $a_{i_1} = u_1$, $a_{i_2} = u_2$, $a_{i_k} = u_k$, где u_j, $(j = 1, 2, \ldots, k,)$ символы информационного вектора u.

Теорема 4.4.1. *Если*

$$G_K = [I_{k\times k}, P_{k\times(n-k)}]$$

есть каноническая форма порождающей матрицы, то каноническая форма проверочной матрицы есть

$$H_K = [-P^T_{(n-k)\times(k)} I_{(n-k)\times(n-k)}]. \tag{4.4.6}$$

Д о к а з а т е л ь с т в о. Пусть в развёрнутом виде

$$G_K = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & p_{11} & p_{12} & \ldots & p_{1,n-k} \\ 0 & 1 & 0 & \ldots & 0 & p_{21} & p_{22} & \ldots & p_{2,n-k} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & p_{k1} & p_{k2} & \ldots & p_{k,n-k} \end{bmatrix}$$

и пусть

$$H_K = \begin{bmatrix} -p_{11} & -p_{21} & \ldots & -p_{k,1} & 1 & 0 & 0 & \ldots & 0 \\ -p_{12} & -p_{22} & \ldots & -p_{k,2} & 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ -p_{1,n-k} & -p_{2,n-k} & \ldots & -p_{k,n-k} & 0 & \ldots & \ldots & 0 & 1 \end{bmatrix}.$$

Найдём скалярное произведение i-й строки матрицы G_K на j-ю строку матрицы H_K. Сравнивая расположение единичных диагональных матриц, легко видеть, что скалярное произведение будет содержать в точности два слагаемых, в которых элементы $-p_{ij}$ и p_{ij} умножаются на 1. Остальные элементы p_{xy} умножаются на нули. Отсюда сразу ясно, что скалярное произведение сводится к сумме $-p_{ij} + p_{ij} = 0$.

В подробном формальном виде доказательство выглядит так:

$$(0, \ldots, 0, 1, 0, \ldots, 0, p_{i1}, p_{i2}, \ldots, p_{ij}, \ldots, p_{i,n-k}) \times$$

$$\times(-p_{1j}, -p_{2j}, \ldots, -p_{ij}, \ldots, -p_{kj}, 0, \ldots, 0, 1, 0 \ldots, 0) =$$

$$= -p_{1j} \cdot 0 - p_{2j} \cdot 0 - \ldots - p_{i-1,j} \cdot 0 - p_{ij} \cdot 1 - p_{i+1,j} \cdot 0 - \ldots - p_{kj} \cdot 0 +$$

$$+0 \cdot p_{i1} + 0 \cdot p_{i2} + \ldots + 0 \cdot p_{i,j-1} + 1 \cdot p_{ij} + 0 \cdot p_{i,j+1} + \ldots + 0 \cdot p_{i,n-k} =$$

$$= -p_{ij} + p_{ij} = 0.$$

Это и означает, что подпространства, порождаемые матрицами G и H, ортогональны. В приведенном выше примере получим

$$H = \left[\begin{array}{cccc:ccc} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right]. \qquad (4.4.7)$$

Это проверочная матрица уже знакомого нам кода Хэмминга. Читатель легко проверит попарную ортогональность строк матриц (4.4.5) и (4.4.7). Та же матрица получится простой перестановкой столбцов матрицы (0.4.11).

Если польза канонической формы порождающей матрицы состоит в том, что с её помощью операция кодирования сохраняет информационные символы кодового вектора, то польза канонической формы проверочной матрицы заключается в том, что она непосредственно демонстрирует. как проверочные символы кодового вектора получаются из данного набора информационных символов, т.е. простейшим образом выполняет процедуру кодирования. В самом деле, пусть кодовый вектор имеет вид

$$v = (u_1, u_2, \ldots, u_k, c_1, c_2, \ldots, c_{n-k}),$$

где первые k символов — информационные. Умножим скалярно этот вектор на i-ю строку проверочной матрицы:

$$-u_1 p_{1i} - u_2 p_{2i} - \ldots - u_k p_{ki} + c_i = 0, \qquad (4.4.8)$$

откуда проверочный символ c_i, $\quad i = 1, 2, \ldots, n-k$, получается в виде суммы.

$$c_i = u_1 p_{1i} + u_2 p_{2i} + \ldots + u_k p_{ki}.$$

Таким образом, для каждого из q^k наборов информационных символов посредством констант проверочной матрицы получаются все проверочные символы кодового вектора.

П р и м е р 4. 2.

Пусть

$$v = (u_1, u_2, u_3, u_4, c_1, c_2, c_3),$$

есть вектор кода с порождающей матрицей (4.4.7). Найдем все три его скалярные произведения на строки матрицы (4.4.7):

$$u_2 + u_3 + c_1 = 0,$$

$$u_1 + u_2 + u_3 + u_4 + c_2 = 0,$$

$$u_1 + u_2 + u_4 + c_3 = 0,$$

откуда (помня, что в $GF(2)$ $-x = x$)

$$c_1 = u_2 + u_3,$$

$$c_2 = u_1 + u_2 + u_3 + u_4,$$

$$c_3 = u_1 + u_2 + u_4.$$

Обращаясь к порождающей матрице G_K , нетрудно заметить, что полученное равенство (4.4.8) есть не что иное, как скалярное произведение вектора $u = (u_1, u_2, \ldots, u_k)$ на i-й столбец матрицы P, что находится в полном согласии со способом получения канонической формы матрицы H_K из канонической формы матрицы G_K.

Да и вообще, любой символ кодового вектора получается как скалярное произведение вектора $u = (u_1, u_2, \ldots, u_k)$ на соответствующий столбец матрицы G, в какой бы форме она ни задавалась.

4.5. Проверочная матрица и минимальное расстояние кода

Определение 4.5.1. *Весом $w(v)$ вектора v называется число отличных от нуля его компонент.*

Теорема 4.5.2. *Любому значению расстояния $d(v_1, v_2)$ между векторами v_1 и v_2 линейного (n, k)-кода отвечает кодовый вектор $v_1 - v_2 = v$, для веса $w(v)$ которого выполняется равенство $w(v) = d(v_1, v_2)$. И, наоборот, каждому значению $w(v)$ веса кодового вектора v отвечает пара кодовых векторов v_1 и v_2 с расстоянием $d(v_1, v_2) = w(v)$, причём таких пар имеется в точности q^k.*

Д о к а з а т е л ь с т в о. В силу того, что код — всегда группа с операцией поразрядного сложения векторов, разность двух кодовых векторов есть снова кодовый вектор: $v_1 - v_2 = v$, и вес $w(v)$ вектора v, т.е. число отличных от нуля его компонент в точности равно расстоянию $d(v_1, v_2)$. Наоборот, пусть вектор v имеет вес $w(v)$. Сложив вектор v с произвольным кодовым вектором v_i, получим, что $d(v+v_i, v_i) = w(v)$, чем и завершается доказательство. Остаётся вспомнить, что кодовых векторов v_i имеется в точности q^k.

Пусть $v = (a_1, a_2, \ldots, a_n)$ — кодовый вектор. Представим проверочную матрицу в виде

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

где h_i есть i-й вектор-столбец проверочной матрицы. Тогда выражение $vH^T = 0$ можно переписать в виде

$$a_1h_1 + a_2h_2 + \ldots + a_nh_n = 0.$$

Иначе говоря, каждый отличный от нуля кодовый вектор v задаёт нетривиальное соотношение линейной зависимости векторов-столбцов проверочной матрицы. Пусть $a_{i_1}, a_{i_2}\ \ldots\ , a_{i_w}$ — все отличные от нуля компоненты вектора v. Тогда равенство превратится в

$$a_{i_1}h_{i_1} + a_{i_2}h_{i_2} + \ldots + a_{i_w}h_{i_w} = 0.$$

Если $w \leq d-1$, то это означает, что имеется такой кодовый вектор, вес которого не превосходит $d-1$, а значит, найдутся такие пары векторов, расстояния между которыми не превосходит $d-1$. Тем более, минимальное расстояние оказывается меньше чем d.

С другой стороны, если любые $d-1$ столбцов проверочной матрицы линейно независимы, то минимальный вес, а значит минимальное расстояние кода, не менее d.

Таким образом, доказана

Теорема 4.5.3. *Для того, чтобы минимальное расстояние линейного кода было не менее, чем d, необходимо и достаточно, чтобы любые $d-1$ и менее столбцов проверочной матрицы были линейно независимы.*

Эта теорема полностью описывает метрические свойства проверочной матрицы.

Есть существенное различие между требованием теоремы 4.5.3 и понятием ранга матрицы. Ранг матрицы по столбцам определяется хотя бы одним максимальным набором её линейно независимых столбцов. От матрицы H же требуется, чтобы *любые* $d-1$ ее столбцов были бы линейно независимыми.

Ясно, конечно, что количество *любых* линейно независимых столбцов не может превосходить ранга матрицы H. Он же в свою очередь не может превосходить размерности $n-k$ пространства всех q^{n-k} векторов длины $n-k$, из которого и набираются столбцы матрицы H.

Этот факт выражается неравенством

$$d-1 \leq n-k, \tag{4.5.9}$$

которое называется границей Синглтона. Граница Синглтона может быть выведена также из канонической формы порождающей матрицы. Действительно, строка порождающей матрицы, вообще, есть кодовый вектор, а строка матрицы в ее канонической форме, в частности, имеет вес, который не может превышать величины $n-k+1$.

Методы построения проверочных матриц будут обсуждаться ниже.

Из теоремы 4.5.3 выводится граница существования линейного над $GF(q)$ кода с параметрами n, k, d.

Будем строить проверочную матрицу H размера $(n-k)\times n$ следующим образом. В качестве первого столбца h_1 можно выбрать любой ненулевой столбец длины $n-k$. Вторым столбцом h_2 может стать любой из оставшихся $q^{n-k}-q$ столбцов, кроме нулевого и $q-1$ столбцов, кратных столбца h_1. Предположим, что выбрано уже j столбцов. Имеется не более

$$(q-1)C_j^1 + (q-1)^2 C_j^2 + \ldots + (q-1)^{d-2} C_j^{d-2}$$

их различных линейных комбинаций, содержащих $d-2$ и менее столбцов. Если эта величина меньше, чем $q^{n-k}-1$, то можно добавить еще один ненулевой столбец, который отличен от всех

этих линейных комбинаций. Тогда никакие $d-1$ столбцов из выбранных $j+1$ столбцов не будут линейно зависимы.

Смысл этого рассуждения состоит в том, что каков бы ни был комплект $d-2$ или менее столбцов из уже выбранных столбцов, добавление к нему еще одного столбца, не принадлежащего ни к одному из этих комплектов, не создаст комплекта из $d-1$ или менее линейно зависимых столбцов.

Подчеркнем основную черту способа выбора столбцов. Он таков, что среди уже выбранных столбцов любые $d-1$ и менее столбцов линейно независимы. Продолжая этот процесс, выбирают $(n-1)$-й столбец и из уже полученных столбцов образуют всевозможные линейные комбинации, содержащие $d-2$ и менее столбцов. Если выполняется соотношение

$$(q-1)C_{n-1}^1 + (q-1)^2 C_{n-1}^2 + \ldots + (q-1)^{d-2} C_{n-1}^{d-2} < q^{n-k} - 1, \tag{4.5.10}$$

то можно добавить еще один ненулевой $n-$й столбец, и любые $d-1$ и менее из этих n столбцов будут линейно независимы. Проверочная матрица построена.

Неравенство (4.5.10) называется границей Варшамова — Гилберта, т.е. границей существования линейного кода с упомянутыми параметрами.

Для реального построения проверочной матрицы этот способ, конечно, не годится, так как главную роль в нем играет перебор всех столбцов, и объем этого перебора имеет порядок q^n . Вполне реальные методы построения проверочной матрицы изложены ниже.

4.6. Декодирование линейного кода

Пусть по каналу связи отправлен кодовый вектор

$$u = (u_1, u_2, \ldots, u_n),\ u \in, A,$$

где, как и прежде, A есть кодовое подпространство, и в канале произошла ошибка, изображаемая вектором

$$e = (e_1, e_2, \ldots, e_n).$$

На приёмном конце принят вектор

$$v = u + e = (u_1 + e_1, u_2 + e_2, \ldots, u_n + e_n),\ v \in V,$$

и декодер вычисляет произведение vH^T. Имеем

$$vH^T = (u+e)H^T = uH^T + eH^T = eH^T,$$

так как

$$uH^T = 0$$

из-за того, что кодовый вектор u принадлежит нулевому подпространству матрицы H. Произведение

$$eH^T = S$$

называется *синдромом.*

Пусть

$$V = A_0 \cup A_1 \cup A_2 \cup \ldots \cup A_{q^{n-k}-1}$$

есть разложение пространства V по подпространству $A = A_0$, и A_i $(i = 0, 1, \ldots, q^{n-k} - 1)$, суть смежные классы (классы вычетов).

Каждый вектор v принадлежит одному и только одному смежному классу.

Теорема 4.6.1. *Все векторы одного и того же смежного класса имеют одинаковые синдромы, и различным смежным классам отвечают различные синдромы.*

Доказательство. Пусть v_1 и $v_2 \in A_i$, и пусть $v_1H^T = S_1$, $v_2H^T = S_2$; $S_1 - S_2 = v_1H^T - v_2H^T = (v_1 - v_2)H^T$. Так как v_1 и $v_2 \in A_i$, то $v_1 - v_2 = v \in A$, и $vH^T = 0$, откуда $S_1 = S_2$. Если же $v_1 \in A_{i_1}$, $v_2 \in A_{i_2}$, то $v_1 - v_2 = v \notin A$, $vH^T \neq 0, v_1H^T \neq v_2H^T, S_1 \neq S_2$, что и требовалось.

Уравнению eH^T, удовлетворяют в точности q^k векторов смежного класса A_i, отвечающего синдрому S_i.

Какой же вектор $a \in A_i$ следует выбрать в качестве вектора-ошибки e? Здесь алгебраические средства теряют силу. Вступают в действие соображения, суть которых восходит к условию (0.1.2). Обращение к этому условию тем более оправдано, что на деле выполняется значительно более сильное неравенство. Его правая часть в реальных условиях передачи информации значительно меньше чем $1/2$. Соглашение, которое лежит в основе выбора вектора-ошибки, основано на принципе декодирования по максимальному правдоподобию. Оно состоит в том, что вектором-ошибкой считается вектор минимального веса в

соответствующем смежном классе. Это означает, что либо в смежном классе содержится единственный вектор e минимального веса $w \le t$, и тогда истинный вектор u получается в виде разности $u = v - e$, либо в смежном классе имеется по крайней мере два вектора e_1, e_2 минимального веса w, и тогда нет оснований определить, какой из этих векторов является вектором-ошибкой. Таким образом, верна

Теорема 4.6.2. *Линейный код исправляет все независимые ошибки кратности t и менее тогда и только тогда, когда все векторы веса t и менее принадлежат различным смежным классам.*

С другой стороны, нам известно, что код, в том числе и линейный, исправляет все независимые ошибки кратности t и менее тогда и только тогда, когда для минимального расстояния выполняется условие $d \ge 2t + 1$.

Таким образом, два равносильных критерия сформулированы в различных выражениях. Естественным является намерение показать, что из одной формулировки следует другая. Справедлива

Теорема 4.6.3. *Минимальное расстояние линейного кода равно $d = 2t + 1$ тогда и только тогда, когда все векторы веса $1, 2, \ldots, t$ принадлежат различным смежным классам.*

Доказательство. Пусть все векторы веса $1, 2, \ldots, t$ принадлежат различным смежным классам. Это означает, в частности, что для произвольного кодового вектора u заведомо $w(u) > t$. Предположим, что расстояние $d < 2t + 1$, т.е. в коде найдётся вектор u веса

$$w(u) \le 2t.$$

Отметим произвольные t отличных от нуля компонент вектора u. В некотором смежном классе по условию заведомо найдётся вектор v_1, все отличные от нуля компоненты которого противоположны отмеченным компонентам вектора u. Тогда этому же смежному классу принадлежит вектор $v_2 = u + v_1$, вес которого $w(v_2) \le t$, что противоречит предположению.

Пусть, наоборот, нашлись два вектора v_1 и v_2 веса $w \le t$, принадлежащие одному смежному классу. Тогда, во-первых, вектор $u = v_1 - v_2$ принадлежит кодовому подпространству A, а, во-вторых, вес $w(v_1 - v_2) \le 2t$, что противоречит условию $d \ge 2t + 1$. Теорема доказана.

Употребляя для описанной выше процедуры термин "синдромное декодирование" , сравним его с декодированием, которое производится с помощью так называемой "книги декодирования". Вообще говоря, принятыми векторами могут быть все q^n векторов длины n с компонентами из алфавита, содержащего q букв. Значит, книга декодирования должна содержать все эти векторы с указанием, какому из возможных передаваемых векторов он соответствует. Применяя в случае линейного кода синдромное декодирование, нам потребуется книга, ставящая в соответствие каждому из q^{n-k} синдромов смежный класс с соответствующим вектором-ошибкой. Ясно, что число страниц книги во втором случае меньше, чем в первом. Однако и q^{n-k} — тоже экспонента. Правда, например, в случае кода Хэмминга $n-k=\log_2(n+1)$, но в общем случае — "с экспонентой не шутят". Ниже будут изложены другие, более экономные, методы декодирования.

Некоторое весьма незначительное упрощение процесса синдромного декодирования можно достигнуть посредством так называемого "стандартного расположения". Оно состоит в следующем.

Выпишем все векторы кодового подпространства слева направо, начиная с нулевого вектора. Затем, помня, что смежный класс определяется любым своим элементом, расположим произвольный смежный класс под кодовым подпространством, начиная с вектора e минимального веса, так чтобы под кодовым вектором u находился вектор $v=u+e$ смежного класса. Вектор e назвется лидером смежного класса. Получим таблицу, верхняя строка которой есть кодовое подпространство, а остальные строки — смежные классы. Приняв вектор v, и вычислив его синдром S, определяем отвечающий ему смежный класс, а в нём находим вектор v. Непосредственно над ним в первой строке таблицы находится передававшийся вектор u.

Стандартная расстановка освобождает от операции вычитания лидера e смежного класса из принятого вектора v. Вычитание выполнено заранее тем, что принятый вектор v расположен под вектором u.

П р и м е р 4. 3.

Пусть линейный код над $GF(2)$ имеет параметры

$$n=5, \quad k=2, \quad n-k=3, \quad d=3.$$

Его порождающая матрица есть

$$G=\left[\begin{array}{ccccccc} 1 & 0 & : & 1 & 1 & 1 \\ 0 & 1 & : & 0 & 1 & 1 \end{array}\right]$$

и проверочная матрица есть

$$H = \left[\begin{array}{ccc:ccc} 1 & 0 & : & 1 & 0 & 0 \\ 1 & 1 & : & 0 & 1 & 0 \\ 1 & 1 & : & 0 & 0 & 1 \end{array}\right].$$

Стандартное расположение имеет вид

$$\mathbf{A_G^0} = (\mathbf{00000}), (\mathbf{10111}), (\mathbf{01011}), (\mathbf{11100})$$

$$\begin{array}{llllll} A_G^1 = A_G^0 + (10000) = \{(10000) & (00111) & (11011) & (01100)\}, \\ A_G^2 = A_G^0 + (01000) = \{(01000) & (11111) & (00011) & (10100)\}, \\ A_G^3 = A_G^0 + (00100) = \{(00100) & (10011) & (01111) & (11000)\}, \\ A_G^4 = A_G^0 + (00010) = \{(00010) & (10101) & (01001) & (11110)\}, \\ A_G^5 = A_G^0 + (00001) = \{(00001) & (10110) & (01010) & (11101)\}, \\ A_G^6 = A_G^0 + (10001) = \{10001) & (00110) & (11010) & (01101)\}, \\ A_G^7 = A_G^0 + (10010) = \{(10010) & (00101) & (11001) & (01110)\}, \end{array}$$

где $\mathbf{A_G^0} - (5,2)-$код с порождающей матрицей G, $\mathbf{S_0}$ — отвечающий ему синдром; $A_G^i, (i = 1, 2, \ldots, 7)$ — смежные классы. Отвечающие им синдромы

$$S_1 = (111), S_2 = (011), S_3 = (100), S_4 = (010), S_5 = (001),$$

$S_6 = (110), S_7 = (101)$.

В этом примере каждый смежный класс A_G^i, $i = 1, 2, \ldots, 5$, в качестве своих лидеров имеет вектор веса 1, и потому код исправляет все одиночные ошибки. В смежных классах A_G^6 и A_G^7 имеется по два вектора веса 2, остальные — большего веса. Лидеров в этих классах нет. Таким образом, целые два смежных класса не используются для исправления ошибок.

Правда, в иных случаях при необходимости можно сформулировать условия, благодаря которым одному из векторов одинакового веса в смежном классе отдаётся предпочтение, и его выбирают лидером.

Всвязи с этим формулируется понятие совершенного кода.

Определение 4.6.4. *Линейный код называется совершенным, если все q^{n-k} смежных классов имеют своих лидеров, и ими являются все векторы веса t и менее.*

Определение 4.6.5. *Линейный код называется квазисовершенным, если все q^{n-k} смежных классов имеют своих лидеров, и ими являются все векторы веса t и менее, а также некоторые векторы веса $t+1$.*

Это означает, что в первом случае код исправляет все ошибки кратности t и менее и не исправляет ошибки никакой иной кратности, а во втором случае код исправляет все те же самые ошибки, *некоторые* ошибки кратности $t+1$ и не исправляет ошибки никакой иной кратности

Совершенными линейными кодами являются:

– Двоичный $(2^m-1, 2^m-m-1, 3)$-код Хэмминга,

– Двоичный (23, 12, 7)-код Голея,

– Троичный (11, 6, 5)-код Голея

– Код, состоящий из одного кодового вектора. Он исправляет все ошибки всех кратностей, не превосходящих длины вектора.

– Всё пространство V. Этот код исправляет все ошибки кратности 0 и не исправляет больше никаких ошибок.

– Другим примером является $(n=2l+1, 1, 2l+1)$-код. Этот код исправляет все ошибки кратности l и менее и не исправляет ошибки никакой иной кратности.

Из теоремы 4.6.2 немедленно следует верхняя граница скорости передачи k/n линейного кода, исправляющего все ошибки кратности t и менее:

$$\sum_{i=0}^{t} C_n^i (q-1)^i \le q^{n-k}. \tag{4.6.11}$$

Отсюда следует, что

$$\frac{k}{n} \le 1 - \frac{1}{n}\log_q \sum_{i=0}^{t} C_n^i (q-1)^i. \tag{4.6.12}$$

Легко сосчитать, что в перечисленных выше случаях совершенных кодов в соотношении (4.6.12) достигается равенство.

Неравенство (4.6.12) есть не что иное, как граница Хэмминга. Правда, выведена она в предположении линейности кода. Однако ее легко вывести и на случай нелинейного кода. Действительно, вспомним понятие шара, которое было определено

в первой главе. Пусть u есть кодовый вектор длины n над алфавитом из q элементов. Назовем этот вектор центром шара, а объемом шара — его центр и совокупность всех $\sum_{i=0}^{t} C_n^i (q-1)^i$ (некодовых!) векторов, находящихся на расстоянии t и менее от центра. Пусть код содержит M векторов. Для правильного декодирования все шары, образованные "вокруг" каждого из кодовых векторов, не должны пересекаться. Поэтому с необходимостью должно выполняться неравенство

$$M \sum_{i=0}^{t} C_n^i (q-1)^i \le q^n. \tag{4.6.13}$$

Отсюда

$$\log_q M + \log_q \sum_{i=0}^{t} C_n^i (q-1)^i \le n.$$

Положив $[\log_q M] = k$, получим

$$\frac{k}{n} \le 1 - \frac{1}{n} \log_q \sum_{i=0}^{t} C_n^i (q-1)^i,$$

что и требовалось.

Предыдущие рассуждения относились к случаю нечетного расстояния $d = 2t + 1$. Граница Хэмминга на случай четного расстояния $d = 2t+2$ будет обсуждаться в следующем разделе.

4.7. Операции над кодами

В практике кодирования возникают разнообразные условия, когда нужно и можно простыми средствами незначительно изменить параметры и свойства кода. Такими средствами являются:

1. Расширение двоичного (n, k, d)**-кода.** Оно состоит в добавлении так называемой проверки на четность. Именно, добавим в конце каждого кодового вектора символ 0, если вес кодового слова четный, и 1 — в противном случае. Если все векторы исходного кода имеют четный вес, то операция расширения не имеет смысла, так как не имеет смысла добавление нуля к каждому вектору. Если минимальный вес

исходного кода нечетный, то минимальный вес полученного кода увеличивается на единицу, и все векторы кода становятся векторами четного веса. Полученный код имеет параметры $n^{'} = n + 1$, $k^{'} = k$, $d^{'} = d + 1$.

К проверчоной матрице исходного кода добавляется один столбец, и одна строка, так как на единицу увеличивается и длина кода и число проверочных символов. Одним из вариантов добавочной строки является сплошь единичная строка. Другой вариант добавочной строки следующий: если вес столбца исходной матрицы четный, то к нему приписывается 1, в противном случае — 0. В любом случае в новой проверочной матрице появится столбец $(00\ldots01)^T$.

П р и м е р 4. 4.

Матрица (4.4.7) — это проверочная матрица $(7, 4, 3)$-кода Хэмминга. Операция расширения с добавлением сплошь единичной строки приведет ее к виду

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \qquad (4.7.14)$$

Матрица (4.7.14) – это проверочная матрица (8, 4, 4)-кода, полученного операцией расширения. Действительно, нули последнего столбца, добавленные к первым трем строкам, никак не влияют на ортогональность векторов кода. Последняя строка ортогональна любым векторам четного веса.

Код исправляет любые одиночные ошибки и обнаруживает все двойные. Действительно, в случае одиночной ошибки синдром равен столбцу с единицей в разряде, отвечающем последней строке. Ошибка исправляется. В случае двойной ошибки ненулевой синдром будет содержать в этом разряде нуль. Ошибка обнаруживается

Второй вариант добавочной строки приводит к матрице

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}. \qquad (4.7.15)$$

Фактически, матрица (4.7.15) получена эквивалентным преобразованием из матрицы (4.7.14): сумма трех ее первых строк прибавлена к четвертой.

Все столбцы матрицы (4.7.15) нечетного веса. В случае одиночной ошибки синдром равен столбцу, отвечающему разряду, в котором произошла ошибка. Ошибка исправляется. В случае двойной ошибки синдром равен сумме столбцов, отвечающих разрядам, в которых произошла ошибка. Сумма двух столбцов нечетного веса имеет четный вес. Ошибка обнаруживается.

Читатель без труда проверит линейную независимость строк матриц (4.7.14) и (4.7.15).

В общем случае рассуждения о корректирующей способности расширенного кода и ортогональности его векторов соответствующим проверочным матрицам аналогичны проведенным выше.

Преобразование порождающей матрицы при расширении кода — тривиально, так как ее строки являются векторами кода, подвергающегося расширению: к каждой строке приписывается сумма ее компонент по модулю два.

2. Выкалывание. Эта операция — обратная расширению и состоит в удалении одного проверочного символа. Длина n кода уменьшается на единицу, размерность k сохраняется. Расстояние d должно уменьшиться на единицу, ибо в противном случае удаленный символ был бесполезным. Действительно, будучи проверочным, он для того и предназначался, чтобы создать расстояние, ради которого и вводится избыточность.

3. Выбрасывание. Удаление из двоичного $(n,\, k,\,)$−кода всех векторов нечетного веса. Все векторы четного веса образуют подгруппу порядка 2^{k-1}. Векторы нечетного веса образуют смежный класс. Длина n кода сохраняется, размерность k уменьшается на единицу, и если кодовое расстояние d было нечетным, то новое расстояние $d^{'} > d$. Действительно, минимальное расстояние — это минимальный вес, и если минимальный вес был нечетным, то все четные веса кода больше него.

4. Пополнение. Добавление к двоичному $(n,\, k,\, d)$-коду сплошь единичного вектора, если он еще не принадлежит коду. Новому $(n,\, k{+}1,\, d^{'})$−коду вместе с вектором $v = (a_1, a_2, \ldots, a_n,)$ будет принадлежать вектор $v^{'} = (a_1{+}1, a_2{+}1, \ldots, a_n{+}1)$. Если максимальное расстояние исходного кода было D, то для нового расстояния $d^{'}$ выполняется соотношение $d^{'} = \min\{d,\, n - D\}$.

5. Удлинение. Эта операция состоит в последовательном выполнении двух операций — пополнения и расширения. При пополнении увеличивается размерность кода, а при расширении увеличивается длина и число проверочных символов.

6. Укорочение. Фиксируем произвольный столбец $(n,\, k,\, d)$-кода и выбираем только те векторы, которые в данном столб-

це содержат 0. Это множество векторов образует подпространство. Его порядок равен q^{k-1}. Отбросив в этом подпространстве нулевой столбец, получим укороченный $(n-1,\, k-1,\, d)-$код.

Познакомившись с операцией выкалывания, можно получить границу Хэмминга для случая четного кодового расстояния.

После выкалывания параметры кода будут $n-1$, k, $d-1 = 2t+2-1 = 2t+1$. Выражение (4.6.11) превратится в такое:

$$\sum_{i=0}^{t} C_{n-1}^{i}(q-1)^i \le q^{n-k-1}, \tag{4.7.16}$$

откуда искомая граница выглядит так

$$\frac{k}{n} \le 1 - \frac{1}{n}\log_q \sum_{i=0}^{t} C_{n-1}^{i}(q-1)^i - \frac{1}{n}. \tag{4.7.17}$$

Аналогичными рассуждениями можно получить границу Хэмминга для четного расстояния нелинейного кода. Только, вместо удаления проверочного символа, придется говорить об удалении координаты во всех кодовых векторах при заведомом сокращении кодового расстояния на единицу.

4.8. Мажоритарное декодирование линейного кода

Начнём с простого примера.

П р и м е р 4. 5.

Пусть проверочной матрицей двоичного линейного $(6, 3)$-кода будет

$$H = \left[\begin{array}{ccc:ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right], \tag{4.8.18}$$

и пусть

$$u = (a_1, a_2, a_3, a_4, a_5, a_6),$$

произвольный кодовый вектор, где первые три символа информационные. Из трёх скалярных произведений вектора u на каждую строку матрицы H получим

$$a_2 + a_3 + a_4 = 0,\ a_1 + a_2 + a_5 = 0,\ a_1 + a_3 + a_6 = 0. \tag{4.8.19}$$

Рассмотрим информационный символ a_1. Из второго и третьего равенства в (4.8.19) после добавления тривиального равенства $a_1 = a_1$ получается система

$$\begin{array}{c} a_1 = a_2 + a_5, \\ a_1 = a_3 + a_6, \\ a_1 = a_1. \end{array} \tag{4.8.20}$$

В реальной передаче на приёмном конце имеют дело с вектором $v = (a'_1, a'_2, a'_3, a'_4, a'_5, a'_6)$, где для любого $i = 1, 2, \ldots, 6.$ символ a'_i может отличаться от a_i.

Система (4.8.20) превратится в такую

$$\begin{array}{c} a'_1 = a'_2 + a'_5, \\ a'_1 = a'_3 + a'_6, \\ a'_1 = a'_1. \end{array} \tag{4.8.21}$$

В первых двух уравнениях системы (4.8.21) символ a'_1 вычисляется посредством остальных символов принятого вектора v. В третьем уравнении его значение берётся непосредственно из принятого вектора v.

Если ошибок не было, то системы (4.8.21) и (4.8.20) совпадают, и левые части всех трёх уравнений обеих систем совпадают с a_1. Положим, произошла одна ошибка. Если в векторе v искажен символ a'_1, то все символы правых частей первых двух уравнений системы (4.8.21) совпадают с нештрихованными символами, и в них $a'_1 = a_1$, а потому символ a_1 по правилу большинства декодируется верно. Если же символ a_1 передан верно, то в правых частях первых двух уравнений системы (4.8.21) искажен только один символ, и потому символ a'_1 не совпадает с a_1 только в одном уравнении. И снова по правилу большинства символ a_1 декодируется верно.

Системы, аналогичные системе (4.8.20), имеют место и для информационных символов a_2, a_3 :

$$\begin{array}{c} a_2 = a_3 + a_4, \\ a_2 = a_1 + a_5, \\ a_2 = a_2. \end{array}$$

$$\begin{array}{c} a_3 = a_2 + a_4, \\ a_3 = a_1 + a_6, \\ a_3 = a_3. \end{array}$$

Таким образом, все три информационных символа по правилу большинства декодируются верно, а проверочные символы нас не интересуют. Для декодирования в этом примере тре-

буется всего шесть двуместных операций суммирования и три мажоритарных элемента.

В общем случае в канонической форме проверочной матрицы линейного кода, исправляющего любую одиночную ошибку мажоритарным образом, подматрица $P^T_{(n-k)\times k}$ состоит из всех C^2_{n-k} столбцов высоты $n-k$, веса 2, и потому параметры кода удовлетворяют равенству $k = C^2_{n-k}$.

Структура каждой соответствующей системы уравнений такова, что данный информационный символ a_i входит во все уравнения системы, а каждый из остальных символов $a_j, j \neq i$ входит только в одно из уравнений системы, чем и обеспечивается верное декодирование символа a_i. Каждое из уравнений системы называется "проверкой" , а вся система называется системой "разделённых проверок".

Легко заметить, что матрица (4.8.18) получается из проверочной матрицы (4.4.7) удалением столбца веса 3, т.е. укорочением (7, 4)-кода Хэмминга. Вообще говоря, любой код рассмотренного класса получается из кода Хэмминга длины $n = 2^m - 1$ укорочением последнего путём удаления из проверочной матрицы всех столбцов веса более двух. Кстати, любое дальнейшее укорочение кода данного класса, т.е. удаление из проверочной матрицы и некоторых столбцов веса 2 сохраняет для оставшихся информационных символов корректирующую способность декодирования по правилу большинства.

Важное значение имеет вопрос, есть ли различие в корректирующей способности кода при мажоритарном декодировании и при декодировании по кодовому расстоянию. Иначе говоря, можно ли посредством мажоритарного декодирования двоичного линейного кода исправить столько же ошибок. сколько и посредством минимального расстояния этого кода. В иных выражениях этот вопрос звучит так: "Реализуется ли кодовое расстояние при мажоритарном декодировании?"

Нетрудно проверить, что для рассмотренного класса кодов любые два столбца проверочной матрицы линейно независимы, и всегда найдутся три линейно зависимых столбца. Поэтому минимальное расстояние этих кодов в точности равно трём.

Существуют коды с мажоритарным декодированием, реализующим значительно бóльшие кодовые расстояния.

Такими кодами, например, являются коды, двойственные (ортогональные) кодам Хэмминга.

П р и м е р 4. 6.

Приведём проверочную матрицу двоичного линейного (15,

4)-кода

$$H = \left[\begin{array}{cccccccccccccccc}
0 & 0 & 1 & 1 & : & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & : & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & : & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & : & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & : & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & : & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & : & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right].$$

Произвольный столбец матрицы $P^T_{11\times 4}$ содержит 4 нуля и 7 единиц. Не ограничивая общности, рассмотрим четвёртый столбец этой матрицы. Он отвечает информационному символу a_4 кодового вектора. Умножим скалярно кодовый вектор

$$v = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{13}, a_{14}, a_{15})$$

на каждую строку матрицы H. Получим:

$$\begin{array}{l}
1.\, a_3 + a_4 + a_5 = 0, \\
2.\, a_2 + a_4 + a_6 = 0, \\
3.\, a_2 + a_3 + a_7 = 0, \\
4.\, a_2 + a_3 + a_4 + a_8 = 0, \\
5.\, a_1 + a_4 + a_9 = 0, \\
6.\, a_1 + a_3 + a_{10} = 0, \\
7.\, a_1 + a_3 + a_4 + a_{11} = 0, \\
8.\, a_1 + a_2 + a_{12} = 0, \\
9.\, a_1 + a_2 + a_4 + a_{13} = 0, \\
10.\, a_1 + a_2 + a_3 + a_{14} = 0, \\
11.\, a_1 + a_2 + a_3 + a_4 + a_{15} = 0,
\end{array}$$

Сложим попарно равенства: 3 и 4, 6 и 7, 8 и 9, 10 и 11. Отвечающие им строки матрицы P^T таковы, что в каждой паре они совпадают, кроме символа в четвёртом столбце. Уравнения 1, 2 и 5 оставим неизменными. Получим относительно информационного символа a_4, (добавив тривиальное уравнение $a_4 = a_4$) следующие 8 уравнений:

$$\begin{array}{l}
a_4 = a_7 + a_8, \\
a_4 = a_{10} + a_{11}, \\
a_4 = a_{12} + a_{13}, \\
a_4 = a_{14} + a_{15}, \\
a_4 = a_3 + a_5, \\
a_4 = a_2 + a_6, \\
a_4 = a_1 + a_9, \\
a_4 = a_4.
\end{array} \qquad (4.8.22)$$

Система (4.8.22) представляет собой восемь разделённых проверок. Любые три ошибки искажают не более трёх значений символа a_4. Декодирование по правилу большинства даст верный результат.

Любые четыре ошибки искажают значения символа a_4 не более чем в четырёх равенствах (4.8.22). В этом случае декодирование по правилу большинства не даст результата. Факт ошибки будет установлен. Но и не более. Такое обстоятельство называется отказом от декодирования.

В общем случае коды, двойственные кодам Хэмминга, имеют параметры $n = 2^m - 1, k = m$.

Теорема 4.8.1. *Код, двойственный коду Хэмминга, допускает построение 2^{m-1} разделённых проверок для каждого информационного символа, т.е. обеспечивает исправление любых $2^{m-2} - 1$ и менее независимых ошибок.*

Доказательство. Каноническая форма проверочной матрицы имеет вид

$$H = [P^T_{(2^m-1-m)\times m} I_{(2^m-1-m)\times(2^m-1-m)}].$$

В матрице $P^T_{(2^m-1-m)\times m}$ отсутствуют все m строк с одной единицей и нулевая строка. Произвольный ее столбец содержит $2^{m-1}-m$ нулей и $2^{m-1}-1$ единиц. Пронумеруем столбцы матрицы $P^T_{(2^m-1-m)\times m}$ в произвольном порядке числами $1, 2, \ldots, m$. Фиксируем столбец с номером i_0 и выберем любую строку матрицы с нулём в этом столбце. Для этой строки найдётся одна и только одна парная ей строка, которая совпадает с данной всюду кроме компоненты в столбце с номером i_0. Таких пар строк имеется в точности $2^{m-1} - m$. В приведённом выше примере такими парами являются 3,4; 6,7; 8,9; 10,11. Сложим строки в каждой паре. Получим $2^{m-1} - m$ строк-сумм, каждая из которых содержит одну единицу в столбце с номером i_0 и еще две единицы, которые расположены на главной диагонали матрицы $I_{(2^m-1-m)\times(2^m-1-m)}$. Каждой такой строке-сумме отвечает проверочная сумма (далее — проверка)

$$a_{i_0} = a_{j_1} + a_{j_2} \tag{4.8.23}$$

компонент кодового вектора.

При этом все индексы j_1, j_2 во всех проверках различны, так как никакие две суммы не содержат одинаковых диагональных элементов, и диагональные элементы внутри проверки также различны. В столбце с номером i_0 имеется еще $m-1$ единиц, не вошедших ни в одну проверку(4.8.23), так как они принадлежат строкам веса 2 матрицы $P^T_{(2^m-1-m)\times m}$. Парными строками для них были бы строки с одной единицей, но, как отмечено выше, они в этой матрице отсутствуют. Зато каждая из этих строк сама по себе доставляет проверку

$$a_{i_0} = a_{i_l} + a_{j_z};\ i_l = 1, 2, \ldots m,\ i_l \neq i_0. \tag{4.8.24}$$

В проверках (4.8.24) все индексы j_z различны и не совпадают ни с одним из индексов остальных диагональных элементов матрицы $I_{(2^m-1-m)\times(2^m-1-m)}$.

Таким образом, для информационного символа a_{i_0} построены $2^{m-1}-1$ проверок, в левой части которых находится символ $a_{i_0,}$ а в правых частях находятся непересекающиеся пары всех остальных 2^m-2 символов. Присовокупив к этим проверкам еще одну тривиальную проверку $a_{i_0} = a_{i_0}$, получим систему 2^{m-1} разделённых проверок, посредством которых можно исправить любые $2^{m-2}-1$ и менее ошибок и обнаружить любые 2^{m-1} ошибок. Тем, что информационный символ a_{i_0} выбран произвольно, завершается доказательство.

Теорема 4.8.2. *Кодовое расстояние кода, двойственного* $(2^m-1, 2^m-1-m)$*-коду Хэмминга, равно в точности* 2^{m-1}.

Иначе говоря, при мажоритарном декодировании данного кода реализуется его кодовое расстояние. Читатель может доказать теорему элементарными комбинаторными средствами. Другое доказательство теоремы, основанное на свойствах циклических кодов, будет дано ниже.

Важнейшим классом кодов, допускающих мажоритарное декодирование являются коды Рида—Маллера.

4.9. Коды Рида—Маллера

Этот замечательный класс кодов получил освещение в весьма серьезных руководствах по теории кодирования и благодаря своему строению, до сих пор служит обильным источником для новых постановок задач. Упомянутые руководства давно стали

библиографической редкостью, и к тому же для первого чтения принятое в них изложение слишком уплотнено и насыщено материалом.

Здесь изложение кодов Рида—Маллера (РМ-коды) следует одной из ранних традиций, и потому есть надежда, что читатель без труда справится с предлагаемым текстом.

Введем в обиход векторное умножение векторов

$$b = (b_1, b_2, \ldots, b_n) \qquad \text{и} \qquad b' = (b_1', b_2', \ldots, b_n')$$

над $GF(2)$:

$$bb' = (b_1 b_1', b_2 b_2', \ldots, b_n b_n'), \tag{4.9.25}$$

где $b_i b_i' = 1$ тогда и только тогда, когда $b_i = b_i' = 1$. На самом деле, эта операция есть поразрядная конъюнкция двух векторов. Такая интерпретация позволяет вести описание РМ-кодов на языке булевой алгебры.

РМ-коды строятся следующим образом. Расположим в виде $(m+1) \times 2^m-$матрицы в лексикографическом порядке слева направо всевозможные 2^m двоичных столбцов длины $m+1$, один символ которых, находящийся в верхней строке, всегда равен 1. Строки этой матрицы обозначим через $v_0, v_m, \ldots, v_1$.

Далее, пользуясь правилом (4.9.25), образуем все возможные произведения векторов v_i, $(i = 0, 1, \ldots, m)$ по одному, по два, ..., по m. Всего получится 2^m различных векторов длины 2^m, которые образуют $(2^m \times 2^m)-$матрицу M_m.

В примере 4.7 на таблице (4.9.28) изображен процесс образования матрицы M_m на случай $m = 5$.

Процесс построения $(2^m \times 2^m)-$матрицы M_m является интерпретацией первоначального метода ее описания посредством булевой алгебры. Именно, 2^m столбцов, заключенных в строках $v_m, \ldots, v_1$, представляют собой наборы значений булевых переменных $v_m, \ldots, v_1$, а все остальные строки представляют собой значения конъюнкций

$$v_m v_{m-1},\ v_m v_{m-2},\ \ldots,\ v_2 v_1,\ \ldots,\ v_m v_{m-1} \cdots v_1$$

рангов[1] $2, 3, \ldots m$. Конъюнкции ранга 1 есть сами переменные, а строка v_0 состоит из одних единиц.

Правило (4.9.25), таким образом, означает, что поразрядная конъюнкция истинна (равна 1) тогда и только тогда, когда истинны (равны 1) все элементы данного разряда.

[1] Рангом конъюнкции называется число входящих в нее конъюнктивных членов

Теорема 4.9.1. *Строки матрицы M_m линейно независимы.*

Доказательство. Переставим строки матрицы, располагая их сверху вниз в следующем порядке:

— Сначала всевозможные различные произведения векторов, не содержащие множителем вектор v_m,

— Затем все остальные.

Получим

$$M_m = \begin{bmatrix} M_{m-1} & M_{m-1} \\ 0 & M_{m-1} \end{bmatrix} \tag{4.9.26}$$

Предположив, что строки матрицы M_{m-1} линейно независимы, читатель легко докажет, что таковы и строки матрицы M_m. Тем самым будет произведен индукционный шаг. Невырожденность же матрицы

$$M_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

очевидна. На случай $m = 5$ матрица (4.9.26) изображена в примере 4.8.

Определение 4.9.2. *Кодом Рида — Маллера r-го порядка называется код длины $n = 2^m$, порождающая матрица которого образована всеми различными произведениями векторов v_i $(i = 0, 1, \ldots, m)$, состоящими из r и менее сомножителей.*

Из этого определения следует, что РМ-код r-го порядка — это линейный код, число информационных символов которого выражается соотношением

$$k = \sum_{j=0}^{r} C_m^j \tag{4.9.27}$$

Порождающую матрицу РМ-кода r-го порядка будем обозначать символом M_m^r. Легко видеть, что РМ-код порядка $r = m$ — это все пространство двоичных векторов длины $n = 2^m$, и потому его рассмотрение не представляет интереса.

П р и м е р 4. 7.
Процесс формирования матрицы M_m на случай $m = 5$.

$$
\begin{array}{rl}
v_0 & 11111111111111111111111111111111 \\
v_5 & 00000000000000001111111111111111 \\
v_4 & 00000000111111110000000011111111 \\
v_3 & 00001111000011110000111100001111 \\
v_2 & 00110011001100110011001100110011 \\
v_1 & 01010101010101010101010101010101 \\
v_5v_4 & 00000000000000000000000011111111 \\
v_5v_3 & 00000000000000000000111100001111 \\
v_5v_2 & 00000000000000000011001100110011 \\
v_5v_1 & 00000000000000000101010101010101 \\
v_4v_3 & 00000000000011110000000000001111 \\
v_4v_2 & 00000000001100110000000000110011 \\
v_4v_1 & 00000000010101010000000001010101 \\
v_3v_2 & 00000011000000110000001100000011 \\
v_3v_1 & 00000101000001010000010100000101 \\
v_2v_1 & 00010001000100010001000100010001 \\
v_5v_4v_3 & 00000000000000000000000000001111 \\
v_5v_4v_2 & 00000000000000000000000000110011 \\
v_5v_4v_1 & 00000000000000000000000001010101 \\
v_5v_3v_2 & 00000000000000000000001100000011 \\
v_5v_3v_1 & 00000000000000000000010100000101 \\
v_5v_2v_1 & 00000000000000000001000100010001 \\
v_4v_3v_2 & 00000000000000110000000000000011 \\
v_4v_3v_1 & 00000000000001010000000000000101 \\
v_4v_2v_1 & 00000000000100010000000000010001 \\
v_3v_2v_1 & 00000010000000100000001000000001 \\
v_5v_4v_3v_2 & 00000000000000000000000000000011 \\
v_5v_4v_3v_1 & 00000000000000000000000000000101 \\
v_5v_4v_2v_1 & 00000000000000000000000000010001 \\
v_5v_3v_2v_1 & 00000000000000000000000100000001 \\
v_4v_3v_2v_1 & 00000000000000010000000000000001 \\
v_5v_4v_3v_2v_1 & 00000000000000000000000000000001
\end{array}
\tag{4.9.28}
$$

П р и м е р 4. 8.
Матрица M_5 в виде (4.9.26) имеет вид

$$M_5 = \left[\begin{array}{cc}
1111111111111111 & 1111111111111111 \\
0000000011111111 & 0000000011111111 \\
0000111100001111 & 0000111100001111 \\
0011001100110011 & 0011001100110011 \\
0101010101010101 & 0101010101010101 \\
0000000000001111 & 0000000000001111 \\
0000000000110011 & 0000000000110011 \\
0000000001010101 & 0000000001010101 \\
0000001100000011 & 0000001100000011 \\
0000010100000101 & 0000010100000101 \\
0001000100010001 & 0001000100010001 \\
0000000000000011 & 0000000000000011 \\
0000000000000101 & 0000000000000101 \\
0000000000010001 & 0000000000010001 \\
0000000100000001 & 0000000100000001 \\
0000000000000001 & 0000000000000001 \\
\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\
0000000000000000 & 1111111111111111 \\
0000000000000000 & 0000000011111111 \\
0000000000000000 & 0000111100001111 \\
0000000000000000 & 0011001100110011 \\
0000000000000000 & 0101010101010101 \\
0000000000000000 & 0000000000001111 \\
0000000000000000 & 0000000000110011 \\
0000000000000000 & 0000000001010101 \\
0000000000000000 & 0000001100000011 \\
0000000000000000 & 0000010100000101 \\
0000000000000000 & 0001000100010001 \\
0000000000000000 & 0000000000000011 \\
0000000000000000 & 0000000000000101 \\
0000000000000000 & 0000000000010001 \\
0000000000000000 & 0000000100000001 \\
0000000000000000 & 0000000000000001
\end{array}\right]$$

В строке-произведении $v_{i_1}v_{i_2}\cdots v_{i_l}$ l сомножителей имеется в точности 2^{m-l} единиц, так как в любых l векторах $v_{i_1}, v_{i_2}, \ldots, v_{i_l}$ имеется именно такое число компонент, которые во всех l векторах содержат единицы. Единственный вектор нечетного веса в матрице M_m, согласно правилу (4.9.25), это вектор $v_0v_mv_{m-1}\ldots v_1$, так как в одном наборе векторов v_0, v_m, v_{m-1}, $\ldots$, v_1, и только в нем существует в точности $2^{m-m} = 1$ компононта, в которой все эти векторы равны 1. Но этот вектор принадлежит только коду порядка $r = m$. Таким образом, в порождающей матри-

це M_m^r нетривиального РМ-кода ($r < m$) все векторы четного веса. Применяя аналогичные рассуждения, получим, что все векторы в любом РМ-коде порядка $r < m$ имеют четный вес.

Произведение двух базисных векторов двух РМ-кодов порядков r и $m - r - 1$ содержит не более чем $m - 1$ различных сомножителей v_i. Отсюда следует, что скалярное произведение двух этих векторов содержит заведомо четное число единиц, а потому равно нулю. Отсюда следует, что РМ-код порядка $r < m$ является нулевым пространством для РМ-кода порядка $m-r-1$. Действительно, размерность РМ-код порядка $m-r-1$ есть

$$\sum_{j=0}^{m-r-1} C_m^j = \sum_{j=r+1}^{m} C_m^j. \tag{4.9.29}$$

Сумма размерностей (4.9.27) и (4.9.29) в точности равна длине $n = 2^m$ кода, т.е. размерности пространства $E_2^{(n)}$:

$$\sum_{j=0}^{r} C_m^j + \sum_{j=r+1}^{m} C_m^j = 2^m.$$

Рассмотрим, порождающую матрицу РМ-кода первого порядка. Нулевым подпространством строк этой матрицы, согласно предыдущим рассуждениям является РМ-код порядка $r = m-2$. Если из этой порождающей матрицы удалить строку v_0 и образовавшийся нулевой столбец, то получим матрицу, в которой все $2^m - 1$ ненулевых столбцов расположены в лексикографическом порядке. Не придавая значения порядку следования столбцов (ибо, как мы знаем, он не влияет на корректирующую способность кода), мы узнаем в этой матрице проверочную матрицу кода Хэмминга. Например, для случая $m = 3$ в этой матрице мы узнаем матрицу (0.4.11) для кода Хэмминга при $m = 3$. Так же будет и в общем случае при любом m. Это означает только, что РМ-код порядка $r = m - 2$ есть расширенный код Хемминга, а порождающая матрица РМ-кода первого порядка есть проверочная матрица расширенного кода Хэмминга. Это означает также, что выколотый РМ-код порядка $r = m - 2$ есть код Хэмминга. Это в свою очередь значит, что код Хэмминга, как и все РМ-коды, о чем подробно изложено ниже, допускает мажоритарное декодирование. Однако в связи с простейшим способом декодирования кода Хэминга,

предложенным в разделе 0.3, здесь выступают заботы о выборе оптимального метода декодирования. Разумеется, важнейшим критерием для выбора метода декодирования является его сложность. Об этом подробно будет сказано в конце главы.

Здесь же обратим внимание на один частный случай, когда $m = 3$. Так как $3 - 2 = 1$, то при $m = 3$ РМ-код $(m-2)-$ го порядка есть также и РМ-код первого порядка. Иначе говоря, расширенный (8, 4)-код Хэмминга (РМ-код первого и одновременно $(m-2)-$ го порядка) самоортогонален (т.е. самодвойственный). В этом, кстати, легко убедиться и непосредственно: числа информационных и проверочных символов совпадают, и скалярные произведения строк в каждой из матриц (4.7.14) и (4.7.15) равны нулю.

Теорема 4.9.3. *Минимальное кодовое расстояние $d(m, r)$ РМ-кода длины 2^m, порядка r выражается равенством*

$$d(m, r) = 2^{m-r}. \tag{4.9.30}$$

Д о к а з а т е л ь с т в о. Предположим, что равенство (4.9.30) справедливо для некоторого $m-1$ при всех $r \leq m-1$, и покажем, что оно справедливо для m при всех $r \leq m$.

Расположим строки матрицы M_m^r в следующем порядке:

— Сначала все $\Sigma_{j=0}^{r} C_{m-1}^{j}$ различных произведений векторов $v_0, v_{m-1}, \ldots, v_1$ по одному, по два,..., по r (вектор v_m не входит ни в одно из них; вектор v_0 есть единственный базисный вектор кода порядка $r = 0$, а его вхождение или невхождение в какое-нибудь произведение строк никак не проявляется: умножение на 1 не вносит в произведение ничего нового: "истинный конъюнктивный член может быть отброшен".)

— Затем $\Sigma_{j=0}^{r} C_{m-1}^{j}$ различных произведений векторов по одному, по два,..., по r, обязательно содержащих сомножитель v_m. После этого матрица M_m^r примет вид

$$M_m^r = \begin{bmatrix} M_{m-1}^r & M_{m-1}^r \\ 0 & M_{m-1}^{r-1} \end{bmatrix}. \tag{4.9.31}$$

Легко видеть, что при $r = m$ совпадают как матрицы M_m и M_m^m, так и способы их построения; надо только принять во внимание что при $r = m$ имеет место $C_{m-1}^r = 0$, и при этом условии $M_{m-1}^r = M_{m-1}^{m-1}$. (Продолжение доказательства на стр. 138).

П р и м е р 4. 9.

Процесс формирования РМ-кода на случай $m = 5$ и $r = 3$.

$$
\begin{array}{rl}
v_0 & 11111111111111111111111111111111 \\
v_4 & 00000000111111110000000011111111 \\
v_3 & 00001111000011110000111100001111 \\
v_2 & 00110011001100110011001100110011 \\
v_1 & 01010101010101010101010101010101 \\
v_4v_3 & 00000000000011110000000000001111 \\
v_4v_2 & 00000000001100110000000000110011 \\
v_4v_1 & 00000000010101010000000001010101 \\
v_3v_2 & 00000011000000110000001100000011 \\
v_3v_1 & 00000101000001010000010100000101 \\
v_2v_1 & 00010001000100010001000100010001 \\
v_4v_3v_2 & 00000000000000110000000000000011 \\
v_4v_3v_1 & 00000000000001010000000000000101 \\
v_4v_2v_1 & 00000000000100010000000000010001 \\
v_3v_2v_1 & 00000001000000010000000100000001 \\
 & ..... \\
v_5 & 00000000000000001111111111111111 \\
v_5v_4 & 00000000000000000000000011111111 \\
v_5v_3 & 00000000000000000000111100001111 \\
v_5v_2 & 00000000000000000011001100110011 \\
v_5v_1 & 00000000000000001010101010101010 \\
v_5v_4v_3 & 00000000000000000000000000001111 \\
v_5v_4v_2 & 00000000000000000000000000110011 \\
v_5v_4v_1 & 00000000000000000000000001010101 \\
v_5v_3v_2 & 00000000000000000000001100000011 \\
v_5v_3v_1 & 00000000000000000000010100000101 \\
v_5v_2v_1 & 00000000000000000001000100010001
\end{array}
\tag{4.9.32}
$$

П р и м е р 4. 10.
Матрица M_5^3 в виде (4.9.31)

$$M_5^3 = \begin{bmatrix} 1111111111111111 & 1111111111111111 \\ 0000000011111111 & 0000000011111111 \\ 0000111100001111 & 0000111100001111 \\ 0011001100110011 & 0011001100110011 \\ 0101010101010101 & 0101010101010101 \\ 0000000000001111 & 0000000000001111 \\ 0000000000110011 & 0000000000110011 \\ 0000000001010101 & 0000000001010101 \\ 0000001100000011 & 0000001100000011 \\ 0000010100000101 & 0000010100000101 \\ 0001000100010001 & 0001000100010001 \\ 0000000000000011 & 0000000000000011 \\ 0000000000000101 & 0000000000000101 \\ 0000000000010001 & 0000000000010001 \\ 0000000100000001 & 0000000100000001 \\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots \\ 0000000000000000 & 1111111111111111 \\ 0000000000000000 & 0000000011111111 \\ 0000000000000000 & 0000111100001111 \\ 0000000000000000 & 0011001100110011 \\ 0000000000000000 & 0101010101010101 \\ 0000000000000000 & 0000000000001111 \\ 0000000000000000 & 0000000000110011 \\ 0000000000000000 & 0000000001010101 \\ 0000000000000000 & 0000001100000011 \\ 0000000000000000 & 0000010100000101 \\ 0000000000000000 & 0001000100010001 \end{bmatrix} \tag{4.9.33}$$

Из определения РМ-кода следует, что РМ-код порядка r_1 целиком содержится в РМ-коде порядка $r_2 > r_1$. Далее, для каждого не сплошь нулевого кодового вектора, являющегося линейной комбинацией строк матрицы (4.9.31), имеет место одно из двух: либо и правый и левый его отрезки длины 2^{m-1} не сплошь нулевые, и тогда по предположению индукции их веса не меньше, чем 2^{m-r-1}, а значит, вес всего вектора не меньше, чем $2 \cdot 2^{m-r-1} = 2^{m-r}$, либо один из этих отрезков сплошь нулевой, и, значит, другой отрезок с необходимостью принадлежит коду длины 2^{m-1} и порядка $r-1$, а тогда по предположению индукции минимальный вес таких отрезков равен в точности $2^{m-1-(r-1)} = 2^{m-r}$. Вспомним теперь, что в линейном коде значения попарных расстояний исчерпываются весами кодовых векторов. Остается заметить, что при $m = r = 1$ равенство $d(1,1) = 1$ очевидно, чем и завершается доказательство.

4.10. Кодирование кода Рида—Маллера

Кодирование для РМ-кода происходит обычным образом (см. (4.2.1)). Подлежащий кодированию информационный вектор $u = (u_1, u_2, \ldots, u_k)$ умножается на порождающую матрицу M_m^r. Результатом процесса кодирования является кодовый вектор $b = b^r = (b_1^r, b_2^r, \ldots, b_n^r) = uM_m^r$. Иначе говоря, каждая строка матрицы M_m^r умножается скалярно на соответствующий символ вектора u, а затем эти произведения складываются поразрядно по модулю два. Помня, что строки матрицы M_m^r обозначены символами [2]

$$v_0,\ v_m,\ \ldots,v_1,\ v_m v_{m-1},\ v_m v_{m-2},\ \ldots,v_r v_{r-1}\cdots v_1,$$

обозначим соответствующие им компоненты вектора u символами

$$u_0,\ u_m,\ \ldots,u_1,\ u_{m,m-1},\ u_{m,m-2},\ \ldots,u_{r,r-1,\ldots,1}. \qquad (4.10.34)$$

Таким образом,

$$b = b^{(r)} = (b_1^{(r)}, b_2^{(r)}, \ldots, b_n^{(r)}) = u_0 v_0 + u_m v_m + \ldots + u_1 v_1 +$$

[2] Вообще говоря, нижние индексы могут располагаться в любом порядке. Расположение их в порядке убывания предпринято здесь только ради сохранения стиля

$$+u_{m,m-1}v_m v_{m-1} + \ldots + u_{21}v_2 v_1 + \ldots + u_{r,r-1,\ldots 1}v_r v_{r-1}\ldots v_1 \tag{4.10.35}$$

Символы (4.10.34) — информационные, и символ $u_{i_l i_{l-1}\ldots i_1}$, на который умножается вектор $v_{i_l}v_{i_{l-1}}\cdots v_{i_1}$, называется информационным символом $l-$го порядка.

Итак, вектор (4.10.35) представляет собой вектор РМ-кода r-го порядка, отвечающий информационному вектору

$$u = (u_0,\, u_m,\, \ldots, u_1,\, u_{m,m-1},\, u_{m,m-2},\, \ldots, u_{r,r-1,\ldots,1}.) \tag{4.10.36}$$

Читатель, знакомый с булевой алгеброй, заметил, что сумма в правой части равенства (4.10.35) есть в точности многочлен Жегалкина от переменных v_i. Информационные символы являются его коэффициентами. В соответствии с (4.10.36) все информационные символы порядка выше, чем r, *заведомо* равны нулю.

Следующие рассуждения имеют дело с вычислением сложности, т.е. числа операций, требуемых при кодировании.

4.11. Сложность кодирования кода Рида — Маллера

Легко видеть, что кодирование для РМ-кода реализуется только посредством операций суммирования, так как операция умножения в (4.10.35) есть операция умножения на константу 1 или 0, а это не требует никаких новых действий. Обращаясь к (4.9.31), видим, что в соответствии со строением матрицы M_m^r суммы (4.10.35) можно разбить на две. К первой, обозначим ее символом $b(\overline{v_m})$, отнесем те векторы-слагаемые, которые не содержат v_m в качестве сомножителя. Эти векторы получаются умножением на соответствующие информационные символы строк верхней полосы матрицы (4.9.31). Для получения вектора-суммы $b(\overline{v_m})$ достаточно просуммировать только правые отрезки длины 2^{m-1} входящих в него векторов-слагаемых. Тем самым будут вычислены и левые отрезки такой же длины, вследствие строения матрицы (4.9.31).

Ко второй сумме, обозначим ее символом $b(v_m)$, отнесем те векторы-слагаемые, которые содержат вектор v_m в качестве сомножителя. Левые отрезки длины 2^{m-1} этих векторов — сплошь нулевые, и потому суммировать следует опять-таки только правые отрезки вектора $b(v_m)$.

Пусть векторы $b(\overline{(v_m)})$ и $b(v_m)$ уже получены. Тогда для получения их суммы $b = b^r = b\overline{(v_m)} + b(v_m)$ требуется еще 2^{m-1} двуместных операций сложения их компонент.

Обозначим через Q_m^r число операций сложения для получения вектора (4.10.35). Тогда из (4.9.31) следует

$$Q_m^r = Q_{m-1}^r + Q_{m-1}^{r-1} + 2^{m-1}. \tag{4.11.37}$$

При этом надлежит помнить, что $Q_{m-1}^m = Q_{m-1}^{m=1}$, так как $M_{m-1}^m = M_{m-1}^{m-1}$,

Опираясь на рекуррентное соотношение (4.11.37), можно найти величину Q_m^r. Этому служит следующая

Теорема 4.11.1.

$$Q_m^r = r(2^m - 2^{r-1}) - \sum_{i=1}^{r-1} i2^{i-1} C_{m-i-1}^{m-r-1}. \tag{4.11.38}$$

Доказательство будем вести индукцией по m. Предположим, что теорема верна для $m-1$ при всех $r \le m-1$. Тогда в силу (4.11.37) имеем

$$Q_m^r = r(2^{m-1} - 2^{r-1}) - \sum_{i=1}^{r-1} i2^{i-1} C_{m-i-2}^{m-r-2} + (r-1)(2^m - 2^{r-2}) -$$

$$- \sum_{i=1}^{r-2} i2^{i-1} C_{m-i-2}^{m-r-1} + 2^{m-1} = r(2^m - 2^{r-1}) - (r-1)2^{m-2} -$$

$$- \sum_{i=1}^{r-1} i2^{i-1} C_{m-i-2}^{m-r-2} - \sum_{i-1}^{r-2} i2^{i-1} C_{m-i-2}^{m-r-1} = r(2^m - 2^{r-1}) -$$

$$- \sum_{i-1}^{r-1} i2^{i-1}(C_{m-i-2}^{m-r-2} + C_{m-i-2}^{m-r-1}) = r(2^m - 2^{r-1}) - \sum_{i-1}^{r-1} i2^{i-1} C_{m-i-1}^{m-r-1},$$

и теорема оказывается верной для m. Последнее равенство имеет место ввиду известного соотношения $C_\alpha^\beta + C_\alpha^{\beta+1} = C_{\alpha+1}^{\beta+1}$.

Для завершения доказательства остается заметить что при $r = m = 1$ теорема очевидна, так как

$$M_1^1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

и для кодирования требуется в точности одна операция сложения.

Так как

$$r(2^m - 2^{r-1}) - \sum_{i-1}^{r-1} i 2^{i-1} C_{m-i-1}^{m-r-1} \leq r(2^m - 2^{r-1}),$$

то $Q_m^r < \frac{n}{2} \log n$. Действительно, если $r \leq [m/2]$, то

$$r(2^m - 2^{r-1}) \leq r2^m \leq [m/2]2^m \leq \frac{n}{2} \log n.$$

Если же $r \geq [m/2]$, то

$$(2^m - 2^{r-1}) \leq 2^{m-1},$$

и

$$r(2^m - 2^{r-1}) \leq r2^{m-1} \leq m2^{m-1} = \frac{n}{2} \log n.$$

Как и всякий линейный двоичный код, любой РМ-код может задаваться в систематической форме. В случае такого задания символы (4.10.34) были бы информационными символами отвечающего им кодового вектора. Однако изложенное выше строение порождающей матрицы M_m^r, как бы ни были расположены ее строки, задает РМ-код не в систематической форме. Поэтому при кодировании символы (4.10.34) в явном виде не сохраняются и не отличаются от остальных, которые в ином случае назывались бы проверочными. Несмотря на это символы (4.10.34) называются информационными; выявляются они только после выполнения процесса декодирования, к описанию которого мы и переходим.

Информационный символ $u_{i_1 i_2 \ldots i_l}$, на который в (4.10.35) умножается вектор $v_{i_1} v_{i_2} \ldots v_{i_l}$ матрицы M_m^r, называется информационным символом $l-$го порядка, $i = 0, 1, \ldots, r$.

4.12. Декодирование кода Рида—Маллера

Декодирование начинается с нахождения информационных символов старшего, т.е. r-го порядка. Имеет место известная

Теорема 4.12.1. *Каждый из C_m^r информационных символов порядка r РМ-кода можно представить в точности 2^{m-r} способами в виде сумм 2^r компонент вектора b так, что каждое слагаемое входит только в одну сумму.*

Эти суммы будем называть проверочными. Доказательство теоремы здесь будет опущено из-за громоздкости и с целью экономии места.

Весь процесс декодирования будет проиллюстрирован подробным примером РМ-кода длины $n = 32$, порядка $r = 3$. Из (4.9.32) и (4.9.33) примеров 4.9. и 4.10. видно, что в векторе (4.10.35) каждый символ b_i^3 есть скалярное произведение вектора (4.10.34) на i-й столбец матрицы (4.9.33) $i = 1, 2, \ldots n$.

Поэтому

$$
\begin{aligned}
b_1^{(3)} &= u_0 \\
b_2^{(3)} &= u_0 + u_1 \\
b_3^{(3)} &= u_0 + u_2 \\
b_4^{(3)} &= u_0 + u_1 + u_2 + u_{21} \\
b_5^{(3)} &= u_0 + u_3 \\
b_6^{(3)} &= u_0 + u_1 + u_3 + u_{31} \\
b_7^{(3)} &= u_0 + u_2 + u_3 + u_{32} \\
b_8^{(3)} &= u_0 + u_1 + u_2 + u_3 + u_{21} + u_{31} + u_{32} + u_{321}.
\end{aligned}
$$

Складывая почленно эти равенства, получим

$$b_1^{(3)} + b_2^{(3)} + b_3^{(3)} + b_4^{(3)} + b_5^{(3)} + b_6^{(3)} + b_7^{(3)} + b_8^{(3)} = u_{321}. \quad (4.12.39)$$

Точно так же:

$$
\begin{aligned}
&b_9^{(3)} = u_0 + u_4\\
&b_{10}^{(3)} = u_0 + u_1 + u_4 + u_{41}\\
&b_{11}^{(3)} = u_0 + u_2 + u_4 + u_{42}\\
&b_{12}^{(3)} = u_0 + u_1 + u_2 + u_4 + u_{41} + u_{42} + u_{21} + u_{421}\\
&b_{13}^{(3)} = u_0 + u_3 + u_4 + u_{43}\\
&b_{14}^{(3)} = u_0 + u_1 + u_3 + u_4 + u_{31} + u_{41} + u_{43} + u_{431}\\
&b_{15}^{(3)} = u_0 + u_2 + u_3 + u_4 + u_{32} + u_{42} + u_{43} + u_{432}\\
&b_{16}^{(3)} = u_0 + u_1 + u_2 + u_3 + u_4 + u_{21} + u_{31} + u_{41} + u_{32} +\\
&+u_{42} + u_{43} + u_{321} + u_{421} + u_{431} + u_{432},
\end{aligned}
$$

откуда

$$b_9^{(3)} + b_{10}^{(3)} + b_{11}^{(3)} + b_{12}^{(3)} + b_{13}^{(3)} + b_{14}^{(3)} + b_{15}^{(3)} + b_{16}^{(3)} = u_{321}. \quad (4.12.40)$$

Аналогично

$$
\begin{aligned}
b_{17}^{(3)} + b_{18}^{(3)} + b_{19}^{(3)} + b_{20}^{(3)} + b_{21}^{(3)} + b_{22}^{(3)} + b_{23}^{(3)} + b_{24}^{(3)} &= u_{321}\\
b_{25}^{(3)} + b_{26}^{(3)} + b_{27}^{(3)} + b_{28}^{(3)} + b_{29}^{(3)} + b_{30}^{(3)} + b_{31}^{(3)} + b_{32}^{(3)} &= u_{321}
\end{aligned}
\quad (4.12.41)
$$

Четыре проверочных суммы (4.12.39), (4.12.40), (4.12.41) позволяют правильно определить символ u_{321}, если неверное значение приняла только одна из них. Таким образом, никакая одиночная ошибка в кодовом векторе (4.10.35) не может повлиять на правильное декодирование символа u_{321}, так как при этом только одна из четырех проверочных сумм принимает неверное значение.

Аналогичные четверки проверочных сумм согласно теореме 4.12.1 составляются для каждого из информационных символов третьего порядка. Например, для информационного символа u_{421} они будут выглядеть так:

$$
\begin{aligned}
b_1^{(3)} + b_2^{(3)} + b_3^{(3)} + b_4^{(3)} + b_9^{(3)} + b_{10}^{(3)} + b_{11}^{(3)} + b_{12}^{(3)} &= u_{421}\\
b_5^{(3)} + b_6^{(3)} + b_7^{(3)} + b_8^{(3)} + b_{13}^{(3)} + b_{14}^{(3)} + b_{15}^{(3)} + b_{16}^{(3)} &= u_{421}\\
b_{17}^{(3)} + b_{18}^{(3)} + b_{19}^{(3)} + b_{20}^{(3)} + b_{25}^{(3)} + b_{26}^{(3)} + b_{27}^{(3)} + b_{28}^{(3)} &= u_{421}\\
b_{21}^{(3)} + b_{22}^{(3)} + b_{23}^{(3)} + b_{24}^{(3)} + b_{29}^{(3)} + b_{30}^{(3)} + b_{31}^{(3)} + b_{32}^{(3)} &= u_{421}
\end{aligned}
\quad (4.12.42)
$$

А для символа u_{531} :

$$b_1^{(3)}+b_2^{(3)}+b_5^{(3)}+b_6^{(3)}+b_{17}^{(3)}+b_{18}^{(3)}+b_{21}^{(3)}+b_{22}^{(3)}=u_{531}$$
$$b_3^{(3)}+b_4^{(3)}+b_7^{(3)}+b_8^{(3)}+b_{19}^{(3)}+b_{20}^{(3)}+b_{23}^{(3)}+b_{24}^{(3)}=u_{531}$$
$$b_9^{(3)}+b_{10}^{(3)}+b_{13}^{(3)}+b_{14}^{(3)}+b_{25}^{(3)}+b_{26}^{(3)}+b_{29}^{(3)}+b_{30}^{(3)}=u_{531}$$
$$b_{11}^{(3)}+b_{12}^{(3)}+b_{15}^{(3)}+b_{16}^{(3)}+b_{27}^{(3)}+b_{28}^{(3)}+b_{31}^{(3)}+b_{32}^{(3)}=u_{531}$$

Общее правило построения проверочных сумм будет указано ниже. В случае, когда в символах b_i произошла только одна ошибка. все информационные символы третьего порядка могут быть правильно декодированы по правилу большинства. В случае двух ошибок, искажающих две проверочные суммы, наступает неопределенность — отказ от декодирования и сообщение об обнаружении ошибки, а в случае трех и более ошибок происходит ошибка декодирования.

После того, как декодированы все информационные символы третьего порядка (в общем случае — символы r-го порядка), следует образовать выражение

$$u_{321}v_3v_2v_1+u_{421}v_4v_2v_1+\ldots+u_{543}v_5v_4v_3, \tag{4.12.43}$$

которое затем вычитается из, быть может, искаженного вектора b:

$$b^{(3)}-(u_{321}v_3v_2v_1+u_{421}v_4v_2v_1+\ldots+u_{543}v_5v_4v_3)=$$
$$=b^{(2)}=(b_1^{(2)},b_2^{(2)},\ldots,b_{32}^{(2)}).$$

Можно считать, что полученный таким образом вектор $b^{(2)}$, представляет собой, быть может, искаженный вектор РМ-кода второго порядка, так как согласно (4.10.35) в его образовании не участвуют информационные символы третьего порядка и векторы-произведения третьего порядка.

Тогда по теореме 4.12.1 символы $b_i^{(2)}$ вектора $b^{(2)}$ можно объединить в проверочные суммы для информационных символов второго порядка. Например:

$$b_1^{(2)}+b_2^{(2)}+b_3^{(2)}+b_4^{(2)}=u_{21},\quad b_1^{(2)}+b_2^{(2)}+b_5^{(2)}+b_6^{(2)}=u_{31},\quad\ldots$$
$$b_5^{(2)}+b_6^{(2)}+b_7^{(2)}+b_8^{(2)}=u_{21},\quad b_3^{(2)}+b_4^{(2)}+b_7^{(2)}+b_8^{(2)}=u_{31},\quad\ldots$$
$$b_9^{(2)}+b_{10}^{(2)}+b_{11}^{(2)}+b_{12}^{(2)}=u_{21},\quad b_9^{(2)}+b_{10}^{(2)}+b_{13}^{(2)}+b_{14}^{(2)}=u_{31},\quad\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\quad\ldots$$
$$b_{29}^{(2)}+b_{30}^{(2)}+b_{31}^{(2)}+b_{32}^{(2)}=u_{21},\quad b_{27}^{(2)}+b_{28}^{(2)}+b_{31}^{(2)}+b_{32}^{(2)}=u_{31},\quad\ldots$$

$$\begin{array}{ll} \ldots & b_1^{(2)} + b_9^{(2)} + b_{17}^{(2)} + b_{25}^{(2)} = u_{54} \\ \ldots & b_2^{(2)} + b_{10}^{(2)} + b_{18}^{(2)} + b_{26}^{(2)} = u_{54} \\ \ldots & b_3^{(2)} + b_{11}^{(2)} + b_{19}^{(2)} + b_{27}^{(2)} = u_{54} \\ \ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots & b_8^{(2)} + b_{16}^{(2)} + b_{24}^{(2)} + b_{32}^{(2)} = u_{54} \end{array} \qquad (4.12.44)$$

Заметим, что в случае правильного декодирования символов третьего порядка вектор (4.12.43) не содержит ошибок, и если искаженным был некоторый символ вектора $b^{(3)}$, то искаженным будет и символ с тем же номером вектора $b^{(2)}$. Иначе говоря, векторы $b^{(3)}$, и $b^{(2)}$ содержат одинаковое число ошибок.

Каждая из десяти систем (4.12.44) состоит из восьми проверочных сумм. Казалось бы, что с их помощью можно правильно декодировать информационные символы даже при трех ошибках. Это было бы действительно так, если бы наш код с самого начала был кодом второго порядка. Однако он не второго, а третьего порядка, и мы обязаны начинать с декодирования символов третьего порядка. Но в силу сделанного выше замечания наличие трех или двух ошибок в векторе $b^{(2)}$ означает, что столько же ошибок содержалось в векторе b, а их-то при декодировании символов третьего порядка мы исправлять и не умеем.

Необходимо отметить, что *некоторые* ошибки более высокой кратности не влияют на правильность декодирования; но это означает только, что РМ-код не является совершенным кодом.

Итак, пусть декодированы информационные символы второго порядка. По аналогии с (4.12.43) составим сумму

$$u_{21}v_2v_1 + u_{31}v_3v_1 + u_{41}v_4v_1 + \ldots + u_{54}v_5v_4, \qquad (4.12.45)$$

которую вычтем затем из, быть может, искаженного вектора $b^{(2)}$;

$$b^{(1)} = b^2 - (u_{21}v_2v_1 + u_{31}v_3v_1 + u_{41}v_4v_1 + \ldots + u_{54}v_5v_4).$$

По вектору $b^{(1)}$ для каждого из пяти информационных символов первого порядка составим по 16 проверочных сумм. В одной сумме объединяются два символа вектора $b^{(1)}$. Например, для информационного символа u_1, как легко проверит

читатель самостоятельно, пользуясь строками со 2-й по 6-ю в (4.9.28) и строками со 2-й по 5-ю и 17-й в (4.9.32) и (4.9.33), получим $u_1 = b^{(1)}_{2t+1} + b^{(1)}_{2t+2}$, $t = 0, 1, \ldots, 15$.

Декодировав символы первого порядка, получим вектор

$$b^{(0)} = b^{(1)} - (u_1 v_1 + u_2 v_{21} + u_3 v_3 + u_4 v_4 + u_5 v_5). \qquad (4.12.46)$$

Если бы ошибок не было вовсе, то можно было бы утверждать, что $b^{(0)} = u_0 v_0$. При этом условии получаем: $u_0 = 0$, если вектор $b^{(0)}$ сплошь нулевой, и $u_1 = 1$, если вектор $b^{(0)}$ сплошь единичный. Если же ошибки были, то не все символы вектора $b^{(0)}$ одинаковы, и значение информационного символа u_0 определится по правилу большинства. На этом декодирование заканчивается.

В общем случае РМ-кода порядка r по вектору $b = b^r$ сначала с помощью 2^{m-r} проверочных сумм, построенных согласно теореме 4.12.1 для каждого информационного символа порядка r, декодируются все C^r_m этих символов. Правильное декодирование возможно при любых $2^{m-r-1}-1$ или менее ошибках в векторе b. Так как минимальное кодовое расстояние $d(m, r)$ РМ-кода порядка r по теореме 4.9.3 равно 2^{m-r}, то число ошибок, исправляемых кодом при помощи мажоритарного декодирования информационных символов порядка r, совпадает с числом ошибок, исправляемых по минимальному расстоянию. Как и в случае кодов, двойственных кодам Хэмминга, мажоритарное декодирование РМ-кодов реализует кодовое расстояние.

Если по вектору $b^{(l)}$ уже декодированы информационные символы порядка l, $l = r, r-1, \ldots, 1$, то следует составить сумму

$$\sum_{i_1, i_2, \ldots, i_l} v_{i_1} v_{i_2} \ldots v_{i_l},$$

после чего вычисляется вектор

$$b^{(l-1)} = b^{(l)} - \sum_{i_1, i_2, \ldots, i_l} v_{i_1} v_{i_2} \ldots v_{i_l}.$$

По вектору $b^{(l-1)}$ строятся проверочные суммы для информационных символов порядка $l-1$, и процесс продолжается, пока не будет найден последний информационный символ u_0.

На этом декодирование заканчивается.[3] Процедура декодирования РМ-кодов называется мажоритарным декодированием в r шагов.

4.13. Сложность декодирования кода Рида—Маллера

Обратимся теперь к выяснению сложности декодирования РМ-кодов. Очевидно, основной вклад в число операций при декодировании РМ-кода вносит вычисление всех проверочных сумм. Согласно теореме 4.12.1 для каждого из C_m^l информационных символов порядка l, $l = r, r-1, \ldots, 1$ в одной проверочной сумме надлежит произвести $2^l - 1$ сложений. Отсюда в случае РМ-кода порядка $r = 1$ точное значение числа сложений для декодирования m информационных символов первого порядка равно числу 2^{m-1} проверочных сумм, умноженному на m, т.е $\frac{n}{2}\log_2 n$. (Декодирование единственного информационного символа нулевого порядка обходится без сложений.)

Случай $r = m$ рассмотрен выше и не представляет интереса. При $r = m-1$ получается код с проверкой на четность, обнаруживающий любую одиночную ошибку, а при $r = m-2$ получается (см. выше) расширенный код Хэммнга. Как будет показано в конце данного раздела, этот код обходится без мажоритарного декодирования значительно более простыми средствами.

Для остальных значений r получим сначала грубую верхнюю оценку числа T_m^r сложений во всех проверочных суммах. Обозначим символом t_m^l число операций сложения при декодировании C_m^l информационных символов порядка $l = 1, 2, \ldots, r$. Помня, что в каждой проверочной сумме знаков сложения на единицу меньше, чем слагаемых, получим

$$t_m^l = (2^l - 1)2^{m-l}C_m^l = (2^m - 2^{m-l})C_m^l. \tag{4.13.47}$$

Откуда

$$T_m^r = \sum_{l=0}^{r} t_m^l < \sum_{l=0}^{m}(2^l - 1)2^{m-l}C_m^l =$$

[3] Доказательство теоремы 4.12.1 и правила построения проверочных сумм можно во всех подробностях найти в книге Ю.Л. Сагаловича "Кодирование состояний и надежность автоматов". М.: "Связь" , 1975. Эта книга не включена в список литературы, так как в него включены только источники, упомянутые в предисловии, и им отведена специальная роль.

$$= \sum_{l=0}^{m} (2^m - 2^{m-l}) C_m^l = 2^{2m} - 3^m < n^2. \qquad (4.13.48)$$

Формула (4.13.47) предполагает, что для различных информационных символов одного порядка числа сложений определяются независимо друг от друга.

На самом деле между проверочными суммами различных информационных символов одного и того же порядка легко усматривается очевидная связь. Например, частичная сумма $b_1^{(3)} + b_2^{(3)} + b_3^{(3)} + b_4^{(3)}$ входит в проверочные суммы (4.12.39) и (4.12.42) для информационных символов u_{321} и u_{421} соответственно. Точно то же самое имеет место и для частичной суммы $b_5^{(3)} + b_6^{(3)} + b_7^{(3)} + b_8^{(3)}$. Для тех же двух информационных символов u_{321} и u_{421} замечаем общую частичную сумму $b_{17}^{(3)} + b_{18}^{(3)} + b_{19}^{(3)} + b_{20}^{(3)}$ в (4.12.39) и (4.12.42). Далее, частичные суммы $b_1^{(2)} + b_2^{(2)}$ и $b_{31}^{(2)} + b_{32}^{(2)}$ входят в проверочные суммы для информационных символов u_{21} и u_{31} в (4.12.44).

Это означает, что, вычислив некоторую частичную сумму для одного информационного символа, ее можно использовать для другого информационного символа в готовом виде, не совершая вторичного ее вычисления. Строение проверочных сумм подробно исследовано в литературе (см. сноску 3 данной главы), и доказана

Теорема 4.13.1. *Для вычисления значений всех $2^{m-l}C_m^l$ проверочных сумм, отвечающим информационным символам порядка l, $l = 1, 2, \ldots, r$, достаточно*

$$t_m^l = \sum_{i=1}^{l} 2^{m-i} C_{m-l+i}^i \qquad (4.13.49)$$

операций сложения.

Суммирование по l величин (4.13.49) при мажорировании по $r = m$ и изменение порядка суммирования на основе известного тождества

$$\sum_{i=0}^{c} C_{a+i}^b = \begin{cases} C_{a+c+1}^{b+1}, \text{ при } b = a \\ C_{a+c+1}^{b+1} - C_a^{b+1}, \text{ при } b \le a-1 \end{cases}$$

дает

$$T_m^r = \sum_{l=1}^{r}\sum_{i=1}^{l} 2^{m-i} C_{m-l+i}^i = \sum_{i=1}^{r} 2^{m-i} \sum_{l=1}^{r} C_{m-(r-l)}^i <$$

$$< \sum_{i=1}^{m} 2^{m-i} \sum_{l=1}^{m} C_{m-(m-l)}^i = \sum_{i=1}^{m} 2^{m-i} C_{m+1}^i < 3^{m+1} = 3 \cdot n^{\log_2 3}.$$

Итак, наиболее трудоемкая часть декодирования — вычисление всех проверочных сумм — требует не более $3 \cdot n^{\log_2 3}$ операций сложения. Порядок всех остальных величин сложности декодирования РМ-кодов по всем этапам декодирования заведомо не превосходит этой величины. Поэтому можно утверждать, что максимальная сложность декодирования РМ-кодов выражается величиной $n^{\log_2 3}$ с небольшой мультипликативной константой.

Рассмотрим теперь вопрос о предпочтительности того или иного способа декодирования кода Хэмминга — мажоритарного или посредством вычисления синдрома. Пусть $n = 2^m - 1$ столбцов проверочной матрицы H_m кода расположены в лексикографическом порядке. Тогда H_m можно представить в виде

$$H_m = \left[\begin{array}{ccc} 00\ldots0 & 1 & 1\ldots1 \\ \hdashline & 0 & \\ H_{m-1} & 0 & H_{m-1} \\ & \vdots & \\ & 0 & \end{array}\right], \tag{4.13.50}$$

что ради наглядности демонстрируется матрицей (15, 11)- кода Хэмминга

$$H_4 = \left[\begin{array}{ccccccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right].$$

В матрице (4.13.50) пунктирная линия подчеркивает то обстоятельство, что к двум матрицам H_{m-1}, разделенным столбцом $(100\ldots0)^T$, добавлена строка длины $n = 2^m - 1$, состоящая из двух отрезков: нулевого длины $2^{m-1} - 1$ и сплошь единичного длины 2^{m-1}.

Теорема 4.13.2. *Вычисление синдрома (n,k)-кода Хэмминга требует в точности $2k$ операций сложения.*

Доказательство. Если для вычисления синдрома скалярное произведение s_i, $i = 1, 2, \dots m-1$ принятого вектора v на $i-$ю строку двух матриц H_{m-1} уже получено, то для получения скалярных произведений вектора v на $i-$е же (считая снизу) строки матрицы H_m требуется еще

1. По одной операции сложения на каждую строку, т.е. ровно $m-1$ операций.

2. Получить скалярное произведение вектора v на новую — верхнюю, т.е. m-ю строку.

Очевидно, для этого требуется получить произведение вектора v только на сплошь единичный отрезок длины 2^{m-1}. Однако произведения на систему 2^{m-1} отрезков длины 2^j, $j = 0, 1, \dots, m-2$, покрывающих $2^{m-1}-1$ единиц этого отрезка, кроме первой слева, уже выполнены в ступенчато расположенных слева направо отрезках той же длины в строках с номерами $j+1$ правой матрицы H_{m-1}. Для соединения этих произведений в одну сумму, включая и первую единицу слева в строке с номером m, требуется $m-1$ операций сложения.

Положим теперь, что теорема верна для $(2^{m-1}-1,\ 2^{m-1}-1-(m-1))$-кода Хэмминга. Для двух матриц H_{m-1} получаем $4k = 4(2^{m-1}-1-(m-1))$. Добавим к этой величине $2(m-1)$ операций сложения: $4(2^{m-1}-1-(m-1)) + 2(m-1) = 2(2^m-1-m)$, а это и есть удвоенное число информационных символов кода Хэмминга длины 2^m-1.

Непосредственно проверяется справедливость теоремы для $m-1=1$. В самом деле, $H_1 = [1]$. Отсюда

$$H_2 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Имеем $n = 3$, $n-k = 2$, $k = 1$, $2k = 2$, что отвечает двум операциям сложения, по одной на каждую строку матрицы H_2.

Итак, сложность вычисления синдрома для кода Хэмминга растет линейно с его длиной. Такова же сложность декодирования по синдрому, который имеет длину $m = \log_2(n+1)$. Таким образом, код Хэмминга, который посредством укорочения может быть получен из РМ-кода длины $n = 2^m$ и порядка $r = m-2$, а потому допускает мажоритарное декодирование в $r = m-2$ шагов, никогда такой процедуре не подвергается. Зато код, двойственный коду Хэмминга, согласно теореме 4.8.1

проще всего декодировать как укороченный РМ-код порядка $r = 1$.

4.14. Матрицы Адамара

Определение 4.14.1. *Матрицей Адамара называется квадратная $(n \times n)$-матрица, элементами которой являются действительные числа $+1$ и -1, и строки попарно ортогональны над полем действительных чисел. Ясно, что скалярное произведение любой строки на самое себя (над полем действительных чисел) равно n.*

В виде формулы определение 4.14.1 выглядит так: $H_n H_n^T = nI_n$. Действительно, все диагональные элементы произведения равны n, так как эти элементы суть скалярные произведения строк на самих себя.

Все остальные элементы матрицы-произведения равны нулю, так как они суть произведения различных строк, которые ортогональны.

Теорема 4.14.2. *Если H_n есть $(n \times n)$-матрица Адамара, то*

$$H'_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$$

есть $(2n \times 2n)-$матрица Адамара.

Доказательство. Положим, u и v есть строки матрицы H_n, $u \neq v$.

Строки матрицы H' имеют вид: $V = (u, u)$, $V = (v, v)$, и также $V = (u, -u), V = (v, -v)$. Пусть строки V пронумерованы числами $j = 1, 2, \ldots, 2n) : V_j$.

Возможны следующие случаи:

1. $j_2 = j_1 + n$.

Тогда $V_{j_1} \cdot V_{j_2} = (v, v) \cdot (v, -v) = vv - vv = n - n = 0$.

2. $1 \leq j_1 \neq j_2 \leq n$.

Тогда $V_{j_1} \cdot V_{j_2} = (v, v) \cdot (u, u) = vu + vu = 0 + 0 = 0$.

3.1 $+ n \leq j_1 \neq j_2 \leq 2n$.

Тогда $V_{j_1} \cdot V_{j_2} = (v, -v) \cdot (u, -u) = vu + vu = 0 + 0 = 0$.

4. $1 \leq j_1 = j_2 \leq n$.

Тогда $V_{j_1} \cdot V_{j_2} = (v, v) \cdot (v, v) = v \cdot v + v \cdot v = n + n = 2n$.

5. $1 + n \leq j_1 = j_2 \leq 2n$.

Тогда $V_{j_1} \cdot V_{j_2} = (v, -v) \cdot (v, -v) = v \cdot v + (-v) \cdot (-v) = n + n = 2n$.

6. $1 \le j_1 \le n$; $1 + n \le j_2 \le 2n$; $j_2 - j_1 \ne n$.

Тогда $V_{j_1} \cdot V_{j_2} = (v, v) \cdot (u, -u) = v \cdot u + (v \cdot -u) = 0 = 0$.

Во всех случаях соблюдены требования определения 4.14.1.

П р и м е р 4. 11.

Для первых значений $n = 1$, 2 имеем $H_1 = [1]$,

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

На основании теоремы 4.14.2 имеем

$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \tag{4.14.51}$$

Вообще, покажем, что матрицы Адамара могут быть только размера (1×1), (2×2) и $4m \times 4m$, где m есть целое положительное число. Действительно, из ортогональности строк матрицы следует, что любые две строки и различаются и совпадают в $n/2$ компонентах. Если число строк матрицы $n > 2$, то существуют по меньшей мере три строки со свойством ортогональности. Пусть это будут строки

$$\begin{array}{cccc} +1+1\ldots+1 & +1+1\ldots+1 & +1+1\ldots+1 & +1+1\ldots+1 \\ +1+1\ldots+1 & +1+1\ldots+1 & -1-1\ldots-1 & -1-1\ldots-1 \\ \underbrace{+1+1\ldots+1}_{i} & \underbrace{-1-1\ldots-1}_{j} & \underbrace{+1+1\ldots+1}_{k} & \underbrace{-1-1\ldots-1}_{l} \end{array}.$$

Из ортогональности строк следует:

$i + j - k - l = 0$, $i - j + k - l = 0$, $i - j - k + l = 0$. Отсюда $i = j = k = l$, а так как $n = i + j + k + l$, то $n = 4i$, что и требовалось.

Из теоремы 4.14.2 следует, что $(2^m \times 2^m)$-матрицы Адамара существуют для любого целого положительного m.

Заменим теперь в матрице Адамара H_n все $+1$ нулями, а все -1 единицами. Получится двоичный эквидистантный код A длины n с расстоянием $d = n/2$, содержащий n векторов. Только при $n = 2^m$ этот код может быть линейным. Сложим по модулю 2 все векторы этого кода со сплошь единичным вектором. Получится код той же длины n, содержащий $2n$ векторов. Если код линейный, то он эквивалентен РМ-коду первого порядка.

Таким способом из матрицы (4.14.51) получается код, который после некоторой перестановки строк примет вид

$$\begin{array}{c}1111\\1100\\1010\\0110\\0011\\0101\\1001\\0000\end{array}$$

Первые три строки есть не что иное, как порождающая матрица РМ-кода длины $n = 2^2$ порядка $r = 1$.

Если в коде A любой столбец сложить по модулю 2 со всеми остальными и с самим собой, то получим код длины $n = 2^m - 1$, содержащий $n = 2^m$ векторов. Это код, двойственный коду Хэмминга. Другой способ получить код, двойственный коду Хэмминга, показан на нижеследующем примере. Он получается из H_4, по правилу теоремы 4.14.2, если затем отбросить нулевой столбец (в данном случае – первый):

$$\begin{array}{c}00000000\\01010101\\00110011\\01100110\\00001111\\01011010\\00111100\\01101001\end{array} \cdot \qquad (4.14.52)$$

Разумеется, метод теоремы 4.14.2 не единственный метод построения матриц Адамара. Одним из полных источников о матрицах Адамара может служить книга [7], к которой мы и отсылаем читателя. Здесь же укажем, что, быть может, за редчайшим исключением были построены матрицы Адамара всех востребованных размеров (кратных четырех!).

Приведем матрицу Адамара размера 12×12.

$$H_{12} = \left[\begin{array}{cccccccccccc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 \\
1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 \\
1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 \\
1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & 1 \\
1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 & -1 & = 1
\end{array}\right].$$

Хорошо видно, что после замены всех символов 1 на 0 и всех символов –1 на единицы левый столбец матрицы становится нулевым и отбрасывается. Получается нелинейный эквидистантный код длины $n = 11$ с расстоянием $d = 6$, состоящий из 12 векторов.

4.15. Заключение

В этой главе изучена общая теория линейных кодов. Если никак не конкретизировать индивидуальные свойства линейных кодов, то бросается в глаза их главный недостаток: экспоненциальная сложность их декодирования. Поэтому на фоне общей теории рассмотрены важные частные случаи. В первую очередь — это коды, допускающие мажоритарное декодирование. Сложность их декодирования имеет порядок, либо линейный относительно длины n, либо $n \log_2 n$, либо $n^{\log_2 3} = 3^{\log_2 n}$. С другой стороны, коды с мажоритарным декодированием проигрывают в скорости передачи. В следующих главах изучаются некоторые коды с замечательными свойствами простоты кодирования и декодирования. Но и они отнюдь не исчерпывают всего многообразия кодов. С богатым арсеналом кодов можно познакомиться в обширной литературе по помехоустойчивому кодированию.

4.16. Задачи к главе 4

4.1. Показать, что любые две линейные комбинации строк порождающей матрицы линейного кода, различающиеся коэффициентами, представляют собой различные кодовые комбинации.

4.2. Показать, что в линейном (n,k)-коде A, содержащем q^k векторов, существует k линейно независимых векторов, но не существует $k+1$ линейно независимых векторов.
4.3. В пространстве V, содержащем q^n векторов длины n, найти число векторов, ортогональных к данному вектору.
4.4. Показать. что множество всех векторов из пространства V, содержащего 2^n векторов длины n, и ортогональных строкам порождающей матрицы линейного двоичного (n,k)-кода, образуют линейное подпространство. Всегда ли оно имеет размерность $(n-k)$?
4.5. Пусть H – проверочная матрица линейного кода. Показать, что смежный класс, синдром которого равен S, содержит вектор веса w тогда и только тогда, когда некоторая линейная комбинация w столбцов матрицы H равна S.
4.6. Показать, что если все векторы q-ичного линейного (n,k)-кода записаны как строки матрицы (без нулевого столбца), то в произвольном ее столбце каждый из элементов $0,\ 1\ldots q-1$ содержится в точности q^{k-1} раз. Показать, что сумма весов всех векторов (n,k)-кода равна $n(q-1)q^{k-1}$.
4.7. Показать, что в двоичном линейном коде либо все векторы имеют четный вес, либо векторов четного и нечетного веса – поровну.
4.8. Показать, что если двоичный линейный (n,k)-код имеет нечетный минимальный вес, то добавление к каждому его вектору суммы всех его символов по модулю два увеличивает минимальный вес кода на 1.
4.9. Показать, что линейный (n,k)-код с минимальным весом w исправляет любые $w-1$ стираний путем решения системы линейных уравнений. Показать, что существует по меньшей мере один случай, когда w стираний не могут быть исправлены.
4.10. Показать, что каков бы ни был равномерный двоичный код, удалив из него не более половины векторов, можно получить код, обнаруживающий любую одиночную ошибку.
4.11. Показать, что число различных базисов пространства всех 2^n двоичных векторов равно $(2^n-1)(2^n-2)\ldots(2^n-2^{n-1})/n!$.
4.12. Показать, что число различных двоичных линейных (n,k)-кодов равно

$$\frac{(2^n-1)(2^n-2)\cdots(2^n-2^{k-1})}{(2^k-1)(2^k-2)\cdots(2^k-2^{k-1})}.$$

4.13. Показать, что кодовое расстояние линейного (n,k)-кода удовлетворяет неравенству

$$d \leq n\frac{q^{k-1}(q-1)}{q^k-1}. \tag{4.16.53}$$

Это неравенство представляет собой верхнюю границу Плоткина для кодового расстояния линейных кодов.

4.14. Показать, что если $n > d$, и $B(n, d)$ есть максимально возможное число кодовых векторов линейного кода длины n с минимальным расстоянием d, то $B(n, d) \leq qB(n-1, d)$.

4.15. Показать, что при $n = 2d-1$ максимальное число кодовых векторов двоичного линейного кода не превосходит $2d$.

4.16. Показать, что максимальное число кодовых векторов двоичного линейного кода длины $n = 9$ с кодовым расстоянием $d = 5$ равно 4.

4.17. Пусть

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

есть порождающая матрица линейного $(7, 4)$-кода.

– Найти ее приведенно-ступенчатую форму.

– Найти проверочную матрицу H этого кода.

– Закодировать информационные символы 1101.

– Исправить ошибку в принятом векторе 1111100.

4.18. Пусть G_1 и G_2 порождающие матрицы линейных кодов с параметрами соответственно n_1, k, d_1 и n_2, k, d_2. Показать, что линейные коды с порождающими матрицами соответственно

$$\begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}$$

и

$$[G_1 \quad G_2]$$

имеют параметры соответственно

$$N = n_1 + n_2, K = 2k, D = \min\{d_1, d_2\}$$

и

$$N = n_1 + n_2,\ K = k,\ D \geq d_1 + d_2.$$

4.19. Доказать, что в линейном $(n, k, d_0)-$коде

$$n \geq d_0 + d_1 + \ldots + d_{k-1},$$

где

$$d_{i+1} = [(d_i - 1)/2],\ i = 0, 1, \ldots, k-2.$$

Глава 5.

Циклические коды

5.1. Циклический код как идеал

Циклические коды — это тот отдел теории помехоустойчивого кодирования, который демонстрирует плодотворное сочетание математического изящества и практической пользы.

Определение 5.1.1. *Циклическим кодом называется линейное векторное подпространство $A \subset V$ — пространства всех векторов длины n над полем $GF(q)$, выдерживающее циклический сдвиг компонент своих векторов,*

Иными словами, если код циклический, и вектор

$$u = (u_0, u_1, \ldots, u_{n-1})$$

принадлежит коду, то и вектор

$$u^* = (u_{n-1}, u_{0,}u_1, \ldots, u_{n-2})$$

также принадлежит коду.

В главе 3 уже было использовано сопоставление вектору того многочлена, набор коэффициентов которого совпадает с этим вектором. В многочленном представлении циклический сдвиг вектора, очевидно, интерпретируется следующим образом: если

$$u(x) = u_0 + u_1 x + \ldots + u_{n-1}x^{n-1},$$

то

$$u^*(x) = u_{n-1} + u_0 x + \ldots + u_{n-2}x^{n-1}.$$

Само собой напрашивается действие, посредством которого многочлен $u(x)$ преобразуется в $u^*(x)$. Рассмотрим произведение $xu(x) = u_0 x + \ldots + u_{n-2}x^{n-1} + u_{n-1}x^n$. Для того чтобы

вектор коэффициентов этого многочлена был равен u^*, коэффициент u_{n-1} должен быть при нулевой степени x, а не при x^n. Это требование выполнилось бы автоматически при условии $x^n = 1$.

В связи со сказанным исследуем кольцо $F[x]$ многочленов над $GF(q)$, а вслед за ним — кольцо $F[x]_{/(x^n-1)}$ классов вычетов многочленов по модулю $x^n - 1$. Одному классу вычетов принадлежат все многочлены, дающие одинаковые остатки при делении на $x^n - 1$. Разумеется, степени всех остатков не превосходят $n - 1$.

Теорема 5.1.2. *Все многочлены степени, не выше, чем $n-1$, принадлежат различным классам вычетов.*

Д о к а з а т е л ь с т в о. Предположим противное, т.е. пусть $f_1(x) \neq f_2(x)$, но $f_1(x) \equiv f_2(x)(\mathrm{mod}\, x^n - 1)$. Тогда получим, что $f_1(x) - f_2(x) \equiv 0(\mathrm{mod} x^n - 1)$. Степень левой части этого сравнения не превосходит $n-1$. Значит, ее делимость на $x^n - 1$ обеспечивается тем, что $f_1(x) - f_2(x) = 0$, откуда $f_1(x) = f_2(x)$, что противоречит условию.

Из теоремы 5.1.2 следует, что $F[x]$ распадается на q^n не пересекающихся классов вычетов. Все многочлены из $F[x]$, которые дают остаток 0, принадлежат тому классу, которому принадлежит и $x^n - 1$. Вместо рассуждений в терминах классов вычетов, будем вести рассмотрение в терминах их нормальных представителей, т.е. в терминах многочленов минимальных степеней в классах.

Впредь слова "вектор" и "многочлен", если не оговорено противное, будут употребляться как синонимы, смотря по случаю, и из соображений удобства. Так, например, правомерным будет словосочетание "вектор $u(x)$".

Теорема 5.1.3. *В кольце $F[x]_{/(x^n-1)}$ линейный код A будет циклическим тогда и только тогда, когда он идеал.*

Д о к а з а т е л ь с т в о. Пусть код A циклический. Тогда вместе с вектором $u(x)$ коду принадлежат и все сдвиги $x^i u(x)$, $i = 1, 2, \ldots, n-1$, а также все произведения $b_i x^i u(x)$, $b_i \in GF(q)$, и их сумма

$$u(x) \sum_{0}^{n-1} b_i x^i,$$

каков бы ни был набор коэффициентов b_i, так как код — подпространство. Это означает, что коду вместе с вектором $u(x)$ принадлежит его произведение на произвольный вектор кольца $F[x]_{/(x^n-1)}$. Остается вспомнить определение 2.16.1 идеала.

В обратную сторону доказательство дается одной строкой: если код A — идеал, то вместе с $u(x)$ ему принадлежит и $xu(x)$, а это и означает, что код циклический.

Доказанная теорема представляет собой критерий, и потому определением циклического кода может служить формулировка: циклический код A — это идеал в кольце $F[x]_{/(x^n-1)}$.

Теорема 5.1.4. *Идеал A содержит единственный нормированный многочлен $g(x)$ минимальной степени r.*

Доказательство. Предположим противное. Пусть кроме многочлена $g(x)$ в A имеется еще нормированный многочлен $f(x)$ той же степени. Ясно, что $(g(x) - f(x)) \in A$, и $\deg(g(x) - f(x)) < r$, так как старшие коэффициенты обоих многочленов $g(x)$ и $f(x)$ одинаковы. Противоречие завершает доказательство.

Следствие 5.1.5. *Все многочлены идеала A кратны многочлену $g(x)$.*

Действительно, предположим противное, т.е.пусть $f(x) \in A$, но $f(x) = q(x)g(x) + r(x)$.

Имеем последовательно: $g(x) \in A$, $q(x)g(x) \in A$, так как A есть идеал; $r(x) = f(x) - q(x)g(x) \in A$. Получается, что $\deg r(x) < \deg g(x)$, чего быть не может. Отсюда $r(x) = 0$.

В соответствии с разделом 2.16 такой идеал называется главным. В кольце $F[x]_{/(x^n-1)}$ все идеалы главные, и оно, таким образом, есть кольцо главных идеалов.

Определение 5.1.6. *Многочлен $g(x)$ называется порождающим многочленом идеала, т.е. циклического кода.*

Теорема 5.1.7. *Любой многочлен $g(x)$, порождающий идеал в кольце $F[x]_{/(x^n-1)}$, делит многочлен $x^n - 1$.*

Доказательство. Пусть $x^n - 1 = q(x)g(x) + r(x)$, где выполняется нервенство $\deg r(x) < \deg g(x)$. Так как в кольце

$F[x]_{/(x^n-1)}$ многочлен x^n-1 принадлежит нулевому классу вычетов, то в нем $q(x)g(x)+r(x)=0$, а потому $q(x)g(x)=-r(x)$, и $r(x)$ принадлежит идеалу, что может быть только при $r(x)=0$ из-за соотношения $\deg r(x) < \deg g(x)$.

5.2. Порождающая матрица циклического кода

Теорема 5.2.1. *Если* $\deg g(x) = r$, *то векторы (многочлены)*

$$g(x),\ xg(x),\ \ldots,\ x^{n-r-1}g(x)$$

линейно независимы.

Действительно, линейная зависимость означала бы существование такого набора не всех равных нулю элементов $u_i \in GF(q)$, $i = 0, 1, \ldots, n-r-1$, что

$$\sum_{i=0}^{n-r-1} u_i x^i g(x) = 0,$$

чего быть не может, так как многочлен под знаком суммы имеет степень не выше, чем $n-1$, а потому не может делиться на x^n-1 и, следовательно, не может принадлежать нулевому классу вычетов кольца $F[x]_{/(x^n-1)}$.

Отсюда немедленно следует, что порождающая матрица циклического кода над $GF(q)$ длины n с порождающим многочленом $g(x) = g_0 + xg_1 + \ldots + x^r g_r$ степени r имеет вид

$$G = \begin{bmatrix} g(x) \\ xg(x) \\ \ldots \\ x^{n-r-1}g(x) \end{bmatrix} = \begin{bmatrix} g_0 & g_1 & \ldots & g_r & 0 & 0 & \ldots & 0 \\ 0 & g_0 & g_1 & \ldots & g_r & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & g_0 & g_1 & \ldots & g_r \end{bmatrix}. \tag{5.2.1}$$

Если u есть информационный вектор длины k, то кодирование (4.2.1), учитывая синоним "вектор — многочлен", принимает форму $v(x) = u(x)g(x)$. Матрица (5.2.1) имеет $k = n-r$ строк и n столбцов, а потому число $r = n-k$ проверочных символов циклического кода равно степени порождающего многочлена.

П р и м е р 5. 1.

Положим $g(x) = 1 + x + x^3$ над $GF(2)$, $-1 = 1$. Можно проверить, что $x^7 - 1 = (1 + x)(1 + x + x^3)(1 + x^2 + x^3)$. Таким образом, $n = 7$, $r = 3$, $k = 4$.

$$G = \begin{bmatrix} 1 + x + x^3 \\ x(1 + x + x^3) \\ x^2(1 + x + x^3) \\ x^3(1 + x + x^3) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}. \quad (5.2.2)$$

Информационными векторами длины $k = 4$ будут

$$\begin{array}{ll} u^{(1)}(x) = 1 + x + x^2 + x^3, & u^{(2)}(x) = 1 + x + x^2, \\ u^{(3)}(x) = 1 + x + x^3, & u^{(4)}(x) = 1 + x^2 + x^3, \\ u^{(5)}(x) = x + x^2 + x^3, & u^{(6)}(x) = 1 + x, \\ u^{(7)}(x) = 1 + x^2, & u^{(8)}(x) = 1 + x^3, \\ u^{(9)}(x) = x + x^2, & u^{(10)}(x) = x + x^3, \\ u^{(11)}(x) = x^2 + x^3, & u^{(12)}(x) = 1, \\ u^{(13)}(x) = x, & u^{(14)}(x) = x^2, \\ u^{(15)}(x) = x^3, & u^{(16)} = 0. \end{array}$$

Прямым умножением находим

$$\begin{array}{l} u^{(1)}(x)g(x) = 1 + x^3 + x^5 + x^6, \\ u^{(2)}(x)g(x) = 1 + x^4 + x^5, \\ u^{(3)}(x)g(x) = 1 + x^2 + x^5, \\ u^{(4)}(x)g(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6, \\ u^{(5)}(x)g(x) = x + x^5 + x^6, \\ u^{(6)}(x)g(x) = 1 + x^2 + x^3 + x^4, \\ u^{(7)}(x)g(x) = 1 + x + x^2 + x^5, \\ u^{(8)}(x)g(x) = 1 + x + x^4 + x^6, \\ u^{(9)}(x)g(x) = x + x^3 + x^4 + x^5, \\ u^{(10)}(x)g(x) = x + x^2 + x^3 + x^6, \\ u^{(11)}(x)g(x) = x^2 + x^4 + x^5 + x^6, \\ u^{(12)}(x)g(x) = 1 + x + x^3, \\ u^{(13)}(x)g(x) = x + x^2 + x^4, \\ u^{(14)}(x)g(x) = x^2 + x^3 + x^5, \\ u^{(15)}(x)g(x) = x^3 + x^4 + x^6, \\ u^{(16)}(x)g(x) = 0. \end{array}$$

В процессе кодирования получен циклический код.

Напомним, что при всех сдвигах показатели степеней приводятся по модулю n.

Нетрудно убедиться, что

$$\begin{array}{ll} xu^{(6)}(x)g(x) = u^{(9)}(x)g(x); & xu^{(9)}(x)g(x) = u^{(11)}(x)g(x); \\ xu^{(11)}(x)g(x) = u^{(1)}(x)g(x); & xu^{(1)}(x)g(x) = u^{(8)}(x)g(x); \\ xu^{(8)}(x)g(x) = u^{(7)}(x)g(x); & xu^{(7)}(x)g(x) = u^{(10)}(x)g(x); \\ xu^{(10)}(x)g(x) = u^{(6)}(x)g(x). & \end{array} \tag{5.2.3}$$

Аналогично

$$\begin{array}{ll} xu^{(3)}(x)g(x) = u^{(12)}(x)g(x); & xu^{(12)}(x)g(x) = u^{(13)}(x)g(x); \\ xu^{(13)}(x)g(x) = u^{(14)}(x)g(x); & xu^{(14)}(x)g(x) = u^{(15)}(x)g(x); \\ xu^{(15)}(x)g(x) = u^{(2)}(x)g(x); & xu^{(2)}(x)g(x) = u^{(5)}(x)g(x); \\ xu^{(5)}(x)g(x) = u^{(3)}(x)g(x). & \end{array} \tag{5.2.4}$$

Назовем орбитой все те векторы кода, которые получаются друг из друга циклическими сдвигами. В приведенном примере 16 векторов кода распадается на 4 орбиты: две орбиты — (5.2.3) и (5.2.4) — по 7 векторов и две по одному вектору. Это $u^{(4)}(x)g(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6$ и 0. Орбиты не пересекаются, но одна из другой получаются сложением каждого вектора одной орбиты с вектором $u^{(4)}(x)g(x)$.

5.3. Проверочная матрица циклического кода

Многочлен $g(x)$ порождает циклический код, т.е. линейное циклическое подпространство (идеал) $A \subset V$, базисом которого является порождающая матрица (5.2.1). Существует ли многочлен, который таким же образом порождает нулевое подпространство нашего кода, и если существует, то как с его помощью изображается проверочная матрица H?

Согласно теореме 5.1.7 не составляет труда найти такой многочлен $h(x)$, что

$$h(x) = (x^n - 1)/g(x) = \sum_{i=0}^{k} h_i x^i. \tag{5.3.5}$$

Это значит, что

$(x^n-1) = g(x)h(x)$. Но в кольце $F[x]_{/(x^n-1)}$ имеем $x^n-1=0$, а потому

$$g(x)h(x) = 0.$$

Это равенство может рассматриваться как аналог равенства (4.3.3), тем более, что слагаемые $n-k$ и k в сумме степеней многочленов $g(x)$ и $h(x)$ такие же, как и размерности ортогональных подпространств, порождаемых матрицами H и G.

В связи с этим многочлен $h(x)$ называется проверочным многочленом кода A. Однако полной аналогии между соотношением "порождающий многочлен $g(x)$ — порождающая матрица G" и соотношением "проверочный многочлен $h(x)$ — проверочная матрица H" нет.

Действительно, пусть $f(x)g(x)-$ произвольный вектор кода A, и $f(x)g(x) = \sum_{i=0}^{n-1} c_i x^i = c(x)$. Найдем (в кольце $F[x]_{/(x^n-1)}$) произведение

$$c(x)h(x) = \sum_{i=0}^{n-1} c_i x^i \sum_{j=0}^{k} h_j x^j = f(x)g(x)h(x) = 0. \qquad (5.3.6)$$

В нем коэффициент при x^j равен

$$a_j = c_0 h_j + c_1 h_{j-1} + \ldots + c_j h_0 +$$

$$+\underline{c_{j+1}h_{n-1} + c_{j+2}h_{n-2} + \ldots + c_{n-1}h_{j+1}}. \qquad (5.3.7)$$

Разумеется, $h_l = 0$ для всех $l \geq k$.

Кроме того, в подчеркнутой части суммы нижние индексы каждого слагаемого таковы, что эти слагаемые являются как будто коэффициентами при x^{n+j}. Однако в кольце $F[x]_{/(x^n-1)}$ имеет место равенство $x^n = 1$.

Равенство (5.3.7) означает, что a_j есть скалярное произведение двух векторов:

$$a_j = (c_0, c_1, \ldots c_{n-1})(h_j, h_{j-1}, \ldots, h_0, h_{n-1}, h_{n-2}, \ldots, h_{j+1}).$$

В первой скобке — кодовый вектор $f(x)g(x)$ кода A. Вторая скобка содержит коэффициенты многочлена $h(x)$, расположенные в порядке *убывания* нижних индексов , а не возрастания, как в (5.2.1), и сдвинутые циклически на $j+1$ шагов вправо. И это скалярное произведение равно нулю ввиду (5.3.6). Таким

образом, все k сдвигов вектора $(h_k, h_{k-1}, \ldots, h_0)$ расположены в $n-k=r$ строках проверочной матрицы H кода A.

$$H = \begin{bmatrix} h(x) \\ xh(x) \\ \ldots \\ x^{n-k-1}h(x) \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 0 & \ldots & 0 & h_k & h_{k-1} & \ldots & h_0 \\ 0 & \ldots & 0 & h_k & h_{k-1} & \ldots & h_0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ h_k & h_{k-1} & \ldots & h_0 & 0 & 0 & \ldots & 0 \end{bmatrix}. \qquad (5.3.8)$$

Из изложенного следует, что код, порождаемый проверочным многочленом $h(x)$, эквивалентен коду B, который ортогонален коду A.

П р и м е р 5. 2.

Для случая циклического кода с порождающим многочленом $g(x) = x^3 + x + 1$ (см. пример 5.1) проверочным многочленом будет $h(x) = (x+1)(x^3+x^2+1) = x^4+x^2+x+1$, $g(x)h(x) = x^7 + 1$, $r = n - k = 3$.

Проверочная матрица такова:

$$H = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}. \qquad (5.3.9)$$

Нетрудно проверить, что все строки матрицы (5.2.2) ортогональны всем строкам матрицы (5.3.9).

Иногда ради краткости обе рассмотренные матрицы будем называть базисными матрицами циклического кода.

5.4. Каноническая форма базисных матриц циклического кода

Имея в виду, что циклический код — это линейный код, и следуя разделу 4.4, представим порождающую матрицу (5.2.1) в приведенно-ступенчатой, т.е. канонической, форме.

Пусть $x^i = q_i(x)g(x) + r_i(x)$, где $r_i(x)$ — остаток от деления x^i на $g(x)$.

Отсюда векторы

$$q_i(x)g(x) = -r_i(x) + x^i,\ i = n-k,\ n-k+1,\ \ldots,n-1,\ldots$$

суть кодовые векторы циклического кода A, порожденного многочленом $g(x)$. При $i = n-k,\ n-k+1,\ \ldots,n-1$ они линейно независимы, так как в слагаемых x^i показатели степеней различны. Поэтому их можно взять в виде базиса подпространства A, т.е. в виде строк порождающей матрицы G. Расположим правые части этих равенств в порядке возрастания индекса i сверху вниз. Кроме того, как это естественным образом получилось в (5.2.1), каждый вектор располагается в порядке возрастания показателей степеней x слева направо. Тогда получим

$$G = [-RI_k]\,, \tag{5.4.10}$$

где I_k — единичная $(k\times k)$-матрица, а $-R$ есть $(k\times(n-k))-$матрица, строка которой с номером $j = i-n+k+1$ образована коэффициентами остатка $-r_i(x)$, $i = n-k,\ n-k+1,\ \ldots,n-1$,

Таким образом, сравнивая матрицы (4.4.4) и (5.4.10), видим: последняя предполагает, что проверочные символы предшествуют информационным, а не наоборот, как это имеет место в (4.4.4).

В таком случае, согласно теореме 4.4.1, каноническая, т.е. приведённо-ступенчатая, форма проверочной матрицы циклического кода немедленно получается в виде

$$H = \left[I_{n-k}R^T\right]. \tag{5.4.11}$$

Здесь I_{n-k} есть единичная $(n-k)\times(n-k)$-матрица, R^T есть $(n-k)\times k$-матрица, столбец которой с номером $j = i-n+k+1$ образован коэффициентами остатка $r_i(x)$, $i = n-k, n-k+1,\ \ldots,n-1$. В полном соответствии с (5.3.8) в каждой строке многочлен расположен в порядке убывания степеней x.

Формулы (5.4.10) и (5.4.11) свидетельствуют, что первые $n-k$ символов кодового вектора, т.е. коэффициенты при степенях $x^0, x,\ \ldots,x^{n-k-1}$ выбраны проверочными, а последние k символов — информационными.

П р и м е р 5. 3.

Пусть снова

$$g(x) = 1+x+x^3, n = 7, k = 4, r = n-k = 3.$$

Имеем

$$\begin{array}{rccccccc}
g(x) = & 1 & +x & & +x^3, & & & \\
xg(x) = & & x & +x^2 & & +x^4, & & \\
(x^2+1)g(x) = & 1 & +x & +x^2 & & & +x^5, & \\
(x^3+x+1)g(x) = & 1 & & +x^2 & & & & +x^6.
\end{array}$$

Базисные векторы-многочлены циклического кода:

$$1+x+x^3, \quad x+x^2+x^4, \quad 1+x+x^2+x^5, \quad 1+x^2+x^6.$$

Каноническая форма порождающей матрицы:

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

Отсюда немедленно следует

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

Построим теперь проверочную матрицу H, пользуясь проверочным многочленом $h(x) = (x^7+1)/g(x) = x^4+x^2+x+1$.

Имеем

$$\begin{array}{rccccccc} (x^2+1)h(x) = & x^6 & & & +x^3 & & +x & +1, \\ xh(x) = & & x^5 & & +x^3 & +x^2 & +x, & \\ h(x) = & & & x^4 & & +x^2 & +x & +1. \end{array}$$

Базисные векторы-многочлены подпространства, ортогонального подпространству строк матрицы G :

$$x^6+x^3+x+1, \quad x^5+x^3+x^2+x, \quad x^4+x^2+x+1.$$

Располагая эти многочлены в порядке *убывания* показателей степеней x, получим проверочную матрицу

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

Совпадение очевидно! Читатель без труда проверит справедливость равенства GH^T.

Читателю полезно усвоить, что обмен местами матриц I и R в канонических формах матриц G и H циклического кода не имеет никакого принципиального значения, но является следствием исключительно соглашения. Недаром выше подчеркивается, что все многочлены в матрице G (5.2.1), как и было

условлено, располагаются по возрастающим степеням x, и как в связи с этим получилось, в матрице H (5.3.8), они располагаются по убывающим степеням x.

Однако впервые в тексте читатель познакомился с канонической формой порождающей матрицы именно в виде (4.4.4), а с канонической формой проверочной матрицы именно в виде (4.4.6). Чтобы вернуться к более привычному виду базисных матриц, именно матрицам (4.4.4) и (4.4.6) следует выбирать первые k символов кодового вектора, т.е. коэффициенты при $x^{n-1}, x^{n-2}, \ldots, x^{n-k}$ информационными, а последние $n-k$ символов, т.е. коэффициенты при $x^{n-k-1}, x^{n-k-2}, \ldots, x^0$ — проверочными. Читатель легко самостоятельно трансформирует на этот случай выкладки примера 5.3.

5.5. Порождающий многочлен с заданными свойствами

Задача, сформулированная в заглавии, напрашивается сама собой, так как порождающий многочлен выступает главной характеристикой циклического кода. Один из признаков многочлена представлен его корнями в некотором поле. Например, в технической диагностике существенным является требование построить многочлен, который бы не делил ни одного из многочленов из заданного множества. Пользуясь кодовой терминологией, такая задача означает построение кода, который не содержал бы многочленов из заданного множества. Простым ее решением было бы построение такого многочлена, минимальный набор корней которого не содержался бы целиком в множествах корней заданных многочленов. Этот пример приведен только для того, чтобы показать, что построение кода с заданным расстоянием является не единственной задачей теории кодирования, хотя в данном руководстве она является основной.

Поэтому в самом общем виде задача выяснить свойства циклического кода — это задача выяснить их по корням порождающего многочлена.

Пусть элементы

$$\alpha_1, \alpha_2, \ldots, \alpha_\rho \in GF(q^m), m = 1, 2, \ldots \tag{5.5.12}$$

суть корни порождающего многочлена $g(x)$ над $GF(q)$. (Положим для простоты, что среди них нет кратных.)

Тогда эти же элементы являются корнями любого многочлена $v(x) = v_0 + v_1 x + v_2 x^2 + \ldots + v_{n-1} x^{n-1}$, принадлежащего

циклическому коду, порождаемому многочленом $g(x)$.

Это следует из того, что, если $v(x) = q(x)g(x)$, то $v(\alpha_i) = q(\alpha_i)g(\alpha_i) = 0$, $i = 1, 2, \ldots, \rho$.

Иначе говоря, $v(\alpha_i) = v_0 + v_1\alpha_i + v_2\alpha_i^2 + \ldots + v_{n-1}\alpha_i^{n-1} = 0$, а это означает, что равно нулю скалярное произведение

$$(v_0, v_1, v_2, \ldots v_{n-1})(1, \alpha_i, \alpha_i^2, \ldots \alpha_i^{n-1}).$$

В свою очередь это означает, что, векторы $(1, \alpha_i, \alpha_i^2, \ldots \alpha_i^{n-1})$ есть не что иное, как строки проверочной матрицы нашего циклического кода:

$$H = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \ldots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \ldots & \alpha_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha_\rho & \alpha_\rho^2 & \ldots & \alpha_\rho^{n-1} \end{bmatrix}, \qquad (5.5.13)$$

Далее, многочлен $g(x)$ по теореме 3.5.5 делится на все минимальные функции $m_i(x)$, $i = 1, 2, \ldots, \rho$, своих корней

$$\alpha_1, \alpha_2, \ldots, \alpha_\rho \in GF(q^m), m = 1, 2, \ldots$$

Следовательно, $g(x)$ делится на наименьшее общее кратное

$$M(m_1(x), m_2(x), \ldots, m_\rho(x)).$$

З а м е ч а н и е. Так как минимальная функция — неприводимый многочлен, требование делимости на наименьшее общее кратное, а не на произведение, обусловлено только тем, что различные элементы α_i имеют одну и ту же минимальную функцию. Последнее не исключено ввиду теоремы 3.5.12. Зато исключена отличная от единицы кратность корней ввиду оговорки, сделанной по поводу элементов (5.5.12).

Так как $\alpha_i \in GF(q^m)$, то $\deg m_i(x) \leq m$.

Отсюда

$$\deg g(x) = r \leq \sum_{i=1}^{\rho} \deg m_i(x) \leq \varrho m.$$

Но степень порождающего многочлена есть в точности число r проверочных символов порождаемого им кода.

Это означает, что число k информационных символов этого кода удовлетворяет соотношению $k \geq n - \rho m$.

Остается найти длину n циклического кода с таким порождающим многочленом $g(x)$, что $g(\alpha_i) = 0$, $i = 1, 2 \ldots, \rho$. Согласно теореме 5.1.7, многочлен $g(x)$ делит многочлен $x^n - 1$. Это значит, что все элементы α_i, будучи корнями многочлена $g(x)$, являются также корнями многочлена $x^n - 1$, т.е. удовлетворяют уравнению $x^n - 1 = 0$. Таким образом, $\alpha_i^n = 1$. Отсюда сразу следует, что число n делится на порядок каждого α_i, а значит, и на наименьшее общее кратное порядков корней (5.5.12). Иными словами, верна

Теорема 5.5.1. *Длина n циклического кода не может быть меньше наименьшего общего кратного порядков корней порождающего многочлена $g(x)$.*

С другой стороны, заметим, что как только длина n' кода по какой-нибудь причине превысит это значение n, двучлен $x^n - 1$ станет принадлежать коду, и его минимальное расстояние будет равным двум.

Так как корнями порождающего многочлена выбраны элементы конечного поля $GF(q^m)$, то помня, что порождающий многочлен $g(x)$ является делителем двучлена $x^n - 1$, следует установить связь между этим двучленом и конечным полем $GF(q^m)$.

Теорема 5.5.2. *Такое число m, что многочлен $x^n - 1$ делит многочлен $x^{q^m-1} - 1$, найдется тогда и только тогда, когда $(q, n) = 1$*

Согласно теореме 3.5.10 надлежит доказать, что $q^m - 1$ делится на n тогда и только тогда, когда $(q, n) = 1$.

Д о к а з а т е л ь с т в о. *Необходимость.*

Пусть $(q, n) = d > 1$. Так как число $q^m - 1$ заведомо не делится ни на один делитель числа q^m, то оно не делится и на число d, а значит и подавно не делится на n.

Достаточность. Пусть $(q, n) = 1$. Тогда по теореме Эйлера

$$q^{\varphi(n)} \equiv 1 (\mathrm{mod} n).$$

Поэтому достаточно взять $m = \varphi(n)$. Однако, если q не есть первообразный корень по модулю n, то в качестве m берут показатель δ, которому число q принадлежит по модулю n. Как известно (см. утверждение 1.11.4), этот показатель делит $\varphi(n)$.

В теории циклических кодов это число m принято называть мультипликативным порядком числа q по модулю n.

Если число m есть показатель, которому которому число q принадлежит по модулю n, то $x^n - 1$, не делит никакого числа вида $q^l - 1$ при $l < m$.

Не обращаясь к теореме Эйлера, существование числа m можно доказать следующим образом. Положим

$$\begin{array}{ll} q = nt_1 + r_1, & (0 < r_1 \leq n-1) \\ q^2 = nt_2 + r_2, & (0 < r_2 \leq n-1) \\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots \\ q^n = nt_n + r_n, & (0 < r_n \leq n-1) \end{array}$$

Получается что среди n чисел r_i, $(i = 1, 2, \ldots, n)$ оказывается не более, чем $n-1$ различных. Значит, не менее двух из них совпадают. Пусть $r_y = r_z = r$, и $y \geq z$. Тогда находим последовательно $q^y - q^z = nt_y + r - nt_z - r$, $q^z(q^{y-z} - 1) = (t_y - t_z)n$. Из $(q, n) = 1$ следует, что n делит $q^{y-z} - 1$, и потому $m = y - z$, что и требовалось.

Длинами n циклических кодов могут быть только делители чисел $q^m - 1$ и, разумеется, само это число. Подобно тому, как в разделе 3.2 поле $GF(q^m)$ называется полем разложения многочлена $x^{q^m} - x$, оно же называется полем разложения многочлена $x^n - 1$.

Если $n = q^m - 1$, то циклический код называется *примитивным.*

На случай $q = 2$ в Таблице П.1 представлены канонические разложения чисел $2^m - 1, m = 1, 2, \ldots, 34$.

Из циклического кода можно построить код меньшей длины путем его укорочения. Новый код уже не будет циклическим. Можно показать, что он будет т.н. квазициклическим кодом в кольце многочленов уже по другому модулю, а не по модулю многочлена $x^n - 1$.

П р и м е р 5. 4.

Пусть корнями порождающего многочлена будут α, α^3, $\alpha^5 \in GF(2^4)$. Элемент α — примитивный. Элементы α^3 и α^5 имеют порядки 5 и 3 соответственно. Длина кода $n = 15$.

На основании примера 3.9 в разделе 3.5 элемент α может быть корнем одного из двух примитивных многочленов 4-й степени над $GF(2)$: $x^4 + x + 1$, или $x^4 + x^3 + 1$ в зависимости от того, по модулю какого из них построено поле $GF(2^m)$.

Пусть, для определенности, это будет $x^4 + x + 1$. Минимальные функции элементов α^3 и α^5 суть $x^4 + x^3 + x^2 + x + 1$ и $x^2 + x + 1$.

Порождающий многочлен

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

$$= 1 + x + x^2 + x^4 + x^5 + x^8 + x^{10}, r = 10, k = 5.$$

Порождающая матрица

$$G = \left[\begin{array}{l} 111011001010000 \\ 011101100101000 \\ 001110110010100 \\ 000111011001010 \\ 000011101100101 \end{array}\right].$$

Проверочная матрица будет

$$H = \left[\begin{array}{ccccccccccccccc} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 & \alpha^8 & \alpha^9 & \alpha^{10} & \alpha^{11} & \alpha^{12} & \alpha^{13} & \alpha^{14} \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} & 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} \\ 1 & \alpha^5 & \alpha^{10} & 1 & \alpha^5 & \alpha^{10} & 1 & \alpha^5 & \alpha^{10} & 1 & \alpha^5 & \alpha^{10} & 1 & \alpha^5 & \alpha^{10} \end{array}\right]. \tag{5.5.14}$$

Заменим теперь в матрице H каждый элемент α^i двоичным вектором-столбцом, пользуясь таблицей поля (3.4.11). Получим изображение проверочной матрицы над полем $GF(2)$:

$$H = \left[\begin{array}{l} 100010011010111 \\ 010011010111100 \\ 001001101011110 \\ 000100110101111 \\ \hline 100011000110001 \\ 000110001100011 \\ 001010010100101 \\ 011110111101111 \\ \hline 101101101101101 \\ 011011011011011 \\ 011011011011011 \\ 000000000000000 \end{array}\right].$$

В этой матрице 12 строк, вопреки тому, что $r = 10$. Противоречие легко снимается, если заметить, что 12-я строка получилась нулевой и потому не имеет смысла, а 10-я и 11-я строки совпадают, и одна из них подлежит удалению.

Пунктирными линиями отделены строки матрицы (5.5.14)

5.6. Циклический код Хэмминга

Матрицы (0.4.9) и (0.4.11), отличающиеся порядком расположения столбцов являются проверочными матрицами (7, 4)-кода. В общем случае от проверочной матрицы кода Хэмминга с параметрами $n = 2^m - 1, k = 2^m - m - 1$ требуется только, чтобы все ее $2^m - 1$ столбцов высоты m были различны, а порядок их следования роли не играет, так как он не влияет на корректирующую способность кода. В таком случае будем трактовать все эти столбцы как векторы длины m, изображающие все ненулевые элементы $\alpha^i \in GF(2^m)$, $i = 0, 1, \ldots, 2^m - 2$, где α — примитивный элемент поля. Расположив столбцы в порядке возрастания показателя степени, проверочную матрицу можно представить в виде

$$H = [\ 1 \quad \alpha \quad \alpha^2 \quad \ldots \quad \alpha^{2^m-2}\ ]. \tag{5.6.15}$$

Если $u(x)$ есть кодовый вектор кода Хэмминга, то $u(\alpha)H^T = 0$. Таким образом, элемент $\alpha \in GF(2^m)$ является корнем произвольного вектора кода Хэмминга с проверочной матрицей (5.6.15). Это значит, что любой кодовый вектор делится на минимальную функцию элемента $\alpha \in GF(2^m)$, а потому она представляет собой порождающий многочлен $g(x)$ степени m над $GF(2)$ циклического кода Хэмминга длины $n = 2^m - 1$. Выше было установлено, что минимальное расстояние кода Хэимминга $d = 3$. В следующей главе будет показано, что в случае циклического кода Хэмминга этот факт есть прямое следствие из того, что порождающий многочлен $g(x)$ является минимальной функцией примитивного элемента $\alpha \in GF(2^m)$.

5.7. Векторы всех циклических кодов простой длины $n = 2^m - 1$

Заранее исключим из рассмотрения коды с порождающими многочленами вида $(x + 1)^c$. Если длина $n = 2^m - 1$ циклического кода есть простое число, то его порождающий многочлен $g(x)$ есть произведение только примитивных многочленов $m(x)$ степени m, и потому каждый вектор такого кода принадлежит некоторому коду Хэмминга простой длины $n = 2^m - 1$, который порожден многочленом $m(x)$. Это означает, что каждый циклический код длины $n = 2^m - 1$ является подкодом (и не одного!) кода Хэмминга той же длины. Это означает в

свою очередь, что множество всех векторов всех кодов Хэмминга простой длины $n = 2^m - 1$ есть в точности множество всех векторов всех циклических кодов той же длины.

Оценим сверху мощность N этого множества. Если n простое, то имеется в точности $\varphi(n) = n-1$ примитивных элементов поля $GF(2^m)$, а потому — в точности $(n-1)/m$ примитивных многочленов степени m, так как все примитивные элементы поля являются корнями примитивных многочленов, которые делят многочлен $x^{2^m} - x$ и потому имеют степень m. Помня, что размерность кода Хэмминга есть $k = n-m = 2^m-1-m$, легко получить

$$N < \frac{(n-1)2^n}{m(n+1)}.$$

Эта оценка допускает уточнения. Например, в правой ее части нулевой и сплошь единичный векторы сосчитаны по $\frac{n-1}{m}$ раз, так как они принадлежат всем кодам Хэмминга. Поэтому

$$N < \frac{n-1}{m}\left(\frac{2^n}{n+1} - 2\right) + 2.$$

Таким образом, циклическим кодам простой длины $n = 2^m - 1$ не принадлежат по меньшей мере

$$2^n - \frac{n-1}{m}\left(\frac{2^n}{n+1} - 2\right) - 2$$

векторов данной длины.

Разумеется, в это число входят и все векторы веса 1, 2, $n-1$, $n-2$, которым в циклических кодах места нет вообще.

Рассмотрим циклические коды длины $n = 7$. Существует два кода Хэмминга такой длины. Их порождающие многочлены — $x^3 + x^2 + 1$ и $x^3 + x + 1$. Каждый код содержит по семь векторов веса $w = 3$ и по семь векторов веса $w = 4$. Общими для двух этих кодов будут только нулевой и сплошь единичный векторы. Таким образом, все циклические коды длины $n = 7$ содержат в точности 30 кодовых векторов из общего числа $2^7 = 128$ векторов. Конечно, этим кодам не полагается содержать 56 векторов веса 1,2,5,6, что очевидно. Далее, подкодам этих кодов, порождаемым многочленами

$$(x^3 + x^2 + 1)(x + 1),\ (x^3 + x + 1)(x + 1), \tag{5.7.16}$$

с кодовым расстоянием $d = 4$ не могут принадлежать векторы веса 3. Всего векторов веса 3 имеется 35, и 14 из них уже принадлежат двум кодам Хэмминга, а потому не могут считаться непринадлежащими циклическим кодам. Но кодам Хэмминга и их подкодам принадлежат только 14 векторов веса 4. Поэтому 21 векторов веса 4, которые могли бы принадлежать кодам с порождающими многочленами (5.7.16), не принадлежат никаким циклическим кодам.

Покажем, что требование простоты длины n кода существенно. Пусть порождающий многочлен есть

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

Его корнями являются элементы α^5, α^3, где α примитивный элемент поля $GF(2^4)$. Порядки этих корней равны 3 и 5 соответственно. Наименьшее общее кратное порядков есть 15, но код, порождаемый нашим многочленом, не содержится целиком ни в одном коде Хэмминга длины $n = 15$. Например, ни в одном циклическом коде Хэмминга длины 15 не содержится вектор, отвечающий данному порождающему многочлену, так как этот многочлен не делится на примитивные многочлены $(x^4 + x^3 + 1)$ и $(x^4 + x + 1)$.

В данном разделе показано, что двоичные циклические коды длины n отнюдь не используют всего ресурса 2^n векторов.

5.8. Задачи к главе 5

5.1. Показать, что минимальный вес двоичного циклического кода длины n, порожденного многочленом $g(x)$, равен, по крайней мере, 3, если n есть наименьшее число, при котором многочлен $x^n - 1$ делится на $g(x)$.

5.2. Показать, что произвольный циклический код над полем $GF(q)$ вместе с каждым вектором $u = (a_0, a_1, \ldots, a_{n-1})$ содержит вектор $v = (a_{n-1}, a_{n-2}, \ldots, a_0)$, если некоторая степень числа q сравнима с -1по модулю n.

5.3. Пусть многочлен $g(x)$ порождает циклический код длины n над полем $GF(p^m)$, и пусть $n \neq 0(\mathrm{mod}\, p)$. Показать, что вектор, состоящий из одних единиц, принадлежит коду тогда и только тогда, когда $g(x)$ не длится на $x - 1$.

5.4. Циклический код порожден многочленом

$$g(x) = x^8 + x^7 + x^6 + x^4 + 1.$$

Показать, что если α корень многочлена $x^4 + x + 1$, то

1) Вместе с элементами α, α^2, α^4 корнем порождающего многочлена будет и α^3.

2) Длина кода равна 15.

Найти порождающую и проверочную матрицы кода в канонической форме.

5.5. Показать, что двоичный циклический код содержит вектор нечетного веса тогда и только тогда, когда он содержит сплошь единичный вектор.

5.6. Показать, что при $n|q^m - 1$ корни многочлена $x^n - 1$ образуют группу, т.е. подгруппу мультипликативной группы поля $GF(q^m)$.

5.7. Пусть r_z есть наибольшая из кратностей r_i, $i = 1, 2, \ldots, \rho$, с которыми элементы (5.5.12) входят в порождающий многочлен $g(x)$ над $GF(q)$, $q = p^m$ Показать, что длина n циклического кода удовлетворяет равенству $n = n_1 p^s$, где n_1 есть общее наименьшее кратное порядков элементов (5.5.12), и s наименьшее целое число, удовлетворяющее неравенству $r_z \leq p^s$.

5.8. В отличие от канала с независимыми ошибками существуют каналы с группирующимися ошибками. Пачка (пакет) ошибок длины не более, чем l — это такой вектор-ошибка (0.1.3), в котором самый левый отличный от нуля символ e_{i_1} и самый правый отличный от нуля символ e_{i_2} таковы, что $i_1 - i_2 \leq l$. Каков циклический код, обнаруживающий все такие векторы-ошибки?

Глава 6.

Коды Боуза—Чоудхури—Хоквингема

6.1. Важнейший класс циклических кодов

Определение 6.1.1. *Кодом Боуза—Чоудхури—Хоквингема (БЧХ) над GF(q) называется такой циклический код, порождающий многочлен $g(x)$ которого имеет своими корнями последовательность идущих подряд степеней некоторого произвольного элемента $\alpha \in GF(q^m)$*

$$\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+\delta-2}, \tag{6.1.1}$$

где b любое целое число и $\delta \geq 2$. Последовательность может содержать также только один элемент α^b.

Найдем длину n циклического кода, порождающий многочлен которого имеет корни (6.1.1).

Теорема 6.1.2. *Либо длина n равна порядку элемента α^b, если $\delta - 2 = 0$, либо порядку элемента α в противном случае, т.е. когда $\delta > 2$.*

Доказательство. При $\delta - 2 = 0$ утверждение тривиально. Вторая часть следует из теоремы 5.5.1. Докажем, что Н.О.К порядков корней (6.1.1) равно в точности порядку элемента α. Действительно, пусть α порождает циклическую группу G порядка n. Именно к ней принадлежат все корни (6.1.1). Легко показать, что не существует такой истинной подгруппы $G' \subset G$, к которой принадлежали бы все корни (6.1.1). Предположим противное. Пусть все корни (6.1.1) принадлежат

к одной подгруппе $G' \subset G$. Тогда все они являются степенями одного элемента α^h, т.е. имеют вид α^{hj}, где h есть делитель n. Рассмотрим два соседних элемента α^{b+i} и α^{b+i+1}. По предположению $b+i = hj_0$, $b+i+1 = hj_1$. Это значит, что $b+i+1-(b+i) = hj_1 - hj_0 = h(j_1 - j_0) = 1$, что возможно только при $h = 1$ для любых соседних целых чисел $b+i = hj_0$, $b+i+1 = hj_1$. Доказанное означает, что порядки элементов $\alpha^{b+i}, i = 0, 1, \ldots, \delta-2$. не являются делителями порядка никакой истинной подгруппы $G' \subset G$. Это значит, что они являются делителями только порядка n группы G, т.е. делителями порядка элемента α.

Следовательно, n является Н.О.К. порядков элементов $\alpha^{b+i}, i = 0, 1, \ldots, \delta-2$, каково бы ни было δ.

На практике чаще всего $b = 1$, или 0. Иногда, в случае $b = 1$ код называют кодом БЧХ *в узком смысле.*

Заметим, что для всех многочленов $u(x)$, принадлежащих коду, заведомо $u(\alpha^{b+i}) = 0$ для всех $i = 0, 1, \ldots \delta-2$.

Определение 6.1.1 открывает возможность регулярным образом строить проверочную матрицу в соответствии с теоремой 4.5.3. Заметим, связь с этой теоремой определяется тем, что последовательность (6.1.1) содержит по меньшей мере $\delta-1$ членов. Слова "по меньшей мере" связаны с тем, что на основании теоремы 3.5.12 последовательность (6.1.1) может оказаться длиннее.

Теорема 6.1.3. *Минимальное расстояние кода БЧХ с корнями* (6.1.1) *порождающего многочлена равно по меньшей мере* δ.

Д о к а з а т е л ь с т в о. Из сравнения последовательностей (5.5.12) и (6.1.1) проверочная матрица кода БЧХ немедленно получается заменой α_i на α^{b+i-1} в проверочной матрице (5.5.13):

$$H = \begin{bmatrix} 1 & \alpha^b & (\alpha^b)^2 & \ldots & (\alpha^b)^{n-1} \\ 1 & \alpha^{b+1} & (\alpha^{b+1})^2 & \ldots & (\alpha^{b+1})^{n-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha^{b+\delta-2} & (\alpha^{b+\delta-2})^2 & \ldots & (\alpha^{b+\delta-2})^{n-1} \end{bmatrix}. \quad (6.1.2)$$

Взяв из матрицы (6.1.2) произвольные различные $\delta-1$ столб-

цов, составим определитель

$$D = \begin{vmatrix} (\alpha^b)^{j_1} & (\alpha^b)^{j_2} & \dots & (\alpha^b)^{j_{\delta-1}} \\ (\alpha^{b+1})^{j_1} & (\alpha^{b+1})^{j_2} & \dots & (\alpha^{b+1})^{j_{\delta-1}} \\ \dots & \dots & \dots & \dots \\ (\alpha^{b+\delta-2})^{j_1} & (\alpha^{b+\delta-2})^{j_2} & \dots & (\alpha^{b+\delta-2})^{j_{\delta-1}} \end{vmatrix}, \qquad (6.1.3)$$

который преобразуется к виду

$$D = \alpha^{b(j_1+j_2+\dots+j_{d-1})} \begin{vmatrix} 1 & 1 & \dots & 1 \\ \alpha^{j_1} & \alpha^{j_2} & \dots & \alpha^{j_{\delta-1}} \\ \dots & \dots & \dots & \dots \\ (\alpha^{j_1})^{\delta-2} & (\alpha^{j_2})^{\delta-2} & \dots & (\alpha^{j_{\delta-1}})^{\delta-2} \end{vmatrix}. \qquad (6.1.4)$$

Множитель перед знаком определителя получится, если из каждого i-го столбца вынести множитель α^{bj_i}.

Положим $\alpha^{b(j_1+j_2+\dots+j_{\delta-1})} = C$, и пусть ради наглядности $\alpha^{j_i} = a_i$. Тогда

$$D = C \begin{vmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_{\delta-1} \\ \dots & \dots & \dots & \dots \\ a_1^{\delta-2} & a_2^{\delta-2} & \dots & a_{\delta-1}^{\delta-2} \end{vmatrix}. \qquad (6.1.5)$$

В этом равенстве справа легко узнать определитель Вандермонда. Известно, что он отличен от нуля тогда и только тогда, когда все a_i различны и принадлежат области целостности. (В нашем случае они даже элементы поля).

Тем самым доказано, что любые $\delta-1$ столбцов проверочной матрицы 6.1.2 линейно независимы.

Остается применить теорему 4.5.3, и теорема 6.1.3 доказана.

Полезно напомнить значение определителя Вандермонда.

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ a_1 & a_2 & \dots & a_c \\ a_1^2 & a_2^2 & \dots & a_c^2 \\ \dots & \dots & \dots & \dots \\ a_1^{c-1} & a_2^{c-1} & \dots & a_c^{c-1} \end{vmatrix} = \prod_{c \ge i > j \ge 1} (a_i - a_j), \qquad (6.1.6)$$

Это равенство можно доказать методом индукции. Для $c = 2$ оно верно, так как

$$\begin{vmatrix} 1 & 1 \\ a_1 & a_2 \end{vmatrix} = a_2 - a_1. \qquad (6.1.7)$$

Положим, что наше утверждение верно для определителя $(c-1)$-го порядка, и поступим следующим образом. Из последней, c-й строки определителя (6.1.6) вычтем $(c-1)$-ю, умноженную на a_1, затем из $(c-1)$-й строки вычтем $(c-2)$-ю, также умноженную на a_1, и т.д., наконец, из второй строки вычтем первую, умноженную на a_1. Определитель в (6.1.6) примет вид

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & a_2-a_1 & a_3-a_1 & \dots & a_c-a_1 \\ 0 & a_2^2-a_1a_2 & a_3^2-a_1a_3 & \dots & a_c^2-a_1a_c \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_2^{c-1}-a_1a_2^{c-2} & a_3^{c-1}-a_1a_3^{c-2} & \dots & a_c^{c-1}-a_1a_c^{c-2} \end{vmatrix}. \tag{6.1.8}$$

Этот определитель равен своему минору M_{11} порядка $(c-1)$. После вынесения общего множителя a_i-a_1 $(i=2,3,\dots c)$ из всех элементов соответствующего столбца определитель примет вид

$$\prod_{i=2}^{c}(a_i-a_1)\begin{vmatrix} 1 & 1 & \dots & 1 \\ a_2 & a_3 & \dots & a_c \\ a_2^2 & a_3^2 & \dots & a_c^2 \\ \dots & \dots & \dots & \dots \\ a_2^{c-2} & a_3^{c-2} & \dots & a_c^{c-2} \end{vmatrix} =$$

$$= \prod_{i=2}^{c}(a_i-a_1)\prod_{c\ge i>j\ge 2}(a_i-a_j) = \prod_{c\ge i>j\ge 1}(a_i-a_j), \tag{6.1.9}$$

где промежуточное равенство справедливо по предположению индукции.

Иногда пользуются другим доказательством равенства (6.1.6).

Обе его части обращаются в нуль одновременно тогда и только тогда, когда в левой части найдутся, по крайней мере, два одинаковых столбца, т.е. тогда, когда найдутся такие i и j, что $a_i=a_j$. Это означает, что разности a_i-a_j входят множителями в левую часть равенства, и, значит, левая часть делится на правую. Из правила вычисления определителя следует, что все $c!$ слагаемых суммы разложения определителя имеют одну и ту же степень $1+2+\dots+c-1=c(c-1)/2$. Поэтому левая и правая части равенства (6.1.6) могут отличаться только постоянным множителем, который равен 1, так как коэффициент при $1a_2a_3^2\cdots a_c^{c-1}$ в обеих частях один и тот же.

П р и м е р 6. 1.

Циклический $(15, 5)$-код примера 5.4. есть не что иное, как код БЧХ. Задав корнями порождающего многочлена элементы α, α^3, $\alpha^5 \in GF(2^4)$, на самом деле задают последовательность α, α^2, α^3, α^4, α^5, α^6, α^8, α^9, α^{10}, α^{12}, которая автоматически получается из первоначальной добавлением к ней сопряженных элементов. Получилось, что подряд идущих степеней элемента α ровно шесть. Это значит, что $\delta - 1 = 6$, и код заведомо исправляет любую комбинацию из трех и менее независимых ошибок. На самом деле здесь $d = \delta$.

Наконец, в полном соответствии с теоремой 6.1.3 справедливо

Утверждение 6.1.4. *Циклический код Хэмминга есть код БЧХ.*

Действительно, порождающий многочлен циклического кода Хэминга длины $n = 2^m - 1$ есть примитивный многочлен (т.е. минимальная функция) степени m. Корень примитивного многочлена есть примитивный элемент $\alpha \in GF(2^m)$. Вместе с α его корнем будет α^2 (см. теорему 3.5.12). Таким образом, последовательность (6.1.1) есть α, α^2, и так как в данном случае $b = 1$, то $\delta - 2 = 1$, $\delta = 3$. Здесь также $d = \delta$.

6.2. Коды, двойственные кодам Хэмминга

Теперь мы в состоянии привести доказательство теоремы 4.8.2.

Так как циклический код Хэмминга порожден примитивным многочленом $m(x)$ степени m, имеющим своими корнями элементы α, α^2, α^4, $\ldots$, $\alpha^{2^{m-1}}$, то проверочный многочлен $h(x) = (x^{2^m-1} - 1)/m(x)$ имеет своими корнями все остальные ненулевые элементы поля $GF(2^m)$. В частности, его корнями будут элементы, расположенные в максимальном промежутке между корнями многочлена $m(x)$. Этот промежуток начинается элементом $\alpha^{2^{m-1}+1}$, оканчивается элементом $\alpha^{2^m-1} = 1$ и содержит в точности $2^m - 1 - 2^{m-1} = 2^{m-1} - 1$ последовательных степеней элемента α. Иными словами, последовательность (6.1.1) имеет вид

$$\alpha^{2^{m-1}+1}, \alpha^{2^{m-1}+2}, \ldots, \alpha^{2^m-1} = 1.$$

Здесь $b = 2^{m-1} + 1$. Порядок элемента α есть $2^m - 1$, так как α является корнем примитивного многочлена $m(x)$. Отсюда следует, что порождающий многочлен кода, двойственного циклическому коду Хэмминга, есть порождающий многочлен кода БЧХ, для которого $\delta - 1 = 2^{m-1} - 1$, а потому $\delta = 2^{m-1}$. Таким образом, для циклического кода, двойственного циклическому коду Хэмминга, теорема 4.8.2 верна. Но она верна и для нециклических кодов, так как они получаются из циклических простой перестановкой компонент кодовых векторов, и эти перестановки не меняют метрических свойств кода. Теорема доказана полностью.

Вообще, пусть циклический код, порожден примитивным многочленом $m(x)$ степени m с корнями α, α^q, α^{q^2}, $\ldots$, $\alpha^{q^{m-1}} \in GF(q^m)$. Двойственный ему код является кодом БЧХ, так как среди корней его порождающего многочлена $h(x) = (x^{q^m-1} - 1)/m(x)$ находятся элементы

$$\alpha^{q^{m-1}+1}, \alpha^{q^{m-1}+2}, \ldots, \alpha^{q^m-1} = 1.$$

Они представляют собой последовательность (6.1.1), содержащую $q^m - 1 - q^{m-1} = q^{m-1}(q-1) - 1 = \delta - 1$ элементов. Рассмотренный выше двоичный $(2^m - 1,\ m,\ 2^{m-1})$-код является частным случаем только что полученного $(q^m - 1,\ m, q^{m-1}(q-1))$-кода при $q = 2$. Построенные коды называются кодами, состоящими из последовательностей максимальной длины. Они эквидистантны.

6.3. Параметры кодов БЧХ

Число $n - k$ проверочных символов циклического кода равно степени его порождающего многочлена. Она в свою очередь равна числу корней многочлена. Так как число $\delta - 1$ идущих подряд степеней (6.1.1) одного и того же корня не может превышать числа всех корней, то $\delta - 1 \leq n - k$, что тривиально, так как это не более, чем граница Синглтона.

В худшем случае все элементы (6.1.1) принадлежат различным минимальным многочленам, степень каждого из которых не превышает m. Поэтому общее число $n - k$ корней порождающего многочлена $g(x)$, а, значит, и его степень удовлетворяет неравенству $\deg g(x) \leq (\delta - 1)m = 2tm$. Отсюда для числа информационных символов имеет место соотношение

$$k \geq n - 2tm. \tag{6.3.10}$$

На самом деле сопряженные корни имеют один и тот же минимальный многочлен. Если $\beta \in GF(q)$ корень, то и β^q корень. Это значит, что каждый q-й элемент последовательности (6.1.1) не принадлежит новому минимальному многочлену, и максимальное число минимальных многочленов не превышает $(\delta - 1)(1 - 1/q)$. В частности, для двоичных кодов БЧХ проверочных символов оказывается в два раза меньше, и (6.3.10) дает $k \geq n - tm$. В литературе содержатся более тонкие утверждения о возможности незначительного улучшения этой оценки.

Последовательность (6.1.1) определяет так называемое, *гарантированное, или конструктивное,* кодовое расстояние. И дело не только в том, что к последовательности (6.1.1) могут автоматически добавиться сопряженные элементы.

П р и м е р 6. 2.

Рассмотрим поле $GF(2^{11})$. Порядок его мультипликативной группы есть $2^{11} - 1 = 89 \times 23$. Пусть α — примитивный элемент группы, и $\beta = \alpha^{89}$. Порядок элемента β есть 23. Минимальная функция $m(x)$ элемента β имеет своими корнями $\beta, \beta^2, \beta^4, \beta^8, \beta^{16}, \beta^{32} = \beta^9, \beta^{18}, \beta^{36} = \beta^{13}, \beta^{\mathbf{26}} = \beta^{\mathbf{3}}, \beta^6, \beta^{12}, \beta^{24} = \beta$.

Среди этих корней имеются четыре последовательные степени β, β^2, β^3, β^4. Положив $m(x)$ порождающим многочленом, получим (23, 12)-код БЧХ длины $n = 23$ с $n - k = 11$ проверочными символами и гарантированным кодовым расстоянием $\delta = 5$.

Но параметры $n = 23$ и $k = 12$ имеет также код Голея (см. конец раздела 4.6). Это означает, что построенный код БЧХ эквивалентен коду Голея и также имеет расстояние $d = 7$, вопреки теореме 6.1.3. Действительное расстояние d превышает гарантированное кодовое расстояние δ, и этот факт не находит объяснения никаким содержанием последовательности корней порождающего многочлена.

Порождающий многочлен (в данном случае — минимальная функция элемента $\beta = \alpha^{89}$) имеет вид

$$m(x) = x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1. \qquad (6.3.11)$$

Многочлен $x^{11} + x^{10} + x^6 + x^5 + x^4 + x^2 + 1$, двойственный многочлену (6.3.11), порождает код, также эквивалентный коду Голея.

Следует усвоить, что в кодах БЧХ всегда $d \geq \delta$. Именно потому в определении 6.1.1 кодов БЧХ и был введен символ δ,

непривычный для обозначения кодового расстояния, что действительное расстояние d может оказаться больше.

Построение многочлена по его корням было подробно рассмотрено в примере 3.9 раздела 3.5. Имея в своем распоряжении таблицу поля $GF(2^{11})$, (при современной вычислительной технике можно строить поля еще бóльших степеней расширения) процедуру упомянутого примера можно было бы распространить и на случай многочлена (6.3.11). Можно, однако, воспользоваться таблицами неприводимых многочленов, содержащихся, например, в известной книге У.У. Питерсона.[1] В нашем случае в той части таблицы, где указаны неприводимые многочлены как минимальные многочлены элементов поля $GF(2^{11})$, находим строчку с числом 89. Это число как раз указывает на элемент $\beta = \alpha^{89}$, являющийся корнем искомого минимального многочлена. Цифровое содержание строчки 5343. Если заменить каждую цифру ее двоичным эквивалентом 101 011 100 011, то единицы укажут отличные от нуля коэффициенты при соответствующих степенях x.

Вернемся к последним абзацам раздела 5.7. Неприводимые делители (x^2+x+1) и $(x^4+x^3+x^2+x+1)$ рассмотренного там порождающего многочлена имеют своими корнями соответственно: α^5, α^{10} и α^3, α^6, α^{12}, $\alpha^9 \in GF(2^4)$. Каждый из этих многочленов порождал бы весьма неинтересные циклические коды длин 3 и 5. Оба кода содержали бы по два кодовых вектора: (000), (111) и (00000), (11111). Вместе же в составе порождающего многочлена $g(x) = (x^2 + x + 1)(x^4 + x^3 + x^2 + x + 1)$ они порождают код длины 15, и каждая из двух пар корней α^5, α^6 и α^9, α^{10} удовлетворяет определению кода БЧХ с $d-1 = 2$. Однако достижением такой код назвать нельзя: исправляя произвольную одиночную ошибку он имеет 9 информационных символов, в то время, как код Хэмминга той же длины и с той же корректирущей способностью имеет 11 информационных символов.

Заметим далее, что, если $b = 0$, то последовательность (6.1.1) начинается с корня $\alpha^0 = 1$. Минимальная функция этого корня есть $x - 1$. Будучи сомножителем порождающего многочлена, она повышает его степень всего на единицу, на столько же увеличивая и минимальное расстояние δ. Для каждого кодового многочлена $u(x) = u_0 + u_1x + \ldots + u_{n-1}x^{n-1}$ тако-

[1] Часть этих таблиц перепечатана в конце книги в разделе "Неприводимые многочлены"

го кода $u(1) = u_0 + u_1 + \ldots + u_{n-1} = 0$. Это означает, что если характеристика поля, которому принадлежат компоненты u_i кодовых векторов, равна p, то в каждом кодовом многочлене $\sum_i u_i \equiv 0 (\mathrm{mod} p)$. В двоичном коде добавление единицы к списку (6.1.1) означает добавление проверки на четность: все кодовые векторы имеют четный вес.

6.4. Декодирование кодов БЧХ

В этой книге изложены основные методы декодирования, которые будут вводиться в рассмотрение по мере появления новых кодов, к которым эти методы наиболее подходят. Ближайшие разделы этой главы имеют дело с алгоритмом декодирования Горенстейна-Питерсона-Цирлера. Затем в применении к кодам Рида-Соломона будет подробно изучен метод, основанный на алгоритме Эвклида, и, наконец более или менее детально будет представлен алгоритм Берлекемпа, сопровождаемый некоторыми важными подробностями.

Итак, пусть по каналу связи отправлен кодовый вектор

$$u = (u_0, u_1, \ldots, u_{n-1})$$

кода БЧХ, и в канале произошла ошибка, изображаемая вектором

$$e = (e_0, e_1, \ldots, e_{n-1}). \tag{6.4.12}$$

На приёмном конце принят вектор

$$v = (u + e),$$

и декодер вычисляет произведение vH^T, где матрица H есть проверочная матрица (6.1.2). Имеем

$$vH^T = (u + e)H^T = uH^T + eH^T = eH^T,$$

так как

$$uH^T = 0.$$

Последнее равенство следует из того, что кодовый вектор u принадлежит нулевому подпространству матрицы H.

Произведение

$$eH^T = S$$

есть синдром. Его элементами, отвечающими i-й строке матрицы (6.1.2), являются

$$S_{i+1} = e(\alpha^{b+i}) = e_0 + e_1\alpha^{b+i} + e_2(\alpha^{b+i})^2 + \ldots + e_{n-1}(\alpha^{b+i})^{n-1}. \quad i = 0,\ 1,\ \ldots, d-2, \tag{6.4.13}$$

и синдром S есть вектор $S = (S_1,\ S_2,\ \ldots,\ S_{d-1})$.

Начиная с этого раздела, будем исходить из действительного кодового расстояния d.

Таким образом, если в общем случае линейных кодов процедура получения синдрома представляет собой вычисление произведения vH^T, то в случае кодов БЧХ эта процедура есть всего-навсего вычисление значений принятого многочлена $v(x)$ при $x = \alpha^{b+i}$, $(i = 0,\ 1,\ \ldots, d-2)$, что и дает немедленно значение $e(\alpha^{b+i})$, так как заведомо $u(\alpha^{b+i}) = 0$ для всех i. (Напомним, что для нас вектор v и многочлен $v(x)$, коэффициенты которого суть компоненты вектора, это одно и то же).

6.5. Декодирование двоичных кодов с исправлением двух ошибок

Начнем с простейшего случая исправления двух ошибок двоичным кодом БЧХ длины n. Если $t = 2$, то $d = 5$. Пусть без ограничения общности $b = 1$. Последовательность (6.1.1) в данном случае будет α, α^2, α^3, α^4.

Положим, в векторе (6.4.12) отличны от нуля две компоненты $e_{j_1} = 1$ и $e_{j_2} = 1$. Тогда в соответствии с правилом вычисления синдрома S

$$\begin{aligned} &S_1 = \alpha^{j_1} + \alpha^{j_2}, \\ &S_2 = (\alpha^2)^{j_1} + (\alpha^2)^{j_2} = (\alpha^{j_1})^2 + (\alpha^{j_2})^2 = (\alpha^{j_1} + \alpha^{j_2})^2 = S_1^2, \\ &S_3 = (\alpha^3)^{j_1} + (\alpha^3)^{j_2}. \end{aligned} \tag{6.5.14}$$

Положим для удобства $\alpha^{j_1} = X_1$, $\alpha^{j_2} = X_2$ и составим уравнение

$$(X - X_1)(X - X_2) = X^2 + (X_1 + X_2)X + X_1X_2 = 0.$$

Выразим коэффициенты уравнения через элементы синдрома: $X_1 + X_2 = S_1; S_3 = X_1^3 + X_2^3 = (X_1 + X_2)(X_1^2 + X_2^2 + X_1X_2) =$

$S_1(X_1X_2 + S_1^2)$, откуда

$$X_1X_2 = S_1^2 + S_3/S_1.$$

Уравнение примет вид

$$X^2 + S_1X + S_1^2 + S_3/S_1 = 0. \qquad (6.5.15)$$

Известная школьная формула решения квадратного уравнения в данном случае неприменима. Зато всех возможных значений X имеется конечное множество, так как они принадлежат тому же конечному полю $GF(2^m)$, что и корни порождающего многочлена. Это означает, что решение будет найдено, если последовательно подставлять в уравнение все отличные от нуля элементы поля $GF(2^m)$. Те элементы α^{j_1} и α^{j_2}, которые обратят левую часть уравнения в нуль, укажут, что символы принятого вектора v искажены, и их следует заменить на противоположные. Существует и более регулярный метод решения квадратного уравнения, который будет изложен в двух следующих разделах этой главы. Однако прежде, чем подставлять в уравнение все элементы поля $GF(2^m)$, следует провести анализ коэффициентов.

1. Если $S_1 = S_3 = 0$, то принимается решение, что ошибок нет.

2. Если $S_1 \neq 0$, $S_1^3 = S_3$, то принимается решение, что произошла одна ошибка. Справедливость такого суждения следует из того. что свободный член обращается в нуль, и уравнение принимает вид

$$X^2 + S_1X = X(X + S_1) = 0.$$

Один корень, пусть это будет X_1, равен нулю, но ни один элемент α^{j_i} не равен нулю, т.е. говоря формально, ни один элемент α^{j_i} не укажет, где ошибка не произошла . Второй корень будет $X_2 = S_1$, и элемент синдрома S_1 в точности указывает место ошибки α^{j_2}.

3. Если $S_1 \neq 0$, $S_3 \neq 0$, $S_1^3 \neq S_3$, то произошла двойная ошибка.

4. Если $S_1 = 0$, $S_3 \neq 0$, то исключаются случаи одиночной, двойной ошибки или отсутствия ошибок. Действительно, если бы ошибок не было, то необходимо $S_1 = 0$, $S_3 = 0$ одновременно. Если бы была одиночная или двойная ошибка, то заведомо $S_1 \neq 0$. Отсюда следует, что имеет место по крайней мере тройная ошибка. Установить ее конкретный вид невозможно, и это означает отказ от декодирования. Ошибка только обнаружена.

5. Может ли случиться, что $S_1 \neq 0$, $S_3 = 0$? Пусть корни порождающего многочлена принадлежат полю $GF(2^4)$, и $e(x) = x^5 + x^{10}$.

Тогда $S_1 = \alpha^5 + \alpha^{10} = 1$, $S_3 = (\alpha^5)^3 + (\alpha^{10})^3 = 1 + 1 = 0$. Квадратное уравнение будет $X^2 + X + 1 = 0$, его корни α^5, α^{10}. На самом деле этот случай укладывается в случай 3. Действительно, если $S_1 \neq 0$, $S_3 = 0$, то это есть частный случай условия $S_1 \neq 0$, $S_3 \neq S_1$.

П р и м е р 6. 3.

Поле $GF(2^4)$ построено по модулю многочлена $p(x) = x^4 + x^3 + 1$. (Поле изображено на табл.(3.4.12)). Корнями порождающего многочлена кода БЧХ являются α и α^3, причём $p(\alpha) = 0$. Длина кода $n = 15$.

В принятом векторе $v = (000101011001000)$ найти искаженные символы в терминах α_i, исправить ошибку и убедиться, что получившийся после исправления вектор принадлежит коду.

Решение. В многочленной форме принятый вектор имеет вид:

$$v(x) = x^3 + x^5 + x^7 + x^8 + x^{11},$$

Проводя операции в поле $GF(2^4)$, построенном по модулю многочлена $p(x) = x^4 + x^3 + 1$, легко вычислить:

$$S_1 = v(\alpha) = \alpha^3 + \alpha^5 + \alpha^7 + \alpha^8 + \alpha^{11} = \alpha^7,$$
$$S_3 = v(\alpha^3) = \alpha^3 + \alpha^9 + \alpha^{15} + \alpha^6 + \alpha^9 = \alpha^{13}.$$

После подстановки найденных элементов $S_1 = \alpha^7$, $S_3 = \alpha^{13}$ синдрома в уравнение (6.5.15) оно станет таким:

$$x^2 + \alpha^7 x + \alpha^{12} = 0.$$

Последовательной подстановкой элементов мультипликативной группы поля $GF(2^4)$ нетрудно убедиться, что корнями уравнения являются элементы α^4, α^8. Искажены 5-й и 9-й символы слева. Вектор-ошибка $e = (000010001000000)$. Переданный вектор $u = (000111010001000)$ принадлежит коду. Проверка этого факта производится подстановкой в $u(x) = x^3 + x^4 + x^5 + x^7 + x^{11}$ корней α и α^3 порождающего многочлена. Легко убедиться, что $u(\alpha) = \alpha^3 + \alpha^4 + \alpha^5 + \alpha^7 + \alpha^{11} = 0$, $u(\alpha^3) = \alpha^9 + \alpha^{12} + \alpha^{15} + \alpha^6 + \alpha^3 = 0$.

Читатель заметил, что ни в данном примере, ни в общем изложении способа декодирования двоичных кодов БЧХ, исправляющих две ошибки, совершенно не использован корень α^4, который в нашем случае замыкает последовательность (6.1.1).

Конечно, вычисление элемента $S_4 = S(\alpha^4)$ синдрома не составило бы никакого труда, так как $S(\alpha^4) = S^2(\alpha^2))$. Просто, в процессе декодирования эта величина не потребовалась.

6.6. Нормальный базис и след элемента поля

Пусть $\xi, \xi^q, \ldots, \xi^{q^{m-1}}$ есть базис поля $GF(q^m)$ над полем $GF(q)$. Так как для любого элемента поля $GF(q^m)$ существует минимальный многочлен, то быть базисом означает, что корни

$$\xi, \xi^q, \ldots, \xi^{q^{m-1}}$$

этого многочлена линейно независимы. Такой базис называется *нормальным*. Нормальный базис существует для любых q и m, в том числе [2]и для $q = 2$. В данном разделе нас будет интересовать именно случай $q = 2$. Например, в поле $GF(2^4)$ нормальным базисом будет

$$\xi^3,\ \xi^6,\ \xi^{12},\ \xi^9. \tag{6.6.16}$$

В поле $GF(2^4)$ это единственный нормальный базис. Практически нормальный базис может быть найден с помощью таблицы неприводимых многочленов в конце книги. Руководствуясь ею, следует отметить многочлены, корни которых линейно независимы. Корни такого многочлена и составляют нормальный базис.

Любой элемент $a \in GF(2^m)$ может быть представлен в виде

$$a = b_0\xi + b_1\xi^2 + \ldots + b_{m-1}\xi^{2^{m-1}},\ b_i \in GF(2),\ i = 0, 1, \ldots m-1. \tag{6.6.17}$$

Выражение

$$T_r(a) = \sum_{i=0}^{m-1} b_i \tag{6.6.18}$$

[2]Доказательство этого факта здесь опущено. Его можно прочитать, например, в [7].

называется следом элемента a над полем $GF(2)$. Так как $b_i \in GF(2)$, то $T_r(a) = 0$ или 1, и ровно 2^{m-1} элементов $a \in GF(2^m)$ имеет след, равный 0, и столько же элементов a имеют след, равный 1. Действительно, среди 2^m наборов значений b_i половина из них имеет одинаковую чётность числа величин b_i, равных единице.

Вычислим все степени $a^{2^i}, i = 0, 1, \ldots m-1$. Имеем

$$\begin{array}{l} a = b_0\xi + b_1\xi^2 + \ldots + b_{m-1}\xi^{2^{m-1}} \\ a^2 = b_0\xi^2 + b_1\xi^4 + \ldots + b_{m-1}\xi^{2^m} \\ a^4 = b_0\xi^4 + b_1\xi^8 + \ldots + b_{m-1}\xi^{2^{m+1}} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a^{2^{m-1}} = b_0\xi^{2^{m-1}} + b_1\xi^{2^{m-1}} + \ldots + b_{m-1}\xi^{2^{2m-2}} \end{array}$$

Напомним, что $b_i \in GF(2)$, и потому $b_i^2 = b_i$. Имея в виду, что $\xi^{2^m} = \xi$, сложим почленно все полученные равенства.

$$\sum_{i=0}^{m-1} a^{2^i} = b_0 \sum_{i=0}^{m-1} \xi^{2^i} + b_1 \sum_{i=0}^{m-1} \xi^{2^i} + \ldots + b_{m-1} \sum_{i=0}^{m-1} \xi^{2^i} = \sum_{i=0}^{m-1} b_i \tag{6.6.19}$$

Последнее равенство имеет место потому, что $\sum_{i=0}^{m-1} \xi^{2^i} = 1$, так как элементы $\xi^{2^i}, i = 0, 1, \ldots, m-1$, линейно независимы, и, значит, не может быть $\sum_{i=0}^{m-1} \xi^{2^i} = 0$. Из (6.6.19) следует, что одновременно $T_r(a) = \sum_{i=0}^{m-1} b_i$ и $T_r(a) = \sum_{i=0}^{m-1} a^{2^i}$. Так как $(a_1 + a_2)^{2^i} = a_1^{2^i} + a_1^{2^i}$, то $T_r(a_1 + a_2) = T_r(a_1) + T_r(a_2)$. Равенство $\sum_{i=0}^{m-1} \xi^{2^i} = 0$ является достаточным, но не необходимым условием линейной зависимости элементов $\xi, \xi^q, \ldots, \xi^{q^{m-1}}$. Равенство$\sum_{i=0}^{m-1} \xi^{2^i} = 1$ является необходимым, но не достаточным условием линейной независимости элементов $\xi, \xi^q, \ldots, \xi^{q^{m-1}}$. В самом деле, для примера рассмотрим поле $GF(2^6)$, построенное по модулю неприводимого многочлена x^6+x+1. Напомним, что это означает $x^6 = x + 1$. В этом поле рассмотрим набор сопряжённых элементов

$$\xi^{11}, \xi^{22}, \xi^{44}, \xi^{25}, \xi^{50}, \xi^{37}. \tag{6.6.20}$$

Все они являются корнями неприводимого примитивного многочлена $x^6+x^5+x^3+x^2+1$, и их сумма равна коэффициенту

при x^5, т.е. 1. Далее, $\xi^{11}+\xi^{25}=\xi^{11}(1+\xi^{14})=\xi^{11}(1+\xi^7)^2=\xi^{11}(1+\xi^6\xi)^2=\xi^{11}(1+(1+\xi)\xi)^2=\xi^{11}(1+\xi+\xi^2)^2=\xi^{11}(\xi^6+\xi^2)^2=\xi^{11}(\xi^2(1+\xi^4)^2=\xi^{11}\xi^4(1+\xi)^8=\xi^{11}\xi^4(\xi^6)^8=\xi^{11}(\xi^4\xi^{48})=\xi^{63}=1$.

Отсюда $0\cdot\xi^{11}+0\cdot\xi^{25}+\xi^{22}+\xi^{44}+\xi^{50}+\xi^{37}=0$, чем и подтверждается линейная зависимость сопряжённых элементов (6.6.20).

Содержание данного раздела используется при решении квадратного уравнения над полем $GF(2^m)$.

6.7. Квадратное уравнение над $GF(2^m)$

Рассмотрим уравнение

$$X^2+X+a=0,\ a\in GF(2^m) \tag{6.7.21}$$

Возведём левую часть уравнения в последовательные степени 2^i, $i=0,1,\ldots,m-1$, и полученные m уравнений сложим почленно

$$\sum_{i=0}^{m-1}(X^2)^{2^i}+\sum_{i=0}^{m-1}(X)^{2^i}+\sum_{i=0}^{m-1}a^{2^i}=0 \tag{6.7.22}$$

Первая сумма преобразуется к виду

$$\sum_{i=0}^{m-1}(X^2)^{2^i}=\sum_{i=0}^{m-1}X^{2^{i+1}}=\sum_{i=0}^{m-1}(X)^{2^i}, \tag{6.7.23}$$

так как $X^{2^m}=X$. Оказывается, две суммы в (6.7.22) совпадают, и таким образом, если уравнение (6.7.21) имеет решение, то

$$T_r(a)=\sum_{i=0}^{m-1}a^{2^i}=0. \tag{6.7.24}$$

Пусть выполняется (6.7.24). Покажем, что сумма

$$y_1=b_1\xi^2+(b_1+b_2)\xi^4+(b_1+b_2+b_3)\xi^8+\ldots+(b_1+b_2+b_3+\ldots+b_{m-1})\xi^{2^{m-1}} \tag{6.7.25}$$

есть решение уравнения (6.7.21). Для этого возведём её в квадрат:

$$y_1^2 = b_1\xi^4+(b_1+b_2)\xi^8+(b_1+b_2+b_3)\xi^{16}+\ldots+(b_1+b_2+b_3+\ldots+b_{m-1})\xi^{2^m} \tag{6.7.26}$$

и, помня, что $(b_1+b_2+b_3+\ldots+b_{m-1}) = T_r(a)+b_0 = b_0$, $\xi^{2^m} = \xi$, найдём ввиду (6.6.17)

$$y_1^2 + y_1 = a, \tag{6.7.27}$$

что и требовалось. Ясно, что вторым решением уравнения будет $y_2 = y_1 + 1$. Таким образом, доказана

Теорема 6.7.1. *Необходимым и достаточным условием решения уравнения* (6.7.21) *является равенство* $T_r(a) = 0$.

Одновременно с этим решается и вопрос о приводимости или нет над $GF(2^m)$ трёхчлена в левой части (6.7.21).

Рассмотрим вопрос о возможности сведения произвольного квадратного трёхчлена к виду (6.7.21).

Пусть трёхчлен

$$X^2 + bX + c \tag{6.7.28}$$

неприводим. Тогда $b, c \neq 0$.

Произведём замену $X = bX'$. Получим трёхчлен

$$b^2((X')^2 + X' + d),\ d = c/b^2. \tag{6.7.29}$$

Он неприводим, т.е. уравнение

$$b^2((X')^2 + X' + d) = 0 \tag{6.7.30}$$

заведомо не имеет корней в поле $GF(2^m)$, и потому, на основании теоремы 6.7.1 $T_r(d) = 1$. Выберем теперь некоторый фиксированный элемент $a \in GF(2^m)$, $T_r(a) = 1$. По свойству следа $T_r(a + d) = T_r(a) + T_r(d) = 0$.

Используя этот факт и теорему 6.7.1, утверждаем, что в поле $GF(2^m)$ существует элемент ζ со свойством

$$\zeta^2 + \zeta + (a + d) = 0. \tag{6.7.31}$$

Заменим в уравнении (6.7.30) $X^{'}$ на $X^{'}+\zeta$: $b^2((X^{'})^2+\zeta^2+X^{'}+\zeta+d)=0$. Учитывая равенство (6.7.31), получим окончательно

$$b^2((X^{'})^2+X^{'}+a)=0. \qquad (6.7.32)$$

Докажем теперь, что любой приводимый над полем $GF(2^m)$ квадратный многочлен, имеющий различные корни, можно привести к виду (6.7.32), где $T_r(a)=0$. В самом деле, коэффициент b трёхчлена (6.7.28) отличен от нуля, так как он есть сумма двух различных элементов поля. Произведём замену $X=bX^{'}$ и вынесем b^2 за скобку. Получим

$$b^2((X^{'})^2+X^{'}+d),\ d=c/b^2.$$

Этот многочлен приводим, и потому $T_r(d)=0$. Находить корни такого многочлена мы умеем.

Резюмируя проведенные рассуждения, скажем, что любой квадратный трёхчлен можно привести к виду (6.7.32), или, что то же, к виду (6.7.21), но только в случае его приводимости $T_r(a)=0$, и, следовательно уравнение имеет решение, а в случае неприводимости $T_r(a)=1$, и решений нет.

Применим полученный результат к задаче декодирования в примере 6.3. Там найден трёхчлен

$$x^2+\alpha^7x+\alpha^{12}=0,$$

последовательной подстановкой в который всех ненулевых элементов поля $GF(2^4)$ отыскиваются два его корня. Ими оказались α^4 и α^8. Специально подчеркнём, что все операции выполнялись в поле (3.4.12), построенном по модулю неприводимого многочлена x^4+x+1.

Поступим, однако более рационально. Произведём замену $x=\alpha^7x^{'}$ и поделим уравнение на α^{14}. Получим $(x^{'})^2+x^{'}+\alpha^{13}=0$. В поле (3.4.12) $T_r(\alpha^{13})=\alpha^{13}+\alpha^{11}+\alpha^7+\alpha^{14}=0$.

Пользуясь таблицей поля (3.4.12), выразим элемент α^{13} через элементы нормального базиса (6.6.16):

$$\alpha^{13}=0\cdot\xi^3+0\cdot\xi^6+1\cdot\xi^{12}+1\cdot\xi^9.$$

Иначе говоря значениями координат b_i элемента α^{13} в этом базисе будут $b_0=b_1=0$; $b_2=b_3=1$.

Отсюда, подставляя эти координаты в (6.7.25), получим корень $y_1 = (b_1 + b_2)\alpha^{12} = \alpha^{12}$. Корень $y_2 = y_1 + 1 = \alpha$.

Так как $x^{'} = \alpha^8 x$, то $x_1 = \alpha^8$, и $x_2 = \alpha^4$, что и требовалось.

Итак, сначала следует привести трёхчлен (6.7.28) к виду (6.7.32) подстановкой $X = bX^{'}$ и вычислить $T_r(a)$. Если $T_r(a) = 1$, то решений нет. Если $T_r(a) = 0$, то один корень $X_1^{'}$ трёхчлена (6.7.32) находится посредством (6.7.25), где b_i, $i = 0, 1, \ldots, m-1$ являются координатами элемента a в принятом нормальном базисе. Второй корень есть $X_2^{'} = X_1^{'} + 1$. Окончательное решение получается обратным преобразованием $X^{'} = b^{-1}X$.

6.8. Общий случай декодирования двоичных кодов БЧХ

Пусть, как и в разделе 6.5, $b = 1$. Положим, в векторе (6.4.12) равны единице t компонент

$$e_{j_1}, e_{j_2}, \ldots, e_{j_t}.$$

Тогда в соответствии с правилом (6.4.13) вычисления синдрома, полагая $\alpha^{j_i} = X_i\,(i = 1, 2, \ldots, t)$ и $d = 2t + 1$, получим

$$\begin{array}{c} S_1 = X_1 + X_2 + \ldots + X_t, \\ S_2 = X_1^2 + X_2^2 + \ldots + X_t^2, \\ \ldots \\ S_{2t} = X_1^{2t} + X_2^{2t} + \ldots + X_t^{2t}. \end{array} \tag{6.8.33}$$

Величины X_i называются локаторами ошибок.

Заметим, что упрощение вычислений достигается тем, что согласно теореме 3.5.1, $S_{2i} = S_i^2$.

Составим многочлен

$$\sigma(z) = (1-X_1z)(1-X_2z)\cdots(1-X_tz) = \sigma_0+\sigma_1z+\sigma_2z^2+\ldots+\sigma_tz^t, \tag{6.8.34}$$

где

$$\begin{array}{c} \sigma_0 = 1, \\ \sigma_1 = -(X_1 + X_2 + \ldots + X_t), \\ \sigma_2 = X_1X_2 + X_1X_3 + \ldots X_{t-1}X_t, \\ \ldots\ldots\ldots\ldots\ldots \\ \sigma_t = (-1)^t X_1X_2 \ldots X_t. \end{array} \tag{6.8.35}$$

Все функции (6.8.35) называются *элементарными симметрическими*, так как они инвариантны относительно $t!$ перестановок индексов. Многочлен $\sigma(z)$ называется многочленом локаторов ошибок. Его корни являются величинами, обратными локаторам ошибок. Если бы были известны коэффициенты σ_i многочлена локаторов ошибок, то его корни можно было бы найти простой подстановкой всех $2^m - 1$ элементов мультипликативной группы поля $GF(2^m)$.

Однако в нашем распоряжении есть только степенны́е суммы (6.8.33) как элементы $S_l = X_1^l + X_2^l + \ldots + X_t^l$, $(l = 1, 2, \ldots, t)$, синдрома $S = (S_1, S_2, \ldots, S_t)$. Связь между элементарными симметрическими функциями σ и степенными суммами, которые являются также симметрическими, ибо инвариантны относительно всех перестановок индексов, реализуется посредством так называемых тождеств Ньютона, вывод которых следует ниже.

Формальная производная многочлена (6.8.34) есть

$$\sigma'(z) = -\sum_i X_i \prod_{j \neq i} (1 - X_j z).$$

Не ограничивая максимальную степень, найдем отношение

$$\begin{aligned} -\frac{z\sigma'(z)}{\sigma(z)} &= \frac{X_1 z}{1 - X_1 z} + \frac{X_2 z}{1 - X_2 z} + \ldots + \frac{X_t z}{1 - X_t z} = \\ &= X_1 z + (X_1 z)^2 + (X_1 z)^3 + \ldots + \\ &+ X_2 z + (X_2 z)^2 + (X_2 z)^3 + \ldots + \\ &\ldots \\ &+ X_t z + (X_t z)^2 + (X_t z)^3 + \ldots + \ldots = \end{aligned} \tag{6.8.36}$$

(меняя порядок суммирования)

$$= z(X_1 + X_2 + \ldots) + z^2(X_1^2 + X_2^2 + \ldots) + \ldots + z^l(X_1^l + X_2^l + \ldots) + \ldots =$$

$$= \sum_{l=1}^{\infty} S_l z^l.$$

Собирая полученные результаты, получим:

$$z\sigma'(z) = \sigma(z)\left(\sum_{l=1}^{\infty} S_l z^l\right).$$

Иначе говоря,

$$(1+\sigma_1 z+\sigma_2 z^2+\ldots)(S_1 z+S_2 z^2+S_3 z^3+\ldots) = z(\sigma_1+2\sigma_2 z+3\sigma_3 z^2+...).$$

Приравнивая коэффициенты при одинаковых степенях z, получим

$$\begin{array}{l} S_1+\sigma_1=0 \\ S_2+S_1\sigma_1+2\sigma_2=0, \\ S_3+S_2\sigma_1+S_1\sigma_2+3\sigma_3=0, \\ S_4+S_3\sigma_1+S_2\sigma_2+S_1\sigma_3+4\sigma_4=0, \\ S_5+S_4\sigma_1+S_3\sigma_2+S_2\sigma_3+S_1\sigma_4+5\sigma_5=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}$$

Это и есть тождества Ньютона. Беря их через одно и заменяя четные числовые коэффициенты нулями, а нечетные — единицами (так как все операции выполняются в поле характеристики 2), получим систему линейных уравнений:

$$\left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & \ldots & 0 \\ S_2 & S_1 & 1 & 0 & \ldots & 0 \\ S_4 & S_3 & S_2 & S_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ S_{2t-4} & S_{2t-5} & S_{2t-6} & S_{2t-7} & \ldots & S_{t-3} \\ S_{2t-2} & S_{2t-3} & S_{2t-4} & S_{2t-5} & \ldots & S_{t-1} \end{array}\right] \left[\begin{array}{c} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \vdots \\ \sigma_{t-1} \\ \sigma_t \end{array}\right] = \left[\begin{array}{c} S_1 \\ S_3 \\ S_5 \\ \vdots \\ S_{2t-3} \\ S_{2t-1} \end{array}\right]. \tag{6.8.37}$$

В дальнейшем матрицу коэффициентов системы (6.8.37) будем обозначать символом M_t.

Решением системы (6.8.37) является набор коэффициентов σ_1, $\sigma_2, \ldots \sigma_t$ многочлена локаторов ошибок (6.8.34). Что делать с известным многочленом (6.8.34), сказано выше.

Прежде, чем обсуждать разрешимость системы (6.8.37), вернемся к случаю $t = 2$. Система (6.8.37) примет вид

$$\left[\begin{array}{cc} 1 & 0 \\ S_2 & S_1 \end{array}\right] \left[\begin{array}{c} \sigma_1 \\ \sigma_2 \end{array}\right] = \left[\begin{array}{c} S_1 \\ S_3 \end{array}\right]. \tag{6.8.38}$$

Решая эту систему, получим $\sigma_1 = S_1$, $\sigma_2 = S_1^2 + S_3/S_1$. Многочлен локаторов ошибок примет вид

$$1 + S_1 X + (S_1^2 + (S_3/S_1))X^2 = 0. \tag{6.8.39}$$

Читатель заметил обратный порядок следования коэффициентов по сравнению с (6.5.15), и без труда понял, почему это произошло.

Займемся общим случаем $t > 2$ ошибок.

Лемма 6.8.1. *Если произошло не более, чем* $t-2$ *ошибки, то* $\sigma_{t-1} = \sigma_t = 0$.

Доказательство. Из условия леммы следует. что нашлись такие j_1, j_2, что $X_{j_1} = X_{j_2} = 0$. Этого достаточно, чтобы все слагаемые в сумме σ_{t-1} (см. (6.8.35)) обратились в нуль, так как в каждом слагаемом окажется, по крайней мере, один нулевой сомножитель. Равенство $\sigma_t = 0$ тривиально.

Лемма 6.8.2. *Если произошло* $t-2$ *ошибок, то*

$$M_t \begin{bmatrix} 0 \\ 1 \\ \sigma_1 \\ \vdots \\ \sigma_{t-2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6.8.40}$$

Доказательство. Выпишем скалярное произведение f-й строки матрицы M_t на вектор-столбец $[0, 1, \sigma_1, \ldots \sigma_{t-1}]^T$:

$$S_{2(f-1)} \cdot 0 + S_{2(f-1)-1} \cdot 1 + S_{2(f-1)-2} \cdot \sigma_1 + \ldots + S_{2(f-1)-(t-1)} \cdot \sigma_{t-2}.$$

Найдем номер строки, для которой последний коэффициент, т.е. коэффициент при σ_{t-2}, равен 1. Очевидно, это будет тогда, когда $2(f-1) - (t-1) = 0$. Отсюда $f = (t+1)/2$. Это значит, что t нечетно, и такая строка единственна. Выше этой строки в последнем столбце матрицы (6.8.37) стоят нули, и потому все скалярные произведения строк с номерами 1, 2, $\ldots (t+1)/2$ совпадают с левыми частями соответствующих тождеств Ньютона. Следовательно, эти скалярные произведения равны нулю. Ниже этой строки в том же столбце стоят коэффициенты при σ_{t-2}, отличные от нуля, но на них соответствующие им левые части тождеств Ньютона обрываются. Однако "недостающие" члены обращаются в нуль из-за того, что $\sigma_{t-1} = \sigma_t = 0$. Следовательно, и эти скалярные произведения равны нулю.

Найдем теперь номер строки, для которой последний коэффициент, т.е. коэффициент при σ_{t-2}, равен S_1. Очевидно, это будет тогда, когда $2(f-1) - (t-1) = 1$. Отсюда $f = t/2 + 1$.

Это значит, что t четно, и такая строка единственна. Выше этой строки в последнем столбце матрицы (6.8.37) стоят нули, и потому все скалярные произведения строк с номерами $1, 2, \dots t/2+1$ совпадают с левыми частями соответствующих тождеств Ньютона. Ниже этой строки в том же столбце стоят коэффициенты при σ_{t-2}, отличные от нуля, но на них соответствующие им левые части тождеств Ньютона обрываются. Однако "недостающие" члены обращаются в нуль из-за $\sigma_{t-1} = \sigma_t = 0$. Следовательно, и эти скалярные произведения равны нулю. Лемма доказана.

Из этих двух лемм следует

Лемма 6.8.3. *Если произошло не более, чем $t-2$ ошибок, то матрица M_t вырождена. (При этом, разумеется, не все $\sigma_i = 0$, так как в противном случае считается, что ошибок нет, и процедура декодирования теряет смысл).*

Действительно, в условиях леммы справедливо равенство (6.8.40), которое представляет собой однородную систему линейных уравнений. Такая система имеет ненулевое решение (а оно действительно существует, так как не все σ_i равны нулю) тогда и только тогда, когда матрица системы вырождена.

Лемма 6.8.4. *Если произошло t ошибок, т.е. если симметрические функции S_l являются суммами $l-$ых степеней t слагаемых X_j, то матрица M_t невырождена.*

Доказательство. Сначала докажем, что в условиях леммы

$$|M_t| = \prod_{j<i}(X_i - X_j). \tag{6.8.41}$$

Действительно, если нашлись такие j_0 и i_0, что $X_{i_0} = X_{j_0}$, то в системе равенств (6.8.33) для всех $l = 1, 2, \dots, 2t$ будет $X_{i_0}^l = X_{j_0}^l$. Так как операции выполняются в поле характеристики 2, то каждая сумма будет содержать $t-2$ слагаемых. Это равносильно тому, что произошло менее, чем $t-1$ ошибок. Но тогда согласно лемме 6.8.3 $|M_t| = 0$. Следовательно, обращение в нуль любой скобки справа в равенстве (6.8.41) обращает в нуль определитель слева. Это значит, что правая часть делит левую. Любое слагаемое в обеих частях имеет одинаковую степень $t(t-2)/2$, а потому правая и левая части в (6.8.41) отличаются только постоянным множителем.

Покажем, что он равен 1. Легко видеть,что справа в (6.8.41) содержится, например, любой член вида

$$X_{i_1} X_{i_2}^2 \dots X_{i_{t-1}}^{t-1}, \tag{6.8.42}$$

где нижние индексы получаются перестановкой

$$\begin{pmatrix} 1, 2, \dots, t-1 \\ i_1, i_2, \dots, i_{t-1} \end{pmatrix}.$$

Очевидно, любой член вида (6.8.42) содержится в (6.8.41) слева в произведении

$$S_1 S_2 \dots S_{t-1} \tag{6.8.43}$$

и только в нем. Одно такое произведение доставляется элементами главной диагонали определителя $\mid M_t \mid$ в его разложении по элементам первой строки.

Легко показать, что никакое другое из $(t-1)!$ слагаемых минора M_{11} определителя $\mid M_t \mid$ не равно произведению (6.8.43). Действительно, пусть в разложении определителя $\mid M_t \mid$ по элементам первой строки общий вид члена минора M_{11} есть

$$a_{1,\alpha_1} a_{2,\alpha_2}, \dots, a_{t-1,\alpha_{t-1}}, \tag{6.8.44}$$

где последовательность

$$\alpha_1, \alpha_2, \dots \alpha_{t-1} \tag{6.8.45}$$

получается перестановкой

$$\begin{pmatrix} 1, 2, \dots, t-1 \\ \alpha_1, \alpha_2, \dots, \alpha_{t-1} \end{pmatrix}.$$

Найдем последовательность (6.8.45), для которой (6.8.44) совпадает с (6.8.43).

Заметим, что элементам S_j синдрома, согласно строению матрицы M_t, равны следующие элементы:

$$a_{k,2k-j}, \ (k = 1, \dots, t-1) \tag{6.8.46}$$

минора M_{11}. Так как нулевые члены нас не интересуют, есть только две возможности: $\alpha_1 = 1$ и 2. Последняя отпадает, так как в противном случае для $t-1$ сомножителей в (6.8.43)

осталось бы $t-2$ строки, что противоречит правилу вычисления определителя. Таким образом, $\alpha_1 = 1$, и $S_1 = a_{11}$. Пусть уже найдено, что $\alpha_2 = 2, \ldots, \alpha_k = k$, т.е. сомножители S_j $(j = 1, 2, \ldots, k)$ в (6.8.43) расположены на главной диагонали минора M_{11}.

Найдем элемент $a_{k+1,\alpha_{k+1}}$. Заведомо $\alpha_{k+1} < k+1$, так как в противном случае, согласно (6.8.46) $2(k+1) - j > k+1$, и $j < k+1$. Но все S_j с такими j уже вошли в произведение (6.8.43)), куда они взяты из столбцов с номерами $1, 2, \ldots, k$. Поэтому элемент S_{k+1} из этих столбцов взят быть не может. Следовательно, $\alpha_{k+1} > k$. Отсюда $\alpha_{k+1} = k+1$. Это означает, что в (6.8.43) вошли только сомножители, расположенные на главной диагонали, т.е. что произведение (6.8.43) может быть построено единственным образом. Это означает, следовательно, что при раскрытии определителя $\mid M_t \mid$ каждое произведение (6.8.42) получается в точности по одному разу и не уничтожается приведением подобных членов. Этим завершается доказательство.

Замечание. Хотя $S_{2j} = S_j S_j$, однако полагая оба экземпляра $S(j)$ различными элементами в $\mid M_t \mid$, нельзя заменить в (6.8.43) один множитель S_{2j} двумя множителями S_j, так как в (6.8.43) должно быть в точности $t-1$ сомножителей. Таким образом, если произошло t ошибок, то матрица M_t невырождена, и система(6.8.37) имеет единственное решение.

С другой стороны, имеет место

Лемма 6.8.5. *Если матрица M_t вырождена, то произошло не более, чем $t-2$ ошибок.*

Действительно, в условиях леммы, по крайней мере, одна скобка в (6.8.41) обращается в нуль. Значит, $X_j = X_i$. Но все ненулевые X_j различны. Значит $X_j = X_i = 0$, и лемма доказана.

Наконец, имеет место

Лемма 6.8.6. *Если произошло $t-1$ ошибок, то матрица M_t остается невырожденной.*

Действительно, в условиях леммы имеем, что для одного из локаторов $X_j = 0$. Заведомо $\sigma_t = 0$, но правая часть в (6.8.41) в нуль не обращается, так как не обращается в нуль ни одна скобка.

Собирая пять последних лемм, получаем как окончательный результат данного раздела, что справедлива

Теорема 6.8.7. *Матрица M_t невырождена, и система* (6.8.37) *имеет единственное решение тогда и только тогда, когда произошло t или $t-1$ ошибок. Матрица M_t вырождена тогда и только тогда, когда произошло менее, чем $t-1$ ошибок.*

Отсюда вытекает такая последовательность действий при декодировании.

1. Зная последовательность (6.1.1) корней порождающего многочлена кода БЧХ, подставляем ее элементы с нечетными показателями степеней в принятый вектор v, в результате чего получаются степенные суммы S_{2i-1}. Степенные суммы с четными индексами получаются возведением в квадрат уже вычисленных. (Напомним, что $S_{2i} = S_i^2$.)

2. Вычисляется определитель $|M_t|$. Если он не равен нулю, то решается система (6.8.37) линейных уравнений относительно σ_i.

3.Составляется многочлен локаторов ошибок. Последовательной подстановкой в него всех ненулевых злементов поля $GF(2^m)$ получаются корни многочлена, как величины, обратные локаторам ошибок.

4. Компоненты вектора v, отвечающие локаторам, заменяются на противоположные.

5. Если $|M_t| = 0$, то это означает, что произошло менее, чем $t-1$ ошибок.

6. В матрице M_t удаляются два последних столбца и две последних строки.

Процесс повторяется i раз до тех пор, пока матрица M_{t-2i} не станет невырожденной. Решается система $t-2i$ линейных уравнений.

П р и м е р 6. 4.

Пусть передавался вектор

$$u = (110100011000100)$$

кода БЧХ длины $n = 15$. Этот код рассмотрен в примере 6.3. Порождающий многочлен кода имеет своими корнями элементы

$$\alpha,\ \alpha^3,\ \alpha^5 \in GF(2^4),$$

а вместе с сопряженными элементами —

$$\alpha,\ \alpha^2,\ \alpha^3,\ \alpha^4,\ \alpha^5,\ \alpha^6, \quad d = 7,\ t = 3.$$

Поле построено по модулю многочлена $1 + x + x^4$, и операции выполняются по правилам таблицы (3.4.11).

1. Произошло 3 ошибки.
Принят вектор

$$v_1 = (011101101000100).$$

Пользуясь таблицей (3.4.11), находим:

$$\begin{array}{l} S_1 = \alpha + \alpha^2 + \alpha^3 + \alpha^5 + \alpha^6 + \alpha^8 + \alpha^{12} = \alpha^{11}, \\ S_3 = \alpha^3 + \alpha^6 + \alpha^9 + 1 + \alpha^3 + \alpha^9 + \alpha^6 = 1, \\ S_5 = \alpha^5 + \alpha^{10} + 1 + \alpha^{10} + 1 + \alpha^{10} + 1 = 0, \\ S_2 = S_1^2 = \alpha^7,\ S_4 = S_2^2 = \alpha^{14},\ S_6 = 1. \end{array}$$

Далее

$$M_3 = \begin{bmatrix} 1 & 0 & 0 \\ \alpha^7 & \alpha^{11} & 1 \\ \alpha^{14} & 1 & \alpha^7 \end{bmatrix},$$

$$|M_3| = \alpha^{18} + 1 = \alpha^3 + 1 = \alpha^{14} \neq 0.$$

Система линейных уравнений относительно σ_i :

$$\begin{array}{l} \sigma_1 = S_1 = \alpha^{11}, \\ \alpha^7\sigma_1 + \alpha^{11}\sigma_2 + \sigma_3 = S_3 = 1, \\ \alpha^{14}\sigma_1 + \sigma_2 + \alpha^7\sigma_3 = S_5 = 0. \end{array}$$

Получим

$$\sigma_1 = \alpha^{11},\ \sigma_2 = \alpha^8,\ \sigma_3 = \alpha^9;$$

Многочлен локаторов ошибок:

$$\sigma(z) = 1 + \alpha^{11}z + \alpha^8 z^2 + \alpha^9 z^3.$$

Его корни как величины, обратные локаторам ошибок:

$$z_1 = \alpha^8,\ z_2 = \alpha^{13},\ z_3 = \alpha^0,$$

Локаторы: α^7, α^2, $\alpha^0 = 1$, что подтверждается сравнением векторов u и v.

2. Произошло 2 ошибки.
Принят вектор

$$v_2 = (111101101000100).$$

Снова пользуясь таблицей (3.4.11), находим:

$$S_1 = 1 + \alpha + \alpha^2 + \alpha^3 + \alpha^5 + \alpha^6 + \alpha^8 + \alpha^{12} = \alpha^{12},$$
$$S_3 = 1 + \alpha^3 + \alpha^6 + \alpha^9 + 1 + \alpha^3 + \alpha^9 + \alpha^6 = 0,$$
$$S_5 = 1 + \alpha^5 + \alpha^{10} + 1 + \alpha^{10} + 1 + \alpha^{10} + 1 = 1,$$
$$S_2 = S_1^2 = \alpha^9,\ S_4 = S_2^2 = \alpha^3,\ S_6 = 0.$$

Далее

$$M_3 = \begin{bmatrix} 1 & 0 & 0 \\ \alpha^9 & \alpha^{12} & 1 \\ \alpha^3 & 0 & \alpha^9 \end{bmatrix},$$

$$|M_3| = \alpha^6 \neq 0.$$

Система линейных уравнений относительно σ_i :

$$\sigma_1 = S_1 = \alpha^{12},$$
$$\alpha^9\sigma_1 + \alpha^{12}\sigma_2 + \sigma_3 = S_3 = 0$$
$$\alpha^3\sigma_1 + \alpha^9\sigma_3 = S_5 = 1$$

Получим:

$$\sigma_1 = \alpha^{12},\ \sigma_2 = \alpha^9,\ \sigma_3 = 0;$$

Многочлен локаторов ошибок:

$$\sigma(z) = 1 + \alpha^{12}z + \alpha^9 z^2.$$

Его корни как величины, обратные локаторам ошибок:

$$z_1 = \alpha^8,\ z_2 = \alpha^{13}.$$

Локаторы: α^7, α^2.

3. Произошла одна ошибка.

Принят вектор

$$v_3 = (110101101000100).$$

Пользуясь таблицей (3.4.11), находим:

$$S_1 = 1 + \alpha + \alpha^3 + \alpha^5 + \alpha^6 + \alpha^8 + \alpha^{12} = \alpha^7,$$
$$S_3 = 1 + \alpha^3 + \alpha^9 + 1 + \alpha^3 + \alpha^9\alpha^6 = \alpha^6,$$
$$S_5 = 1 + \alpha^5 + 1 + \alpha^{10} + 1 + \alpha^{10} + 1 = 1\alpha^5$$
$$S_2 = S_1^2 = \alpha^{14},\ S_4 = S_{14}^2 = \alpha^{13},\ S_6 = \alpha^{12}.$$

Далее

$$M_3 = \begin{bmatrix} 1 & 0 & 0 \\ \alpha^{14} & \alpha^7 & 1 \\ \alpha^3 & \alpha^6 & \alpha^{14} \end{bmatrix},$$

$$|M_3| = 0, |M_1| = |1| = 1.$$

Получим

$$\sigma_1 = S_1 = \alpha^7,\ \sigma_2 = 0,\ \sigma_3 = 0;$$

Многочлен локаторов ошибок:

$$\sigma(z) = 1 + \alpha^7 z.$$

Его корень

$$z_1 = \alpha^8.$$

Локатор: α^7.

В примерах 6.3. и 6.4. специально указывалось, по модулю какого многочлена построено поле. В примере 6.3. это был многочлен $1 + x^3 + x^4$, в примере 6.4. — многочлен $1 + x + x^4$. Необходимость такого указания можно подтвердить примером, когда отсутствие указания приводит к неверному декодированию. Рассмотрим циклический код Хэмминга длины $n = 15$ с порождающим многочленом $1 + x + x^4$, корнями которого являются элементы α, α^2, α^4, α^8. Этот код исправляет любую одиночную ошибку, или обнаруживает любую двойную.

Пусть передан вектор

$$u = (110010000000000),$$

и принят вектор

$$v = (100110000000000).$$

Произошли две ошибки, которые заведомо обнаруживаются. Если будет указано поле (3.4.11), то подстановка в вектор v элемента α по правилам этого поля даст $S_1 = 1 + \alpha^3 + \alpha^4 = \alpha^9 \neq 0$. На исправление двойной ошибки рассчитывать нельзя, но обнаружена она заведомо будет.

Если же подстановка α в v будет выполняться по правилам поля (3.4.12), то тривиальным образом $S_1 = 1 + \alpha^3 + \alpha^4 = 0$, и ошибка не обнаруживается.

В заключение раздела сделаем следующее замечание. Из принадлежности к классу кодов БЧХ циклического кода, двойственного циклическому коду Хэмминга, вовсе не следует, что и декодировать его следует посредством изложенной в данном разделе процедуры. Имеют ли дело с кодом, двойственным коду Хэмминга (циклическим или нет), с укороченным ли РМ-кодом первого порядка, все эти коды эквивалентны, и наилучший для них способ — мажоритарное декодирование.

6.9. Общий случай декодирования q-ичных кодов БЧХ

Основное отличие декодирования кодов над $GF(q)$ от случая двоичных кодов состоит в том, что кроме положения ошибочных символов, т.е. локаторов ошибок, необходимо установить еще и *значения* ошибок. Каждый символ кодового вектора может принимать q значений из поля $GF(q)$. Каждое значение ошибки может быть только ненулевое, ибо нулевое значение ошибки есть ее отсутствие. Как и в двоичном случае, положение ошибки изображается элементом мультиплкативной группы расширения степени m поля $GF(q)$, т.е. поля $GF(q^m)$.

Таким образом,

$$e(x^{b+i}) = e_0 + e_1 x^{b+i} + e_2 (x^{b+i})^2 + \ldots + e_{n-1}(x^{b+i})^{n-1}; \quad (6.9.47)$$

$i = 0, 1, \ldots, d-2$, где $e_i \in GF(q)$. Будем считать, что отличных от нуля слагаемых имеется не более, чем t, и пусть их коэффициенты таковы

$$e_{j_1}, e_{j_2}, \ldots, e_{j_t} \in GF(q).$$

Иначе говоря, положив для простоты и без потери общности $b = 1$, получим

$$e(x) = e_{i_1} x^{i_1} + e_{i_2} x^{i_2} + \ldots + e_{i_t} x^{i_t}.$$

На деле может оказаться, что произошло менее, чем t ошибок. В процессе декодирования это обстоятельство обнаружится и будет учтено автоматически.

Теперь, помня, что элементы (6.4.13) синдрома имеют вид

$$S_{i+1} = e(\alpha^i),$$

и полагая

$e_{j_i} = Y_i$, $\alpha^{j_i} = X_i$, получим (ср. с (6.8.33)):

$$\begin{gathered} S_1 = Y_1X_1 + Y_2X_2 + \ldots + Y_tX_t, \\ S_2 = Y_1X_1^2 + Y_2X_2^2 + \ldots + Y_tX_t^2, \\ \ldots \\ S_{2t} = Y_1X_1^{2t} + Y_2X_2^{2t} + \ldots + Y_tX_t^{2t}, \\ Y_i \in GF(q), \quad X(i) \in GF(q^m). \end{gathered} \tag{6.9.48}$$

Величины X_i, как и прежде, называются локаторами ошибок.

Вычисления упрощаются, если принять во внимания, что согласно теореме 3.5.1,

$$S_{iq} = Y_1X_1^{iq}+Y_2X_2^{iq}+\ldots+Y_tX_t^{iq} = (Y_1X_1^i+Y_2X_2^i+\ldots+Y_tX_t^i)^q = S_i^q.$$

Система (6.9.48) — это система линейных уравнений относительно Y_i, $i = 1, 2, \ldots, t$. Если произошло t ошибок, то все X_i, $i = 1, 2, \ldots, t$, различны и отличны от нуля. В этих условиях определитель первых t уравнений системы, как будет показано ниже, отличен от нуля. Следовательно, первые t уравнений системы (6.9.48) линейно независимы, и они разрешимы относительно неизвестных Y_i, $i = 1, 2, \ldots, t$.

Обратимся к нахождению локаторов ошибок.

Воспользуемся многочленом (6.8.34):

$$\sigma(z) = \sigma_0 + \sigma_1 z + \sigma_2 z^2 + \ldots + \sigma_t z^t, \tag{6.9.49}$$

который, как и в (6.8.34) получается раскрытием скобок и приведением подобных членов в произведении

$$(1 - X_1z)(1 - X_2z)\cdots(1 - X_tz).$$

Напомним, что коэффициенты σ_i — это элементарные симметрические функции (6.8.35). Корни многочлена (6.9.49) — это величины X_i^{-1}, обратные локаторам ошибок. Если бы величины σ_i были известны, то подстановкой в многочлен σ_i последовательно ненулевых элементов поля $GF(q^m)$ можно было бы найти все корни многочлена как величины, обратные локаторам ошибок. Однако они неизвестны. Поступим следующим образом:

Подставим в (6.9.49) $z = X_i^{-1}$.

$$\sigma(X_i^{-1}) = 1 + \sigma_1X_i^{-1} + \sigma_2X_i^{-2} + \ldots + \sigma_tX_i^{-t} = 0, \tag{6.9.50}$$

Затем помножим тождество (6.9.50) на $Y_i X_i^{j+t}$:

$$Y_i(X_i^{j+t} + \sigma_1 X_i^{j+t-1} + \sigma_2 X_i^{j+t-2} + \ldots + \sigma_t X_i^j) = 0. \quad (6.9.51)$$

При фиксированном j просуммируем все тождества (6.9.51) по $i = 1, 2, \ldots, t$:

$$\sum_{i=1}^{t} Y_i(X_i^{j+t} + \sigma_1 X_i^{j+t-1} + \sigma_2 X_i^{j+t-2} + \ldots + \sigma_t X_i^j) = 0.$$

Раскрывая скобки и меняя порядок суммирования, получим

$$\sum_{i=1}^{t} Y_i X_i^{j+t} + \sigma_1 \sum_{i=1}^{t} Y_i X_i^{j+t-1} + \sigma_2 \sum_{i=1}^{t} Y_i X_i^{j+t-2} + \ldots + \sigma_t \sum_{i=1}^{t} Y_i X_i^j = 0.$$

Но

$$\sum_{i=1}^{t} Y_i X_i^{j+t-l} = S_{j+t-l},\ l = 0, 1, \ldots, t.$$

Поэтому для фиксированного j имеем:

$$S_{j+t} + \sigma_1 S_{j+t-1} + \ldots + \sigma_t S_j = 0,\ j = 1, 2, \ldots, t.$$

Таких уравнений будет t штук, и вместе они составляют систему уравнений

$$\begin{array}{l} \sigma_1 S_t + \sigma_2 S_{t-1} + \ldots + \sigma_t S_1 = -S_{t+1}, \\ \sigma_1 S_{t+1} + \sigma_2 S_t + \ldots + \sigma_t S_2 = -S_{t+2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sigma_1 S_{2t-1} + \sigma_2 S_{2t-2} + \ldots + \sigma_t S_t = -S_{2t} \end{array} \quad (6.9.52)$$

или в матричном виде

$$\begin{bmatrix} S_t & S_{t-1} & \ldots & S_2 & S_1 \\ S_{t+1} & S_t & \ldots & S_3 & S_2 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ S_{2t-1} & S_{2t-2} & \ldots & S_{t+1} & S_t \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_t \end{bmatrix} = \begin{bmatrix} -S_{t+1} \\ -S_{t+2} \\ \vdots \\ -S_{2t} \end{bmatrix}. \quad (6.9.53)$$

Обозначим матрицу системы символом M_t.

Теорема 6.9.1. *Система уравнений* 6.9.52 *имеет единственное решение тогда и только тогда, когда произошло* t *ошибок.*

Доказательство. Пользуясь системой (6.9.48), в которой элементы S_l синдрома представлены в виде степенных сумм, нетрудно проверить рутинными выкладками, что $M_t = VLV^T$, где

$$L = \begin{bmatrix} Y_1X_1 & 0 & 0 & \dots & 0 \\ 0 & Y_2X_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & Y_tX_t \end{bmatrix}, \tag{6.9.54}$$

$$V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_t \\ \vdots & \vdots & \vdots & \vdots \\ X_1^{t-1} & X_2^{t-1} & \dots & X_t^{t-1} \end{bmatrix}. \tag{6.9.55}$$

Если произошло t ошибок, то все Y_i и X_i отличны от нуля, определитель матрицы L, равный произведению элементов главной диагонали, отличен от нуля. При тех же условиях, поскольку отличные от нуля величины X_i заведомо различны, определители матриц V и V^T, отличны от нуля, как определители Вандермонда.

Таким образом, матрица M_t невырождена, и система (6.9.52) имеет единственное решение.

Если произошло менее, чем t ошибок, то хотя бы один элемент главной диагонали матрицы L равен нулю. В таком случае определитель матрицы L равен нулю, и, значит, система (6.9.52) не имеет решения. Теорема доказана.

Найдя из системы (6.9.52), если это возможно, все величины $\sigma_1, \sigma_2, \dots, \sigma_t$, (напомним, что всегда $\sigma_0 = 1$), следует составить многочлен локаторов ошибок. Читатель уже знает, что нужно с ним делать. Подстановкой в него всех ненулевых элементов поля $GF(q^m)$ получаются величины, обратные локаторам ошибок, а с ними и сами локаторы.

Теперь вернемся к системе (6.9.48). Ее определитель таков:

$$\begin{vmatrix} X_1 & X_2 & \dots & X_t \\ X_1^2 & X_2^2 & \dots & X_t^2 \\ \vdots & \vdots & \vdots & \vdots \\ X_1^t & X_2^t & \dots & X_t^t \end{vmatrix} = X_1X_2\dots X_t \begin{vmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_t \\ \vdots & \vdots & \vdots & \vdots \\ X_1^{t-1} & X_2^{t-1} & \dots & X_t^{t-1} \end{vmatrix}.$$

Справа находится определитель Вандермонда. При наличии t ошибок локаторы X_i различны, и определитель отличен от нуля. Система (6.9.48) имеет единственное решение.

Окончательно процедура декодирования выглядит следующим образом.

1. Подставляя в принятый вектор v корни порождающего многочлена, вычисляют элементы S_i синдрома. Если все они равны нулю, то считается, что ошибок нет, и процедура окончена.

2. В противном случае из элементов синдрома составляется система уравнений (6.9.52). Если ее матрица M_t в формуле (6.9.53) не вырождена, вычисляются коэффициенты σ_1, $\sigma_2, \ldots \sigma_t$ многочлена локаторов ошибок.

3. Отыскиваются t корней многочлена локаторов ошибок последовательной подстановкой в него элементов поля $GF(q^m)$. Величины, обратные корням, есть локаторы ошибок.

4. Составляется система (6.9.48) уравнений, которая имеет единственное решение, так как ее матрица не вырождена. Решением системы являются значения ошибок.

5. В соответствии с каждой ненулевой (!) парой X_i, Y_i из i-го символа вектора v вычитается величина Y_i. Восстановлен передававшийся вектор u. Процедура закончена.

6. Если матрица M_t в (6.9.53) вырождена, система не разрешима. Это означает, что произошло не более, чем $t-1$ ошибок. Из матрицы M_t в (6.9.53) следует удалить последние строку и столбец, положить $\sigma_t = 0$, а из системы (6.9.52) последнее уравнение. Вся процедура выполняется снова после замены t на $t-1$.

Сравнивая только что изложенную процедуру с процедурой декодирования двоичных кодов БЧХ, видим, что в обоих случаях центральным пунктом является отыскание коэффициентов σ_1, $\sigma_2, \ldots \sigma_t$ многочлена локаторов ошибок. Всякий раз, когда основная матрица системы уравнений относительно этих коэффициентов оказывается вырожденной, делается вывод, что в действительности произошло меньшее число ошибок, чем предполагалось вначале. Порядок системы уравнений снижается: в двоичном случае — на две единицы, в недвоичном случае — на одну единицу. Таким образом, "скорость" продвижения к разрешимости системы уравнений в двоичном случае в два раза выше.

К тому же, в недвоичном случае следует находить еще и значения ошибок, кроме их расположения.

Разумеется, при желании можно не рассматривать отдельно способ декодирования двоичных кодов, так как они есть про-

сто частный случай q-ичных кодов.

Изложенный в разделах 6.9 и 6.8, алгоритмом Питерсона — Горенстейна — Цирлера по мнению большинства исследователей лучше всех остальных способов декодирования проясняет алгебраическую идею процесса.

Другие способы декодирования, например, итеративный алгоритм Берлекемпа, будучи несколько более экономичными, оказываются менее прозрачными. Иные соображения на этот счет можно прочитать в литературе по алгебраическим методам кодирования. Как упомянуто выше в связи с примерами декодирования кодов Рида-Соломона, которые являются частным случаем кодов БЧХ, несколько последних разделов следующей главы, посвящены декодированию посредством алгоритма Эвклида. А Выбор на практике того или иного метода декодирования, как и, вообще, выбор той или иной системы, будь то передача информации или постройка дома, — это всегда решение оптимизационной задачи по многим параметрам.

6.10. Коды БЧХ и исправление стираний

В разделе 0.3 рассмотрен тривиальный метод исправления стираний способом замены стертых символов всевозможными двоичными комбинациями. Метод декодирования кодов БЧХ с минимальным расстоянием d дает возможность произвести только одну такую замену, а затем исправить любые $d-1$ или менее стираний посредством решения системы линейных уравнений. Локаторы стираний X_1, X_2, $\ldots$, X_ν известны заранее. Достаточно подставить на все ν стертых позиций, например, нули и для полученного таким образом вектора вычислить все степенные суммы, т.е. элементы синдрома, S_1, S_1, $\ldots$, $S_{2\nu}$. Тогда истинные значения стертых символов получатся как решения Y_1, $Y_2, \ldots$ Y_ν системы (6.9.48) линейных уравнений.

В случае двоичных кодов возникает некоторая особенность. Она состоит в том, что система (6.9.48) рассматривалась в связи с недвоичными кодами. Конечно, система (6.9.48) годится и в двоичном случае. Но замена стёртых символов нулями в двоичном случае даёт возможность забыть про известные локаторы ошибок, найти их заново для вектора, полученного заменой стёртых символов нулями (ибо эта замена есть ни что иное , как искусственное внесение ошибок), и тогда уже некоторые нули заменить единицами без обращения к системе (6.9.48).Такой образ мыслей может показаться абсурдным, тем не менее иметь его в виду не помешает.

Можно исправлять не только стирания и не только ошибки, но пойти дальше и исправлять ошибки и стирания. Частично эта тема рассмотрена в разделе 0.3. Во всех подробностях она представлена в следующей главе, в изложении проблемы декодирования посредством алгоритма Эвклида.

6.11. Задачи к главе 6

6.1. Построить двоичный код БЧХ размерности 12 с гарантированным расстоянием $\delta = 5$.
6.2. Определить размерность БЧХ-кода над полем $GF(3)$, исправляющего 5 ошибок и имеющего длину 80.
6.3. Показать, что любой двоичный код, порожденный многочленом $g(x) = g^*(x)$, имеющим 1 среди своих корней, и который обладает корнем порядка самое меньшее 5, есть БЧХ-код с минимальным расстоянием 6. Показать, что утверждение применимо к любому из двух двоичных циклических (17, 8)-кодов.
6.4. Показать, что если в определении БЧХ-кодов элемент α заменить на другой элемент α' того же порядка, то получится эквивалентный код.
6.5. Как соотносятся коды БЧХ, у которых при одинаковых b имеет место неравенство $\delta_1 \geq \delta + 2$?
6.6. Всегда ли двойственные коды одновременно являются или не являются кодами БЧХ?

Глава 7.

Коды МДР

7.1. Коды на границе Синглтона

Коды МДР — это коды с максимально достижимым кодовым расстоянием. Строго говоря, с такими кодами мы уже встречались. Например, в полном смысле этого слова, кодами с максимально достижимым кодовым расстоянием можно назвать все совершенные коды. Действительно, например, при тех значениях параметров n и $k = n - \log_2(n+1) = 2^m - 1 - m$, которыми обладают коды Хэмминга, максимальным расстоянием будет $d = 3$, и оно достигается. Но называть таким именем принято не коды Хэмминга, лежащие на неасимптотической границе Хэмминга, а совсем другие коды.

Определение 7.1.1. *Код МДР — это линейный код, для минимального расстояния которого выполняется соотношение* $d = n - k + 1$.

Иначе говоря, это код, который лежит на границе Синглтона. Коды МДР существуют, но пока, не обращаясь к конкретному воплощению этих кодов в реальных объектах, займемся их свойствами, вытекающими только из определения

Определение 7.1.2. *Информационной совокупностью линейного* (n, k)*-кода над GF(q) называется множество номеров компонент кодового вектора, в которых все* q^k *кодовых векторов различны.*

П р и м е р 7.1.

Рассмотрим порождающую матрицу кода Хэмминга

$$G = \left[\begin{array}{cccc:ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right].$$

Легко видеть, что номера 1, 2, 3, 4 составляют информационную совокупность, так как эти части строк матрицы линейно независимы, и все 16 векторов, порождаемых данной матрицей, в первых четырех компонентах различны. С другой стороны, номера 1, 4, 5, 6 не составляют информационную совокупность, так как 2-я и 3-я строки матрицы в компонентах с этими номерами совпадают.

Теорема 7.1.3. *Линейный код лежит на границе Синглтона тогда и только тогда, когда любая совокупность k номеров его компонент является информационной.*

Доказательство. *Необходимость.* Пусть минимальное расстояние линейного кода удовлетворяет равенству $d = n - k + 1$, но пусть при этом некоторая совокупность k номеров компонент кодового вектора не является информационной. Тогда найдутся два кодовых вектора u и v, которые в этих k компонентах совпадают. Значит, для расстояния $d(u, v)$ между ними выполняется неравенство $d \leq n-k$, вопреки условию.

Достаточность. Пусть любая совокупность k номеров компонент кодового вектора является информационной, но минимальное расстояние не удовлетворяет равенству $d = n - k + 1$. Тогда найдутся два таких вектора u и v кода, что $d(u, v) < n - k + 1$. В таком случае векторы u и v совпадают в $n - d > n - (n - k + 1) = k - 1$, т.е. по крайней мере, в k компонентах, совокупность номеров которых, таким образом, оказывается не информационной, вопреки условию. Теорема доказана.

Таким образом, оба утверждения: "Код удовлетворяет границе Синглтона" и "Любая совокупность k номеров компонент — информационная" эквивалентны и являются равносильными определениями кода МДР.

Двоичных кодов МДР имеется только два. Первый — это $(n, n-1, 2)$-код, т.е. код с проверкой на четность. Он содержит 2^{n-1} векторов, и все они четного веса. Второй — это двойственный ему $(n, 1, n)$-код. Он состоит из двух векторов: нулевого и сплошь единичного.

Теорема 7.1.4. *Код, двойственный коду МДР, есть также код МДР.*

1-е Доказательство. Пусть параметры кода МДР есть n, k, d и $d = n - k + 1$. Любые $d-1$ столбцов проверочной матрицы линейно независимы. Пусть параметры двойственного кода есть n, $k^{'}$, $d^{'}$. Тогда $k^{'} = n - k = d - 1$, и, значит, линейно независимы любые $k^{'}$ столбцов проверочной матрицы. Значит, отличен от нуля любой минор порядка $k^{'}$ проверочной матрицы. Значит, любой ненулевой вектор двойственного кода содержит не более, чем $k' - 1$ нулей. Значит, любой ненулевой вектор двойственного кода имеет вес $w \geq n - k^{'} + 1$. Значит минимальный вес двойственного кода, и, следовательно, его минимальное расстояние $d^{'} = n - k^{'} + 1$, что и требовалось.

2-е Доказательство. Любые k столбцов порождающей матрицы $(n, k)-$кода МДР линейно независимы, так как они образуют минор, строкам которого принадлежат компоненты информационной совокупности. Значит, расстояние двойственного кода есть $d^{'} \geq k + 1 = n - k^{'} + 1$. С другой стороны всегда $d^{'} \leq n - k^{'} + 1$. Остается знак равенства: $d^{'} = n - k^{'} + 1$. Двойственный код лежит на границе Синглтона, что и требовалось.

Читателю предлагается придумать новое доказательство.

Теорема 7.1.5. *Матрица $G = [I_{k\times k}, P_{k\times(n-k)}]$ порождает код МДР, если и только если любой минор матрицы $P_{k\times(n-k)}$ отличен от нуля.*

Доказательство. *Необходимость.* Пусть минор L порядка $l\,(1 \leq l \leq k)$ расположен на пересечении строк матрицы G с номерами r_1, r_2, ... , r_l и столбцов матрицы $[P_{k\times(n-k)}]$ с номерами c_1, c_2, ... , c_l. Если минор $L = 0$, то это значит, что вектор кода, равный некоторой линейной комбинации строк с номерами r_1, r_2, ... , r_l, имеет *в точности* l нулевых компонент с номерами c_1, c_2, ... , c_l. И этот же самый вектор имеет $l^{'} \leq l$ отличных от нуля компонент на отрезке матрицы $[I_{k\times k}]$.

Таким образом, нашему коду МДР принадлежит вектор веса $w = n - k - l + l^{'} \leq n - k$, что противоречит условию $w = d = n - k + 1$.

Отсюда следует также, что все элементы матрицы $[P_{k\times(n-k)}]$ отличны от нуля, что, впрочем, усматривается непосредственно: строка матрицы $G = [I_{k\times k}, P_{k\times(n-k)}]$ сама есть кодовый вектор, и ее вес не менее, чем $w = d = n - k + 1$.

Достаточность. На отрезке кодового вектора, отвечающем матрице $[P_{k\times(n-k)}]$, любая совокупность k номеров компо-

нент является информационной, так как произвольный минор матрицы $[P_{k\times(n-k)}]$ отличен от нуля. Совокупность k номеров компонент на отрезке, отвечающем матрице $[I_{k\times k}]$, является информационной по определению. Если же совокупность k номеров компонент состоит из k_1 номеров на отрезке, отвечающем матрице $[I_{k\times k}]$, и k_2, $(k_1 + k_2 = k)$ номеров на отрезке, отвечающем матрице $[P_{k\times(n-k)}]$, то произведение отличного от нуля минора порядка k_2 на отличное от нуля его алгебраическое дополнение порядка k_1 само будет отлично от нуля, так как это алгебраическое дополнение в каждой строке и каждом столбце содержит в точности по одной единице. Таким образом, обсуждаемая совокупность также информационная.

Теорема 7.1.6. *При выкалывании или укорочении кода МДР снова получается код МДР.*

Доказательство. Пусть параметры исходного кода есть n, k, d. После выкалывания получается код с параметрами

$$n^{'} = n - 1,\ k^{'} = k,\ d \geq d^{'} \geq d - 1.$$

Имеем последовательно:

$$d = n - k + 1 = n^{'} + 1 - k^{'} + 1 = n^{'} - k^{'} + 2;\ d - 1 = n^{'} - k^{'} + 1.$$

Наконец, так как $d^{'} \geq d - 1$, получаем $d^{'} = n^{'} - k^{'} + 1$, что и требовалось.

После укорочения $n^{'} = n - 1$, $k^{'} = k - 1$, $d + 1 \geq d^{'} \geq d$. Имеем последовательно: $n^{'} - k^{'} + 1 = n - k + 1 = d \leq d^{'}$, что и требовалось.

7.2. Коды Рида—Соломона

Определение 7.2.1. *Кодом Рида — Соломона (РС) называется код БЧХ, если корни $\alpha^b, \alpha^{b+1}, \ldots, \alpha^{b+d-2}$ его порождающего многочлена $g(x) = g_0 + g_1 x + \ldots + g_r x^r$ принадлежат тому же полю $GF(q)$, что и коэффициенты.*

Сравнивая данное определение с определением 6.1.1, видим, что степень m расширения поля $GF(q^m)$, которому принадлежат корни порождающего многочлена, удовлетворяет условию $m = 1$.

Найдем степень минимальной функции элемента, который является корнем порождающего многочлена.

Так как корни порождающего многочлена кода БЧХ, вообще говоря, принадлежат полю $GF(q^m)$, то они удовлетворяют уравнению $x^{q^m-1}-1=0$. По теореме 3.5.16 степень минимальной функции не превосходит m. Значит, при $m=1$ степень каждой минимальной функции каждого корня α^i порождающего многочлена равна единице, и имеет вид $x-\alpha^i$.

Теорема 7.2.2. *Код Рида—Соломона есть код МДР.*

Д о к а з а т е л ь с т в о. Порождающий многочлен кода БЧХ есть наименьшее общее кратное минимальных функций корней многочлена. В нашем случае

$$g(x)=(x-\alpha^b)(x-\alpha^{b+1})\cdots(x-\alpha^{b+d-2}), \qquad (7.2.1)$$

так как сопряженных корней нет, и разные корни имеют разные и взаимно простые минимальные функции. Таким образом, с одной стороны степень порождающего многочлена равна числу $n-k$ проверочных символов кода, а с другой стороны, она равна $d-1$ по числу скобок в (7.2.1). Поэтому $n-k=d-1$, и теорема доказана.

Следует отдать себе отчет, в чем причина того, что код РС лежит на границе Синглтона. Каждый корень порождающего многочлена циклического кода вообще, и кода БЧХ, в частности, добавляет один проверочный символ, но при этом отнюдь не каждый корень увеличивает кодовое расстояние. Более того, если корни порождающего многочлена над $GF(q)$ принадлежит полю $GF(q^m)$, то при выводе границы (6.3.10) "закладываются" на максимальное число m сопряженных корней каждой из минимальных функций, входящих сомножителями в порождающий многочлен. Поэтому и получается, что расстояние $2t+1$ требует $2tm$ проверочных символов (tm при $q=2$). В случае же кода РС каждая новая единица кодового расстояния требует в точности одного корня порождающего многочлена, так как $m=1$, а потому и в точности одного проверочного символа. Это и означает, что $n-k=d-1$.

Будем считать, что в определении кода РС $b=1$.

Если α примитивный элемент поля $GF(q)$, то согласно теореме 6.1.2 длина кода РС $n=q-1$.

П р и м е р 7.2.

Пусть $q = 5,\ n = q - 1 = 4$. Мультипликативная группа поля $GF(5)$ — это приведенная система вычетов по модулю 5. Положим $\alpha = 2,\ \alpha^2 = 4$ — корни порождающего многочлена. Тогда $g(x) = (x-2)(x-4) = 3 - x + x^2$. Порождающая матрица кода РС будет

$$G = \begin{bmatrix} 3 & -1 & 1 & 0 \\ 0 & 3 & -1 & 1 \end{bmatrix}.$$

Ее приведенно-ступенчатая форма

$$G = \begin{bmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & 3 & 2 \end{bmatrix}.$$

7.3. Кодирование кода РС

Разумеется, кодирование кода РС можно выполнить двояко, опираясь на методы, известные по предыдущему изложению. Во-первых, полагая код РС обычным линейным кодом и умножая информационный вектор на порождающую матрицу. Во-вторых, полагая код РС циклическим кодом, и умножая информационный многочлен на порождающий. Однако справедлива

Теорема 7.3.1. *Положим $a_i \in GF(q)$, $(i = 0, 1 \ldots, k-1)$, $\alpha \in GF(q)$, и пусть $a = (a_0, a_1, \ldots, a_{k-1})$– вектор информационных символов, а значит, $a(x) = a_0 + a_1 x + \ldots + a_{k-1}x^{k-1}$ — информационный многочлен. Тогда вектором кода РС будет*

$$u = (a(1), a(\alpha), \ldots, a(\alpha^{q-2}))$$

Д о к а з а т е л ь с т в о. Рассмотрим многочлен

$$u(x) = a(1) + a(\alpha)x + \ldots + a(\alpha^{q-2})x^{q-2} =$$

$$\begin{aligned} &= (a_0 + a_1 \ldots + a_{k-1}) + \\ &+ (a_0 + a_1\alpha \ldots + a_{k-1}\alpha^{k-1})x + \\ &+ (a_0 + a_1\alpha^2 \ldots + a_{k-1}\alpha^{2(k-1)})x^2 + \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &+ (a_0 + a_1\alpha^{q-2} \ldots + a_{k-1}\alpha^{(q-2)(k-1)})x^{q-2} = \end{aligned}$$

(после изменения порядка суммирования)

$$
\begin{array}{l}
= a_0(1 + x + x^2 + \ldots + x^{q-2}) + \\
+a_1(1 + \alpha x + \alpha^2 x^2 + \ldots + (\alpha x)^{q-2}) + \\
+a_2(1 + \alpha^2 x + \alpha^4 x^2 + \ldots + (\alpha^2 x)^{q-2}) + \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
+a_{k-1}(1 + \alpha^{k-1} x + (\alpha^{k-1} x)^2 + \ldots + (\alpha^{k-1} x)^{q-2}).
\end{array}
$$

Найдем отсюда $u(\alpha^{-i})$ для всех $i = 0, 1, \ldots, k-1$.

$$
\begin{array}{l}
u(\alpha^{-i}) = a_0(1 + \alpha^{-i} + \alpha^{-2i} + \ldots + \alpha^{(q-2)i}) + \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
+a_i(1 + \alpha^{-i}\alpha^i + \alpha^{-2i}\alpha^{2i} + \ldots + (\alpha^i \alpha^{-i)})^{q-2}) + \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
+a_{k-1}(1 + \alpha^{k-1}\alpha^{-i} + (\alpha^{k-1}\alpha^{-i})^2 + \ldots + (\alpha^{k-1}\alpha^{-i})^{q-2}) = -a_i
\end{array}
\qquad (7.3.2)
$$

Действительно, только в одной из скобок (7.3.2) все слагаемые обращаются в единицу; слагаемых в скобке ровно $q-1$ штук, величина q есть степень (быть может и первая) характеристики поля. Поэтому $q - 1 = -1$.

В каждой из остальных скобок содержится сумма членов геометрической прогрессии вида

$$1 + \alpha^j + \alpha^{2j} + \ldots + \alpha^{(q-2)j} = (\alpha^{j(q-1)} - 1)/(\alpha^j - 1) = 0.$$

В самом деле, так как $\alpha \in GF(q)$, то α^j есть корень двучлена $x^{q-1} - 1$. Поэтому $(\alpha^{j(q-1)} - 1) = 0$. Но $\alpha^j - 1 \neq 0$.

Как только $i \geq k$, в сумме (7.3.2) больше не будет скобки с $q-1$ единичными слагаемыми. Поэтому при $k \leq i \leq q-2$ нулевой будет каждая скобка, и, значит, $u(\alpha^{-i}) = 0$.

Но $\alpha^{-i} = \alpha^{q-1-i}$. Следовательно, $u(\alpha^j) = 0$ при $j = 1, 2, \ldots, q-1-k$. Иначе говоря, вектор u есть вектор кода БЧХ с минимальным расстоянием $d = q - k = q - 1 - k + 1 = n - k + 1$, а значит, и кода РС. Теорема доказана.

Значение этой теоремы не только и не столько в простоте и практичности процедуры кодирования. Этот способ кодирования показывает, что, имея дело с кодом РС, можно вполне обойтись без порождающего многочлена с его корнями, и без порождающей матрицы. Наоборот, и более того: из самого способа кодирования получаются корни любого кодового многочлена (в том числе и порождающего) кода РС. И именно поэтому изложенный способ кодирования первоначально выступил

в качестве определения кода РС. Суть в том, что корректирующая способность кода РС полностью определяется его длиной n и размерностью k. За их разностью $n-k$ полностью скрывается и порождающий многочлен $g(x)$, и кодовое расстояние d.

Однако такая процедура кодирования не сохраняет информационные символы, и код не является систематическим.

Восстановление информационного вектора из кодового вектора содержится в самом доказательстве теоремы 7.3.1. Действительно, согласно формуле (7.3.2) $u(\alpha^{-i}) = -a_i$ для всех $i = 0, 1, \ldots, k-1$.

Заметим, что размерность k кода определяется степенью информационного многочлена $a(x)$. Как отнестись к тому, что степени k_1 и $k_2 < k_1$ двух информационных многочленов, соответственно $a_1(x)$ и $a_2(x)$ различны? Может показаться, что из доказательства теоремы следует, будто кодовые многочлены $u_1(x)$ и $u_2(x)$, отвечающие многочленам, соответственно $a_1(x)$ и $a_2(x)$, принадлежат различным кодам РС. Ведь из доказательства теоремы вытекает, что корнями многочлена $u_1(x)$ будут α^j при $j = 1, 2, \ldots, q-1-k_1$, а корнями многочлена $u_2(x)$ будут α^j при $j = 1, 2, \ldots, q-1-k_2$. Один список корней содержится в другом, ответ на поставленный вопрос должен быть положительным. И коды действительно разные! На самом деле вектор $u_2(x)$ принадлежит коду РС размерности k_2. Однако этот код содержится в другом коде РС большей размерности k_1.

Такое соотношение между кодами РС называется вложением кодов РС.

П р и м е р 7.3.

Читатель построил поле $GF(3^2)$ по модулю многочлена x^2+x+2, решая задачу 3.1. Если $\alpha^2+\alpha+2=0$, то таблица поля имеет вид:

$$\begin{array}{rrrl}
0 = 0 & & & = (00),\\
\alpha^0 = & 1 & & = (10),\\
\alpha^1 = & & \alpha & = (01),\\
\alpha^2 = & 1 & +2\alpha & = (12),\\
\alpha^3 = & 2 & +2\alpha & = (22),\\
\alpha^4 = & 2 & & = (20),\\
\alpha^5 = & & 2\alpha & = (02),\\
\alpha^6 = & 2 & +\alpha & = (21),\\
\alpha^7 = & 1 & +\alpha & = (11),\\
\alpha^8 = & 1 & & = (10).
\end{array} \qquad (7.3.3)$$

Производя вычисления в этом поле и положив информационный многочлен

$$a(x) = 1 + \alpha x + \alpha^2 x^2 + \alpha^3 x^3, \tag{7.3.4}$$

получим

$$a(1) = \alpha^2,\ a(\alpha) = 0,\ a(\alpha^2) = \alpha^6,\ a(\alpha^3) = 0,\ a(\alpha^4) = \alpha^5,$$

$$a(\alpha^5) = 0,\ a(\alpha^6) = \alpha^7,\ a(\alpha^7) = 1,$$

Построен вектор кода РС с параметрами $n = 8,\ k = 4,\ d = 5$:

$$u = (\alpha^2, 0, \alpha^6, 0, \alpha^5, 0, \alpha^7, 1). \tag{7.3.5}$$

Если же информационный многочлен есть $a(x) = 2 + \alpha x$, то $a(1) = \alpha^6,\ a(\alpha) = \alpha : 5,\ a(\alpha^2) = \alpha^2,\ a(\alpha^3) = 1,\ a(\alpha^4) = \alpha^3,\ a(\alpha^5) = \alpha^7,\ a(\alpha^6) = \alpha,\ a(\alpha^7) = 0,$ и вектор кода РС с параметрами $n = 8,\ k = 2,\ d = 7$ есть

$$u = (\alpha^6, \alpha^5, \alpha^2, 1, \alpha^3, \alpha^7, \alpha, 0). \tag{7.3.6}$$

Второй код содержится в первом, т.е. вложен в первый. Вложение кодов фактически уже было отмечено в задаче 6.5. Там рассматривались вложенные коды БЧХ. Частным случаем вложения кодов БЧХ является вложение кодов РС, так как сами коды РС являются частным случаем кодов БЧХ.

7.4. Удлинение кодов РС

"Обычные" коды БЧХ допускают переход к существенно новым длинам без изменения алфавита. Например, алфавитом может быть поле $GF(q)$, а корни порождающего многочлена принадлежат полю $GF(q^m)$, и при этом степень m расширения может расти. Зато обычные коды БЧХ не принадлежат к классу кодов МДР.

Коды РС — это "не обычные" коды БЧХ и такого удлинения не допускают: стоит существенно увеличить длину $q - 1$, как немедленно изменится алфавит. Поэтому заслуживает внимания каждое такое удлинение, при котором сохраняется и алфавит, и принадлежность кода к классу кодов МДР.

В этом разделе будет обсуждаться удлинение кода РС на один и два, а в некоторых случаях даже и на три символа. Пусть

$$v = (c_0, c_1, \ldots, c_{q-2}) \tag{7.4.7}$$

вектор кода РС, порождающий многочлен которого есть

$$g(x) = (x-\alpha)(x-\alpha^2)\cdots(x-\alpha^{d-1}),\ \deg g(x) = d-1 = r = n-k, \tag{7.4.8}$$

где $n = q-1$.

Положим

$$c_{q-1} = -\sum_{0}^{q-2} c_i. \tag{7.4.9}$$

Если $w(v) = d$, то добавление к вектору v символа c_{q-1}, увеличит его вес до $d+1$, если $c_{q-1} \neq 0$. Покажем, что $c_{q-1} \neq 0$. Для этого рассмотрим многочлен

$$v(x) = (c_0 + c_1 x + \ldots + c_{q-2}x^{q-2}). \tag{7.4.10}$$

Он принадлежит коду РС. Поэтому $v(x) = g(x)z(x)$. Согласно (7.4.9) $v(1) = g(1)z(1) = -c_{q-1}$.

$g(1) \neq 0$, так как 1 не принадлежит списку корней порождающего многочлена. Но и $z(1) \neq 0$, так как в противном случае $(x-1)g(x)|v(x)$, и согласно границе кодов БЧХ, $w(v) = d+1$, вопреки условию. Получилось: $n' = n+1$, $k' = k$, $d' = d+1$.

Иначе говоря, $d' = n' - k' + 1$, т.е. удлиненный код снова лежит на границе Синглтона, а потому он есть код МДР. Не забудем, что $n = q-1$, $n' = q$.

Итак, удлиненный вектор кода РС есть

$$v' = (c_0, c_1, \ldots, c_{q-2}, c_{q-1}), \tag{7.4.11}$$

где c_{q-1} получается как (7.4.9).

Удлинение кода РС на один символ называется "1-удлинение".

Так как проверочная матрица кода РС с порождающим многочленом (7.4.8) имеет вид

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \ldots & \alpha^{q-2} \\ 1 & \alpha^2 & \alpha^4 & \ldots & \alpha^{2(q-2)} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha^{d-1} & \alpha^{2(d-1)} & \ldots & \alpha^{(d-1)(q-2)} \end{bmatrix}, \tag{7.4.12}$$

то проверочная матрица H_1 1-удлиненного кода РС будет

$$H_1 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{q-2} & 0 \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(q-2)} & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \alpha^{d-1} & \alpha^{2(d-1)} & \dots & \alpha^{(d-1)(q-2)} & 0 \end{bmatrix}. \quad (7.4.13)$$

Легко видеть, что скалярные произведения 1-удлиненного кодового вектора (7.4.11) на все строки матрицы (7.4.13), начиная со второй, совпадают со скалярными произведениями этого вектора на все строки матрицы (7.4.12), так как нуль в конце каждой ее строки вносит в произведение только нулевой вклад. Зато, учитывая (7.4.9), произведение вектора на первую строку равно в точности

$$c_0 + c_1 + \ldots + c_{q-2} + c_{q-1} = c_0 + c_1 + \ldots + c_{q-2} - \sum_{0}^{q-2} c_i = 0.$$

что и требуется.

С другой стороны, легко убедиться, что любые d столбцов матрицы (7.4.13) линейно независимы. Действительно, если в число этих d столбцов последний столбец не входит, то они образуют определитель Вандермонда, который отличен от нуля, так как элементы его второй строки различны. Если же столбец $(0\,0\,\ldots\,0\,1)^T$ в число произвольно выбранных d столбцов входит, то разлагая полученный определитель по элементам указанного столбца, получим, что определитель равен своему минору порядка $d-1$, а он есть снова определитель Вандермонда.

Обратимся теперь к случаю 2-удлинения кодов РС.

Положим, что еще одним символом c_q будет

$$c_q = -\sum_{i=0}^{q-2} c_i \alpha^{i(d)}. \quad (7.4.14)$$

Он заведомо отличен от нуля, так как α^d не принадлежит к множеству корней порождающего многочлена (7.4.8). См. также определение 7.2.1.

Вектором 2-удлиненного кода РС будет

$$v'' = (c_0, c_1, \ldots, c_{q-2}, c_{q-1}, c_q), \tag{7.4.15}$$

где c_{q-1} и c_q выражаются соответственно формулами (7.4.9) и (7.4.14).

Покажем, что проверочной матрицей 2-удлиненного кода РС будет матрица

$$H_2 = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 & 1 & 0 \\ 1 & \alpha & \alpha^2 & \ldots & \alpha^{q-2} & 0 & 0 \\ 1 & \alpha^2 & \alpha^4 & \ldots & \alpha^{2(q-2)} & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \alpha^{d-1} & \alpha^{2(d-1)} & \ldots & \alpha^{(d-1)(q-2)} & 0 & 0 \\ 1 & \alpha^d & \alpha^{2d} & \ldots & \alpha^{d(q-2)} & 0 & 1 \end{bmatrix}. \tag{7.4.16}$$

Действительно, как и в случае 1-удлинения, легко видеть, что скалярные произведения 2-удлиненного кодового вектора (7.4.15) на все строки матрицы (7.4.16) до предпоследней включительно, совпадают со скалярными произведениями этого вектора на все строки матрицы (7.4.13), так как нули двух последних столбцов в эти произведениях никакого вклада не вносят. Зато произведение вектора (7.4.15) на последнюю строку в точности равно

$$c_0 + c_1\alpha^d + \ldots + c_{q-2}\alpha^{d(q-2)} - \sum_{i=0}^{q-2} c_i\alpha^{i(d)} = 0,$$

что и требуется.

С другой стороны, легко убедиться, что любые $d+1$ столбцов матрицы (7.4.16) линейно независимы. Действительно, если в число этих $d+1$ столбцов последние два столбца не входят, то они образуют определитель Вандермонда, который отличен от нуля, так как все элементы его второй строки различны. Если в число этих $d+1$ столбцов входит только один из двух последних столбцов, то разлагая определитель по элементам этого столбца, найдем, что он равен своему минору порядка d. Но этот минор есть также определитель Вандермонда. Если же в число выбранных $d+1$ столбцов входят оба последних столбца, то разлагая определитель сначала по элементам одного, а затем и по элементам другого столбца, получим снова определитель Вандермонда, но теперь уже порядка $d-1$.

Строение матрицы (7.4.16) не оставляет места для дальнейших попыток удлинения. Некуда, так сказать, поместить еще один столбец с одной единицей. Это препятствие снимается для случая поля $GF(2^m)$ при $k = 3$ и $k = 2^m - 1$. Именно, верна

Теорема 7.4.1. *Существуют* 3*-удлиненные коды РС с параметрами*

$$n = 2^m + 2,\ k = 2^m - 1,\ d = 4 \quad u \quad n = 2^m + 2,\ k = 3,\ d = 2^m.$$

Д о к а з а т е л ь с т в о. Рассмотрим проверочную матрицу

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{q-2} \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(q-2)} \end{bmatrix}$$

$(q-1,\ q-3,\ 3)$-кода РС, где $q = 2^m - 1$, и порождающий многочлен есть $g(x) = (x - \alpha)(x - \alpha^2.)$

Из нее можно получить согласно (7.4.13) проверочную матрицу

$$H_1 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{q-2} & 0 \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(q-2)} & 0 \end{bmatrix} \tag{7.4.17}$$

1-удлиненного кода РС с параметрами $n^{'} = q = 2^m, k = q - 3$, $d = 4$, и новый символ имеет вид (7.4.9).

Добавим к матрице (7.4.17) два столбца, не увеличивая числа строк:

$$H_3 = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 & 1 & 0 & 0 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{q-2} & 0 & 1 & 0 \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(q-2)} & 0 & 0 & 1 \end{bmatrix}. \tag{7.4.18}$$

Теорема 7.4.2. *Любые три столбца матрицы* (7.4.18) *линейно независимы.*

Д о к а з а т е л ь с т в о. Любые три столбца из первых $q-1$ столбцов матрицы образуют определитель Вандермонда, который отличен от нуля, так как в любой его строке элементы, отличные от единицы, различны. Три последних столбца также линейно независимы. Остается рассмотреть "смешанный" состав тройки столбцов. Если в определитель входят

какие-нибудь два из трех последних столбцов, то после двух последовательных разложений определителя по элементам этих столбцов, получится минор первого порядка, и он заведомо есть $\alpha^i \neq 0$, где $i = 0, 1, \ldots, q-2$. Рассуждения на случай вхождения в определитедь столбцов $(001)^T$ или $(100)^T$ аналогичны случаю матрицы (7.4.16). Интерес представляет случай вхождения в определитель столбца $(010)^T$. Отвечающий ему минор

$$M = \begin{bmatrix} 1 & 1 \\ \alpha^{2i} & \alpha^{2j} \end{bmatrix}$$

заведомо отличен от нуля, так как (и это центральная деталь доказательства) в поле характеристики 2, и только этой характеристики, все вторые степени различны (см. раздел 3.7),

Итак, матрица (7.4.18) является проверочной для первого кода теоремы 7.4.1. Она же служит порождающей для второго кода теоремы. Оба кода лежат на границе Синглтона.

7.5. Декодирование кодов РС

Коды РС — это коды БЧХ. Поэтому к ним применимы все методы декодирования, в том числе и метод, изложенный в разделе 6.9. Здесь сначала будет применен именно классический метод Питерсона— Горенстейна —Цирлера, который, как уже отмечалось, по распространённому мнению, лучше других раскрывает алгебраическую сущность процесса декодирования.До конца этой главы будет изложен метод, основанный на алгоритме Эвклида. При этом будут исследованы случаи и ошибок и стираний в принятом векторе.

П р и м е р 7.4.

Рассмотрим код РС над $GF(3^2)$ с корнями α, α^2, α^3, α^4 порождающего многочлена.

Поле $GF(3^2)$ изображено на таблице (7.3.3). Длина кода $n = 8$, и минимальное расстояние $d = 5$. Код исправляет любые ошибки кратности 1 и 2. Проверочная матрица кода есть

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & 1 & \alpha^2 & \alpha^4 & \alpha^6 \\ 1 & \alpha^3 & \alpha^6 & \alpha & \alpha^4 & \alpha^7 & \alpha^2 & \alpha^5 \\ 1 & \alpha^4 & 1 & \alpha^4 & 1 & \alpha^4 & 1 & \alpha^4 \end{bmatrix}. \qquad (7.5.19)$$

Положим, принят вектор

$$u = (0, 0, 0, 0, \alpha^5, 0, \alpha^7, 1) \tag{7.5.20}$$

над полем (7.3.3), в котором производятся все дальнейшие вычисления.

Разумеется, одного взгляда достаточно, чтобы по весу этого вектора понять, что он содержит по меньшей мере две ошибки. Однако для указания истинного вектора необходима формальная процедура.

Скалярные произведения этого вектора на строки проверочной матрицы дают элементы синдрома:

$$S_1 = \alpha^7,\ S_2 = \alpha^2,\ S_3 = \alpha,\ S_4 = 0. \tag{7.5.21}$$

Подставив эти значения в (6.9.52), получим систему уравнений относительно коэффициентов σ_1, σ_2 многочлена локаторов ошибок.

$$\begin{aligned} \alpha^2\sigma_1 + \alpha^7\sigma_2 &= -\alpha, \\ \alpha\sigma_1 + \alpha^2\sigma_2 &= 0. \end{aligned}$$

Отсюда $\sigma_1 = \alpha^7$, $\sigma_2 = \alpha^2$.

Многочлен локаторов ошибок

$$\sigma(z) = \alpha^2 z^2 + \alpha^7 z + 1. \tag{7.5.22}$$

Легко проверить, что его корни, которые являются величинами, обратными локаторам X_1, X_2 ошибок, есть $z_1 = \alpha^0 = 1$, $z_2 = \alpha^6$.

Отсюда $X_1 = \alpha^0 = 1$, $X_2 = \alpha^2$. Это означает, что в принятом векторе (7.5.20) первый и третий нули — это неверные символы. Иначе говоря, вектор-ошибка имеет вид $e(x) = e_1 + e_2\alpha^2$. Коэффициенты $e_1 = Y_1$, $e_2 = Y_2$ этого вектора пока неизвестны. Для их отыскания подставим известные величины

$$S_1 = \alpha^7,\ S_2 = \alpha^2,\ X_1 = 1,\ X_2 = \alpha^2$$

в систему (6.9.48):

$$\begin{aligned} Y_1 + \alpha^2 Y_2 &= \alpha^7, \\ Y_1 + \alpha^4 Y_2 &= \alpha^2. \end{aligned}$$

Отсюда $Y_1 = \alpha^6$, $Y_2 = \alpha^2$. Таковы значения ошибок, которые надлежит вычесть, соответственно из первого и третьего символов принятого вектора: $0 - \alpha^6 = \alpha^2$, $0 - \alpha^2 = \alpha^6$.

Истинным оказывается вектор

$$u = (\alpha^2, 0, \alpha^6, 0, \alpha^5, 0, \alpha^7, 1),$$

который есть не что иное, как построенный выше вектор (7.3.5).

Чтобы "замкнуть круг" рассуждений, вычислим информационный вектор $a = (a_0\, a_1 \ldots, a_{k-1})$, подставляя в

$$u(x) = \alpha^2 + \alpha^6 x^2 + \alpha^5 x^4 + \alpha^7 x^6 + x^7$$

последовательно $x = \alpha^{-i}, i = 0, 1, 2, 3$, и пользуясь при этом таблицей (7.3.3) поля $GF(3^2)$. Читатель может убедиться, что

$$\begin{gathered} u(1) = 1 + 1 + 1 + 1 + 1 = 2 = -1, a_0 = 1, \\ u(\alpha^{-1}) = \alpha^2 + \alpha^4 + \alpha + \alpha + \alpha = \alpha^5 = -\alpha,\ a_1 = \alpha, \\ u(\alpha^{-2}) = \alpha^2 + \alpha^2 + \alpha^5 + \alpha^3 + \alpha^2 = \alpha^6 = -\alpha^2,\ a_2 = \alpha^2, \\ u(\alpha^{-3}) = \alpha^2 + 1 + \alpha + \alpha^5 + \alpha^3 = \alpha^7 = -\alpha^3,\ a_3 = \alpha^3. \end{gathered}$$

Остается сравнить полученный результат с (7.3.4).

Некоторые замечания к декодированию кода РС. Принадлежность многочлена $u(x)$ к коду РС над $GF(q)$ той или иной размерности на передающем конце, разумеется, известна, хотя результат кодирования от этого знания согласно теореме 7.3.1 никак не зависит. Тем не менее в рассмотренном примере декодирования все параметры кода РС обозначены полностью. Но если при декодировании знание этих параметров необходимо, то стоит ли подчеркивать возможность пренебрежения ими при кодировании. И, наоборот, если безразличие к параметрам кода на передающем конце столь значимо, то не является ли явной асимметрией опора при декодировании именно на параметры кода РС ?

П р и м е р 7.5.

Пусть принят вектор

$$u' = (0,\ 0,\ 0,\ 1, \alpha^3, \alpha^7, \alpha, 0,), \tag{7.5.23}$$

представляющий собой вектор (7.3.6), в котором искажены первые три символа. Сам вектор (7.3.6) получен кодированием информационного вектора $a = (2,\ \alpha)$, или в многочленной форме, $a(x) = 2 + \alpha x$. Так как $k = 2$, $n - k = 8 - 2 = 6 = d - 1$, $t = 3$, то

код РС исправляет любые три ошибки. Проверочная матрица на этот случай будет

$$H = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7 \\ 1 & \alpha^2 & \alpha^4 & \alpha^6 & 1 & \alpha^2 & \alpha^4 & \alpha^6 \\ 1 & \alpha^3 & \alpha^6 & \alpha & \alpha^4 & \alpha^7 & \alpha^2 & \alpha^5 \\ 1 & \alpha^4 & 1 & \alpha^4 & 1 & \alpha^4 & 1 & \alpha^4 \\ 1 & \alpha^5 & \alpha^2 & \alpha^7 & \alpha^4 & \alpha & \alpha^6 & \alpha^3 \\ 1 & \alpha^6 & \alpha^4 & \alpha^2 & 1 & \alpha^6 & \alpha^4 & \alpha^2 \end{bmatrix}. \qquad (7.5.24)$$

Она получилась добавлением к матрице (7.5.19) двух строк, отвечающих еще двум корням любого многочлена кода РС размерности $k = 2$. Элементы синдрома

$$S_1 = \alpha,\ S_2 = 1,\ S_3 = \alpha^3,\ S_4 = \alpha^5,\ S_5 = 1,\ S_6 = \alpha^5.$$

Система трех линейных уравнений относительно коэффициентов σ_1, σ_2, σ_3 выглядит так

$$\begin{aligned} &\alpha^3\sigma_1 + \sigma_2 + \alpha\sigma_3 = \alpha, \\ &\alpha^5\sigma_1 + \alpha^3\sigma_2 + \sigma_3 = \alpha^4, \\ &\sigma_1 + \alpha^5\sigma_2 + \alpha^3\sigma_3 = \alpha. \end{aligned}$$

Отсюда

$$\sigma_1 = 1,\ \sigma_2 = \alpha^5,\ \sigma_3 = \alpha^7,$$

и многочлен локаторов ошибок имеет вид

$$\sigma(z) = 1 + z + \alpha^5 z^2 + \alpha^7 z^3. \qquad (7.5.25)$$

Легко проверить, что его корни, которые являются величинами, обратными локаторам X_1, X_2, X_3 ошибок, есть

$$z_1 = \alpha^0 = 1,\ z_2 = \alpha^7,\ z_3 = \alpha^6$$

и

$$X_1 = 1,\ X_2 = \alpha,\ X_3 = \alpha^2,$$

что фактически нам известно из условий задачи.

Коэффициенты $e_1 = Y_1$, $e_2 = Y_2$, $e_3 = Y_3$, многочлена-ошибки, конечно, также известны из условий задачи и усматриваются непосредственно из (7.3.6): $Y_1 = \alpha^2$, $Y_2 = \alpha$, $Y_3 = \alpha^6$.

Однако для соблюдения формализмов их можно найти из системы первых трех уравнений системы (6.9.48):

$$\begin{aligned} Y_1 + \alpha Y_2 + \alpha^2 Y_3 &= \alpha, \\ Y_1 + \alpha^2 Y_2 + \alpha^4 Y_3 &= 1, \\ Y_1 + \alpha^3 Y_2 + \alpha^6 Y_3 &= \alpha^3. \end{aligned}$$

Вычисления предоставляются читателю.

Доводя рассуждения до абсурда, можно было бы потребовать, чтобы впереди каждого передаваемого вектора следовал "флаг" , который сообщал, какой степени k информационный многочлен закодирован в данном акте передачи. Тогда декодер приготовится исправлять $t = [(q-1-k)/2]$ ошибок.

Число строк в проверочной матрице будет варьироваться. И хотя в каждом векторе, принадлежащем коду размерности k_1 можно было бы исправить большее число ошибок, чем в векторе, принадлежащем коду размерности $k_2 > k_1$, все же здравый смысл подсказывает настроить декодирование на максимальное для данной передачи число k, тем более, что, во-первых, появление большего числа ошибок мало вероятно, а во-вторых, с убыванием k убывает и число векторов, в которых можно было бы исправлять большее число ошибок. На самом деле источником только что предпринятых рассуждений является все тот же факт вложения кодов РС.

7.6. Алгоритм Эвклида для многочленов

Теперь мы переходим к изложению того метода декодирования, который носит название декодирования посредством алгоритма Эвклида. Эта тема была заявлена в конце раздела 6.7, и ей посвящены разделы 7.7 – 7.10.

В гл. 1 алгоритм Эвклида был представлен как средство нахождения наибольшего общего делителя (Н.О.Д.) двух целых чисел a и b, $b < a$. Существует алгоритм Эвклида также и для многочленов. Его принцип действия точно такой же, что и для целых чисел, с той лишь разницей, что вместо соотношений между целыми числами, рассматриваются соотношениям между многочленами, являющихся делимыми, делителями,

частными и остатками, и сравниваются их степени.

$$\begin{array}{ll} & \deg b(z) \le \deg a(z) \\ a(z) = b(z)q_0(z) + r_0(z), & \deg r_0(z) < \deg b(z) \\ b(z) = r_0(z)q_1(z) + r_1(z), & \deg r_1(z) < \deg r_0(z) \\ r_0(z) = r_1(z)q_2(z) + r_2(z), & \deg r_2(z) < \deg r_1(z) \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots \\ r_k(z) = r_{k+1}(z)q_2(z) + r_{k+2}(z), & \deg r_{k+2}(z) < \deg r_{k+1}(z) \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots \\ r_{n-2}(z) = r_{n-1}(z)q_n(z) & 0 = r_n(z) \end{array} \tag{7.6.26}$$

Последнее равенство, где $r_n(z) = 0$, неизбежно ввиду уменьшения степеней остатков на каждом следующем шаге деления. Если $\deg r_{n-1}(z) = 0$, то заведомо $r_n(z) = 0$, так как константа поля делит любой многочлен над этим полем всегда без остатка.

Отсюда в полном соответствии с (1.2.3) имеем

$$(a(z), b(z)) = (b(z), r_0(z)) = (r_0(z), r_1(z)) = \ldots = (r_{n-1}(z), 0) = r_{n-1}(z).$$

Пример 7.6. Пусть над полем $GF(2)$ $a(z) = z^4 + z^3 + z^2 + 1$, $b(z) = z^4 + z^2 + z + 1$.

$$\begin{array}{l} z^4 + z^3 + z^2 + 1 = (z^4 + z^2 + z + 1) + (z^3 + z), \\ (z^4 + z^2 + z + 1) = (z^3 + z)z + (z + 1), \\ z^3 + z = (z + 1)(z^2 + z). \end{array} \tag{7.6.27}$$

Таким образом, $z + 1 = (a(z), b(z))$,

$$q_0 = 1,\ r_0 = (z^3 + z);\ q_1 = z,\ r_1 = z + 1;\ q_2 = z^2 + z,\ r_2 = 0.$$

На случай многочленов имеет место также аналог равенства (1.2.4). Оно получается процедурой (1.2.5)—(1.2.14).

Следует только заменить символы q_k, u_k, v_k, r_k символами $q_k(z), u_k(z), v_k(z), r_k(z)$.

Например, системы (1.2.5), (1.2.6) и (1.2.9), (1.2.10), (1.2.11) заменяются системами, соответственно

$$\begin{array}{l} u_{-2}(z) = 0,\ u_{-1}(z) = 1. \\ v_{-2}(z) = 1,\ v_{-1}(z) = 0, \end{array} \tag{7.6.28}$$

$$\begin{array}{l} u_k(z) = q_k(z)u_{k-1}(z) + u_{k-2}(z), \\ v_k(z) = q_k(z)v_{k-1}(z) + v_{k-2}(z). \end{array} \tag{7.6.29}$$

и

$$\begin{array}{l} v_{k-1}(z)r_k(z) + v_k(z)r_{k-1}(z) = r_{-1}(z), \\ u_{k-1}(z)r_k(z) + u_k(z)r_{k-1}(z) = r_{-2}(z), \\ v_k(z)u_{k-1}(z) - u_k(z)v_{k-1}(z) = (-1)^k. \end{array} \tag{7.6.30}$$

Нетрудно посчитать

$$\begin{array}{l} \deg u_i(z) = \sum\limits_{k=1}^{i} \deg q_k(z), \\ \deg r_{i-1}(z) = \deg r_{-1}(z) - \sum\limits_{k=1}^{i} \deg q_k(z), \\ \deg u_i(z) = \deg r_{-1}(z) - \deg r_{i-1}(z) < \deg r_{-1}(z), \\ \deg v_i(z) = \sum\limits_{k=1}^{i} \deg q_k(z), \\ \deg v_i(z) = \deg r_{-1}(z) - \deg r_{i-1}(z) < \deg r_{-1}(z). \end{array} \tag{7.6.31}$$

Утверждение 7.6.1. *Многочлены $u_k(z)$ и $v_k(x)$ взаимно просты.*

Действительно, если бы оба слагаемых в левой части последнего равенства в (7.6.30) делились на какой-нибудь многочлен, отличный от константы, то и правая часть обладала бы таким свойством, что невозможно.

Возвращаясь к примеру 7.6, получаем

$$\begin{array}{l} u_{-2}(z) = 0,\ u_{-1}(z) = 1,\ u_0(z) = 1,\ u_1(z) = z^+1, \\ v_{-2}(z) = 1,\ u_{-1}(z) = 0,\ v_0(z) = 1,\ v_1(z) = z. \end{array} \tag{7.6.32}$$

Многочленный аналог равенства (1.2.4) принимает вид

$$(z+1) = z(z^4 + z^3 + z^2 + 1) + (z+1)(z^4 + z^2 + z + 1). \tag{7.6.33}$$

Вместо заключительного рассуждения процедуры (1.2.5)—(1.2.14), удобно, однако, поступить следующим образом. Решим систему (7.6.30) относительно $r_k(z)$. Получим

$$r_k(z) = (-1)^k(-v_k(z)r_{-2}(z) + u_k(z)r_{-1}(z)) \tag{7.6.34}$$

Результат (7.6.33) примера 7.6. получится здесь при $k = 1$.

7.7. Вывод ключевого уравнения

Пусть задан код РС над $GF\,(q)$ длины $n = q-1$ и размерности $k < n$. Напомним обозначения:

$u = (u_0, u_1, ...u_{n-1})$, $u_i \in GF(q)-$ кодовый вектор, переданный в канал связи,

$v = (v_0, v_1, ..., v_{n-1})$, $v_i \in GF(q)-$ полученный из канала вектор, в котором могут быть ошибки ,

$\mathbf{e} = (e_0, e_1, ..., e_{n-1}), e_i \in GF\,(q)$ $-$ вектор-ошибка, такой, что $v = u + e$.

Синдром принятого вектора имеет вид (ср. с (6.9.48))

$$S = (S_1, S_2, \dots S_{2t}),\ S_j \in GF(q^m) \tag{7.7.35}$$

$$\begin{array}{c} S_1 = Y_1X_1 + Y_2X_2 + \ldots + Y_tX_t, \\ S_2 = Y_1X_1^2 + Y_2X_2^2 + \ldots + Y_tX_t^2, \\ \cdots \\ S_{2t} = Y_1X_1^{2t} + Y_2X_2^{2t} + \ldots + Y_tX_t^{2t}, \\ Y_j \in GF(q), \quad X(j) \in GF(q^m). \end{array} \tag{7.7.36}$$

Величины X_j, как и прежде, называются локаторами ошибок, а величины Y_j – значениями ошибок. В случае кодов РС $m = 1,\ 2t = n - k = r$.

Вычисления упрощаются, если принять во внимания, что согласно теореме 3.5.1,

$$S_{jq} = Y_1X_1^{jq} + Y_2X_2^{jq} + \ldots + Y_tX_t^{jq} = (Y_1X_1^j + Y_2X_2^j + \ldots + Y_tX_t^j)^q = S_j^q.$$

Соотношений (7.7.36) достаточно для декодирования методом главы 6. Вспомним, что центральной целью там было нахождение многочлена $\sigma(z)$ локаторов ошибок. Как только это было сделано, всё дальнейшее выполнялось почти автоматически. Однако при большом числе ошибок нахождение многочлена локаторов ошибок предпочтительней выполнить, опираясь на алгоритм Эвклида для многочленов. Более того, считается, что [7] "в любом случае декодирование с помощью алгоритма Эвклида является наиболее простым для понимания, и по быстродействию оно, по меньшей мере сравнимо с другими методами при $n < 10^6$."

Сопоставим вектору (7.7.35), как это не раз делалось в предыдущих главах, многочлен от z :

$$S(z) = \sum_{j=1}^{j=2t} S_j z^{j-1} \tag{7.7.37}$$

Многочлен (7.7.37) иногда называют синдромным многочленом.

Из (7.7.37) и (7.7.36), изменив порядок суммирования и пользуясь формулой для суммы геометрической прогрессии, получим:

$$S(z) = z^{2t} \sum_{i=1}^{t} Y_i X_i \frac{X_i^{2t}}{X_i z - 1} - \sum_{i=1}^{i=t} \frac{Y_i X_i}{X_i z - 1}. \tag{7.7.38}$$

Полагая

$$\begin{aligned} \sigma(z) &= \prod_{i=1}^{l=t} (X_l z - 1) = \sum_{i=0}^{i=t} \sigma_i z^i, \ \sigma_0 = 1. \\ \omega(z) &= \sum_{i=1}^{i=t} Y_i X_i \prod_{l=1, l\neq i}^{l=t} (X_l z - 1), \\ \Phi(z) &= \sum_{i=1}^{i=t} Y_i X_i^{2t+1} \prod_{l=1, l\neq i}^{l=t} (X_l z - 1), \end{aligned} \tag{7.7.39}$$

после приведения всех дробей в (7.7.38) к общему знаменателю получим:

$$S(z) = z^{2t} \frac{\Phi(z)}{\sigma(z)} - \frac{\omega(z)}{\sigma(z)} \tag{7.7.40}$$

Выражение (7.7.40) называют *ключевым уравнением.* Придадим ему иной вид

$$S(z)\sigma(z) = z^{2t}\Phi(z) - \omega(z) \tag{7.7.41}$$

Иначе говоря,

$$-\omega(z) \equiv \sigma(z)S(z) (\mathrm{mod}\, z^{2t}) \tag{7.7.42}$$

Ограничение по модулю z^{2t} вызвано тем, что компоненты S_j синдрома вычислены и используются только до значения $j = 2t$, а потому все степени с показателями выше чем $2t - 1$ отбрасываются. Поэтому же удалён и член $z^{2t}\Phi(z)$, кратный z^{2t}. Как нам известно, к каждой части сравнения можно прибавить число, кратное модуля (см. 1.4.4). Многочлены $\sigma(z)$ и $\omega(z)$ называются многочленами, соответственно, локаторов ошибок и значений ошибок.

7.8. Решение ключевого уравнения

Для решения ключевого уравнения воспользуемся соотношением (7.6.34), придав ему вид

$$r_k(z) \equiv (-1)^k (u_k(z) r_{-1}(z) (\mathrm{mod} r_{-2}(z)) \tag{7.8.43}$$

Это соотношение справедливо при всех k, однако нас будет интересовать строго определённое значение этого индекса.

Сопоставляя сравнения (7.7.42) и (7.8.43), видим, что естественно положить

$$r_{-2}(z) = z^{2t},\ r_{-1}(z) = S(z). \tag{7.8.44}$$

Тогда при некотором k решение сравнения

$$r_k(z) \equiv (-1)^k (u_k(z) S(z) (\mathrm{mod} z^{2t}) \tag{7.8.45}$$

даст

$$\chi u_k(z) = \sigma(z) \tag{7.8.46}$$

и

$$(-1)^k \chi r_k(z) = \omega(z). \tag{7.8.47}$$

Учитывая, что степени остатков r_k строго убывают, значение k выбирается из следующих соображений. Степень ϵ многочлена $\sigma(z)$ должна быть наименьшей и во всяком случае не превосходить t. Это будет достигнуто, когда впервые выполнится неравенство $\deg r_k \leq t-1$. В самом деле, степень правой части сравнения (7.8.45) равна $\deg S(z)+\epsilon$. И так как $\epsilon \leq t,\ \deg S(z) = 2t-1$, то сравнение $\deg S(z)+\epsilon \equiv \deg r_k (\mathrm{mod}\, 2t)$ возможно при $\deg r_k \leq t-1$. С другой стороны, при $\deg r_k \geq t$ окажется, что $\epsilon \geq t+1$, вопреки условию. При выполнении условия $\deg \sigma(z) \leq t$ из сравнения (7.7.42) немедленно следует $\deg \omega(z) \leq t-1$. Действительно, $\deg \omega(z) = (2t-1) + \deg \sigma(z) - 2t \leq 3t - 1 - 2t \leq t-1$. Этот факт находится в точном соответствии с формулой (7.7.39), из которой видно, что с необходимостью выполняется неравенство $\deg \omega(z) \leq t-1$. Результат рассуждений не изменится, если отправляться не от условия $\deg \sigma(z) \leq t$, а от условия $\deg \omega(z) \leq t-1$. Заметим, что слова "впервые выполняется неравенство $\deg r_k \leq t-1$" означают ни что иное, как "пока выполняется неравенство $\deg r_{k-1} \geq t = (d-1)/2$."

Делимость многочленов не зависит от константы поля, на которую умножен многочлен. Поэтому её выбор определяется специальными требованиями. В данном случае, как это следует из (7.7.39), элемент χ выбирается из условия $\sigma(0) = \sigma_0 = 1$. При этом, разумеется, по-прежнему будут выполнены сравнения (7.7.42) и неравенства $\deg\sigma(z) \le t,\ \deg\omega(z) \le t-1$.

Теорема 7.8.1. *При $\sigma(0) = 1$, выполнении условий $\deg\sigma(z) \le t$ и $\deg\omega(z) \le t-1$ многочлены $\sigma(z)$, $\omega(z)$ из (7.8.46) и (7.8.47) единственны, и многочлен $\sigma(z)$ имеет минимальную степень.*

Д о к а з а т е л ь с т в о. Пусть имеются два решения $\sigma_1(z),\ \omega_1(z)$ и $\sigma_2(z),\ \omega_2(z)$ сравнения (7.7.42). Это означает, что $-\omega_1(z) \equiv \sigma_1(z)S(z)(\mathrm{mod}\, z^{2t})$ и $-\omega_2(z) \equiv \sigma(z)_2 S(z)(\mathrm{mod}\, z^{2t})$.

Умножив первое сравнение на $\sigma_2(z)$, а второе на $\sigma_1(z)$, получим два сравнения, у которых правые части совпадают. Значит сравнимы их левые части: $\sigma_2(z)\omega_1(z) \equiv \sigma_1(z)\omega_2(z)(\mathrm{mod}\ z^{2t}$.) Однако степень произведения многочленов в этом сравнении не превосходит $2t$. Поэтому сравнение переходит в равенство $\sigma_2(z)\omega_1(z) = \sigma_1(z)\omega_2(z)$.

Отсюда $\sigma_2(z)/\sigma_1(z) = \omega_2(z)/\omega_1(z) = \mu(z)$, где $\mu(z)-$ некоторый многочлен. Если в решении (7.8.46) и (7.8.47) степень многочлена $\sigma(z)$ не минимальна, то из предыдущего следует, что $\sigma_2(z) = \mu(z)\sigma_1$, $\omega_2(z) = \omega_1(z)\mu(z)$ есть также решение сравнения (7.7.42). Собирая (7.6.30), (7.6.34) и (7.8.47), получим для $\omega_2(z)$:

$$\omega_2(z) = (-1)^k \chi r_k(z) = \chi v_k(z) z^{2t} + \chi^{-1}\sigma(z)S(z) \qquad (7.8.48)$$

Но так как $\omega_2(z) = \mu(z)\omega_1$, то

$$\mu(z)\omega_1(z) = \chi v_{k(z)} z^{2t} + \mu(z)\sigma_1(z)S(z) \qquad (7.8.49)$$

и $\mu(z)$ делит $v_k(z)$. Однако согласно (7.8.46)

$$\chi u_k(z) = \mu(z)\sigma(z), \qquad (7.8.50)$$

и $\mu(z)$ делит также $u_k(z)$, что в силу утверждения 7.6.1 может быть, только когда $\mu(z)$ константа. Теорема доказана.

Пример 7.7. Вернёмся к примеру 7.4. Согласно (7.5.21) имеем

$$r_{-1}(z) = S(z) = \alpha z^2 + \alpha^2 z + \alpha^7,\ r_{-2}(z) = z^4. \qquad (7.8.51)$$

Следуя процедуре (7.6.28), (7.6.29) (7.8.43)–(7.8.47), получим

$$r_{-2}(z) = z^4,\ r_{-1}(z) = \alpha z^2+\alpha^2 z+\alpha^7,\ q_0(z) = \alpha^7 z^2+\alpha^4 z+\alpha^5,\ r_0(z) = \alpha^4. \tag{7.8.52}$$

Действительно, выполняя операции в поле $GF(3^2)$ (7.3.3), нетрудно проверить, что

$$z^4 = (\alpha z^2 + \alpha^2 z + \alpha^7)(\alpha^7 z^2 + \alpha^4 z + \alpha^5) + \alpha^{4.} \tag{7.8.53}$$

При $k = 0$ тривиальным образом впервые выполняется неравенство $\deg r_k \le t - 1$, где в нашем случае $t = 2$.

Имеем далее $u_{-2} = 0;\ u_{-1} = 1, \quad u_0 = q_0.u_{-1}+u_{-2} = \alpha^7 z^2+ \alpha^4 z + \alpha^5$. Согласно формуле (7.8.46) следует положить $\chi = \alpha^3$. Тогда получится свободный член многочлена $\sigma(z)$, и он будет равен $\sigma(0) = 1$. Окончательно получим многочлен локаторов ошибок $\sigma(z) = \alpha^2 z^2 + \alpha^7 z + 1$. Он совпадает с многочленом (7.5.22).

Пример 7.8. Перейдём к примеру 7.5. Из него получаем:

$$r_{-2}(z) = z^6,$$

$$\begin{array}{ll} r_{-1} = S(z) = \alpha^5 z^5 + z^4 + \alpha^5 z^3 + \alpha^3 z^2 + z + \alpha, & q_0(z) = \alpha^3 z + \alpha^2, \\ r_0 = \alpha^7 z^4 + \alpha z^3 + \alpha^2 z^2 + \alpha z + \alpha^7, & q_1(z) = \alpha^6 z + \alpha^6, \\ r_1 = \alpha^7 z^3 + \alpha z^2 + \alpha^6, & q_2(z) = z \\ r_2 = \alpha^2 z^2 + z^5 z + \alpha^7, & q_3(z) = \alpha^5 z + \alpha \\ r_3 = \alpha^3 z + \alpha^3, & q_4(z) = \alpha^7 z + \alpha \\ r_4 = \alpha^6 & \end{array} \tag{7.8.54}$$

Далее

$u_{-2}(z) = 0,\ u_{-1}(z) = 1$
$u_0(z) = q_0 u_{-1}(z) + u_{-2}(z) = \alpha^3 z + \alpha^2 + 0$
$u_1(z) = q_1(z)u_0(z) + u_{-1}(z) = (\alpha^6 z + \alpha^6)(\alpha^3 z + \alpha^2) + 1 = \alpha z^2 + \alpha^7 z + \alpha^4$
$u_2(z) = q_2(z)u_1(z) + u_0(z) = z(\alpha z^2 + \alpha^7 z + \alpha^4) + (\alpha^3 z + \alpha^2) = \alpha z^3 + \alpha^7 z^2 + \alpha^2 z + \alpha^2$

Видим,что впервые неравенство $\deg r_k \le (t - 1)$ выполняется при $k = 2$, так как $t = 3$. Следовательно,

$$\sigma(z) = \chi u_2(z).$$

Для того, чтобы выполнялось условие $\sigma(0) = 1$, положим $\chi = \alpha^6$. Тогда многочлен локаторов ошибок будет

$$\sigma(z) = \alpha^7 z^3 + \alpha^5 z^2 + z + 1,$$

что совпадает с многочленом (7.5.25).

7.9. Вывод ключевого уравнения на случай ошибок и стираний

Перейдём теперь к случаю, когда при передаче в канале происходят и ошибки и стирания. Вспоминая формулу (0.3.8) в разделе 0.3, положим, что произошло $l \leq (d-1)$ стираний и $t \leq (d-l-1)/2$ ошибок. Сначала объединим термином "искажение" оба явления, и ошибки и стирания. Заменим стёртые символы принятого вектора a, например, нулями и будем обращаться с ним, как с вектором, содержащим только ошибки (искажения!) с той лишь разницей, что локаторы l стертых символов нам известны заранее.

Синдром принятого вектора имеет вид

$$S = (S_1, S_2, \dots S_{d-1}),\ S_j \in GF(q^m) \tag{7.9.55}$$

По аналогии с (7.7.37) сопоставим вектору (7.9.55) многочлен

$$S(z) = \sum_{j=1}^{j=d-1} S_j z^{j-1} \tag{7.9.56}$$

Без ограничения общности преобразуем систему (7.7.36) в такую:

$$\begin{gathered}
S_1 = Y_1X_1 + Y_2X_2 + \ldots + Y_tX_t + Y_{t+1}X_{t+1} + Y_{t+2}X_{t+2} + \ldots + Y_{t+l}X_{t+l},\\
S_2 = Y_1X_1^2 + Y_2X_2^2 + \ldots + Y_tX_t^2 + Y_{t+1}X_{t+1}^2 + Y_{t+2}X_{t+2}^2 + \ldots + Y_{t+l}X_{t+l}^2,\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
S_{d-1} = Y_1X_1^{d-1} + Y_2X_2^{d-1} + \ldots + Y_tX_t^{d-1} + Y_{t+1}X_{t+1}^{d-1} + Y_{t+2}X_{t+2}^{d-1} + \ldots + Y_{t+s}X_{t+s}^{d-1},\\
Y_j \in GF(q), \quad X(j) \in GF(q^m).
\end{gathered} \tag{7.9.57}$$

Здесь $X_1, X_2, \dots, X_t$ – локаторы ошибок, $X_{t+1}, X_{t+2}, \dots, X_{t+l}$ – известные локаторы стираний и $Y_1, Y_2, \dots, Y_{t+l}$ – значения искажений.

Казалось бы, теперь остаётся поступить точно так же, как и в случае одних только ошибок, изменяя лишь индексы суммирования и умножения. В самом деле, из (7.9.56) и (7.9.57), изменив порядок суммирования и пользуясь формулой для суммы геометрической прогрессии, получим:

$$S(z) = z^{d-1} \sum_{i=1}^{t+l} Y_i X_i \frac{X_i^{d-1}}{X_i z - 1} - \sum_{i=1}^{i=t+l} \frac{Y_i X_i}{X_i z - 1}. \tag{7.9.58}$$

Полагая

$$\begin{aligned} \tilde{\sigma}(z) &= \prod_{i=1}^{i=t+l} (X_i z - 1) = \sum_{i=1}^{i=t+l} \sigma_i z^i. \\ \tilde{\omega}(z) &= \sum_{j=1}^{j=t+l} Y_j X_j \prod_{i=1, i\neq j}^{i=t+l} (X_i z - 1), \\ \tilde{\Phi}(z) &= \sum_{i=1}^{i=t+l} Y_i X_i^d \prod_{c=1, c\neq i}^{c=t+l} (X_c z - 1), \end{aligned} \tag{7.9.59}$$

после приведения всех дробей в (7.9.59) к общему знаменателю получим:

$$S(z) = z^{d-1} \frac{\tilde{\Phi}(z)}{\tilde{\sigma}(z)} - \frac{\tilde{\omega}(z)}{\tilde{\sigma}(z)} \tag{7.9.60}$$

Выражение (7.9.60), как и выше, называют *ключевым уравнением*. Придадим ему иной вид

$$S(z)\tilde{\sigma}(z) = z^{d-1}\tilde{\Phi}(z) - \tilde{\omega}(z) \tag{7.9.61}$$

Иначе говоря,

$$-\tilde{\omega}(z) \equiv \tilde{\sigma}(z) S(z) (\mathrm{mod}\, z^{d-1}), \tag{7.9.62}$$

что очень похоже на уравнение (7.7.42). Заметим, однако, что многочлен $\tilde{\sigma}(z)$ представляет собой произведение двух многочленов $\tilde{\sigma}(z) = \nu(z)\sigma(z)$, где $\sigma(z)-$ это по-прежнему многочлен неизвестных нам локаторов ошибок, а многочлен $\nu(z)-$это многочлен известных локаторов стираний. Действительно, множество корней многочлена $\tilde{\sigma}(z)$ состоит из двух подмножеств —

известных и неизвестных. Теперь введём в рассмотрение произведение $S(z)\nu(z) = \tilde{S}(z),\ \deg \tilde{S}(z) = 2t + l - 1$.

В новых обозначениях ключевое уравнение (7.9.62) примет вид

$$-\tilde{\omega}(z) \equiv \sigma(z)\tilde{S}(z)(\mathrm{mod} z^{d-1}) \tag{7.9.63}$$

Выражение $S(z)\nu(z) = \tilde{S}(z)$ иногда называют модифицированным синдромным многочленом.

7.10. Решение ключевого уравнения на случай ошибок и стираний

В этом разделе почти дословно повторяются рассуждения раздела 7.8.

Для решения ключевого уравнения воспользуемся соотношением (7.6.34), придав ему вид

$$r_k(z) \equiv (-1)^k(u_k(z)r_{-1}(z)(\mathrm{mod} r_{-2}(z)). \tag{7.10.64}$$

Сопоставляя сравнения (7.9.62) и (7.10.64), видим, что естественно положить

$$r_{-2} = z^{d-1},\ r_{-1} = \tilde{S}(z). \tag{7.10.65}$$

Тогда при некотором k решение сравнения

$$r_k(z) \equiv (-1)^k(u_k(z)\tilde{S}(z)(\mathrm{mod} z^{d-1}) \tag{7.10.66}$$

даст

$$\chi u_k(z) = \sigma(z) \tag{7.10.67}$$

и

$$(-1)^k\chi r_k(z) = \tilde{\omega}(z). \tag{7.10.68}$$

Учитывая, что степени остатков r_k строго убывают, значение k выбирается из следующих соображений. Степень ϵ многочлена $\sigma(z)$ должна быть наименьшей и во всяком случае не превосходить t. Это будет достигнуто, когда впервые выполнится неравенство $\deg r_k \leq t + l - 1$. В самом деле, степень правой части сравнения (7.10.66) равна $\deg \tilde{S}(z) + \epsilon$. И так как $\epsilon \leq t$, и $\deg \tilde{S}(z) = d + l - 2$, то сравнение $\deg \tilde{S}(z) + \epsilon \equiv \deg r_k(\mathrm{mod}(d-1))$ возможно при $\deg r_k \leq t + l - 1$. С другой стороны, при $\deg r_k \geq t$ окажется, что $\epsilon \geq t + 1$, вопреки

условию. При выполнении условия $\deg\sigma(z) \leq t$ из сравнения (7.9.62) немедленно следует $\deg\tilde{\omega}(z) \leq t+l-1$. Действительно, $\deg\omega(z) = (d-2) + \deg\sigma(z) - d + 1 \leq t + l - 1$. Этот факт находится в точном соответствии с формулой (7.9.59), из которой видно, что с необходимостью выполняется неравенство $\deg\tilde{\omega}(z) \leq t + l - 1$.

Как и в разделе 7.8, приведенное рассуждение можно обратить. Заметим, что слова "впервые выполняется неравенство $\deg r_k \leq t-1$" означают ни что иное, как "пока выполняется неравенство $\deg r_{k-1} \geq t$."

Делимость многочленов не зависит от константы поля, на которую умножен многочлен. Поэтому её выбор определяется специальными требованиями. В данном случае, как это следует из (7.9.59), элемент χ выбирается из условия $\sigma(0) = \sigma_0 = 1$. При этом, разумеется, по-прежнему будут выполнены сравнения (7.9.62) и неравенства $\deg\sigma(z) \leq t + l,\ \deg\tilde{\omega}(z) \leq t + l - 1$.

Теорема 7.10.1. *При* $\sigma(0) = 1$, *выполнении условий* $\deg\sigma(z) \leq t + l$ *и* $\deg\tilde{\omega}(z) \leq t + l - 1$ *многочлены* $\sigma(z)$, $\tilde{\omega}(z)$ *из* (7.10.67) *и* (7.10.68) *единственны, и многочлен* $\sigma(z)$ *имеет минимальную степень.*

Д о к а з а т е л ь с т в о. Пусть имеются два решения $\sigma_1(z)$, $\tilde{\omega}_1(z)$ и $\sigma_2(z)$, $\tilde{\omega}_2(z)$ сравнения (7.9.62). Это означает, что $-\tilde{\omega}_1(z) \equiv \sigma_1(z)\tilde{S}(z)(\mathrm{mod} z^{d-1})$ и $-\tilde{\omega}_2(z) \equiv \sigma(z)_2\tilde{S}(z)(\mathrm{mod} z^{d-1})$. Умножив первое сравнение на $\sigma_2(z)$, а второе на $\sigma_1(z)$, получим два сравнения, у которых правые части совпадают. Значит сравнимы их левые части: $\sigma_2(z)\tilde{\omega}_1(z) \equiv \sigma_1(z)\tilde{\omega}_2(z)(\mathrm{mod} z^{d-1}.)$ Однако степень произведения многочленов в этом сравнении не превосходит $d-1$. Поэтому сравнение переходит в равенство $\sigma_2(z)\tilde{\omega}_1(z) = \sigma_1(z)\tilde{\omega}_2(z)$. Отсюда $\sigma_2(z)/\sigma_1(z) = \tilde{\omega}_2(z)/\tilde{\omega}_1(z) = \mu(z)$, где $\mu(z)-$ некоторый многочлен. Если в решении (7.10.67) и (7.10.68) степень многочлена $\sigma(z)$ не минимальна, то из предыдущего следует, что $\sigma_2(z) = \mu(z)\sigma_1,\ \tilde{\omega}_2(z) = \tilde{\omega}_1(z)\mu(z)$ есть также решение сравнения (7.9.62). Собирая (7.6.30), (7.6.34) и (7.10.68), получим для $\tilde{\omega}_2(z)$:

$$\tilde{\omega}_2(z) = (-1)^k\chi r_k(z) = \chi v_k(z)z^{2t} + \chi^{-1}\sigma(z)\tilde{S}(z) \qquad (7.10.69)$$

Но так как $\tilde{\omega}_2(z) = \mu(z)\tilde{\omega}_1$, то

$$\mu(z)\tilde{\omega}_1(z) = \chi v_{k(z)}z^{2t} + \mu(z)\sigma_1(z)\tilde{S}(z) \qquad (7.10.70)$$

и $\mu(z)$ делит $v_k(z)$. Однако согласно (7.10.67)

$$\chi u_k(z) = \mu(z)\sigma(z), \tag{7.10.71}$$

и $\mu(z)$ делит также $u_k(z)$, что в силу утверждения 7.6.1 может быть, только когда $\mu(z)$ константа. Теорема доказана.

Читатель помнит, что, получив многочлен локаторов ошибок $\sigma(z)$, следует найти его корни, обратные значения которых и дадут ошибочные компоненты принятого вектора. В случае стираний к локаторам ошибок добавятся еще l локаторов стираний, которые указаны явно при приёме, и назначаются корнями многочлена $\nu(z)$.

Остаётся вычислить значения $Y_i,\ i = 1, 2, \ldots, t+l$, ошибок.

Пример 7.9. Пусть передан вектор

$$u = (\alpha^6, \alpha^5, \alpha^2, 1, \alpha^3, \alpha^7, \alpha, 0).$$

кода РС над $GF(3^2),\ d = 7$, и принят вектор

$$v = (0, 0, *, 1, \alpha^3, \alpha^7, \alpha, *).$$

Заранее видим, что значения ошибок суть

$$Y_1 = -\alpha^6 = \alpha^2,\ Y_2 = -\alpha^5 = \alpha.$$

Заменим стирания нулями. Новый вектор будет иметь вид:

$$v^{'} = (0, 0, 0, 1, \alpha^3, \alpha^7, \alpha, 0).$$

Он совпадает с вектором (7.5.23), для которого в примере 7.5 посредством матрицы (7.5.24) уже вычислены элементы S_i синдрома. Отсюда — известный нам синдромный многочлен

$$S(z) = \alpha^5 z^5 + z^4 + \alpha^5 z^3 + \alpha^3 z^2 + z + \alpha.$$

Локаторы стираний: $\alpha^2,\ \alpha^7,$ многочлен локаторов стираний $\nu(z) = (1 - \alpha^2 z)(1 - \alpha^7 z) = \alpha z^2 + z + 1$. Таким образом, модифицированный синдромный многочлен $\tilde{S}(z) = S(z)\nu(z)$ будет:

$$\begin{aligned}\tilde{S}(z) = (\alpha^6 z^7 + \alpha^5 z^4 + \alpha^3 z^3 + \alpha^7 z^2 + \alpha^7 z + \alpha)(\mathrm{mod}\, z^6) = \\ = \alpha^5 z^4 + \alpha^3 z^3 + \alpha^7 z^2 + \alpha^7 z + \alpha.\end{aligned}$$

Имеем

$$r_{-2} = z^6,\ r_{-1} = \tilde{S}(z)(\mathrm{mod} z^6).$$

Далее

$$z^6 = (\alpha^5 z^4 + \alpha^3 z^3 + \alpha^7 z^2 + \alpha^7 z + \alpha)(\alpha^3 z^2 + \alpha^5 z + \alpha^2) + (\alpha^7 z + \alpha^5).$$

$$q_0(z) = \alpha^3 z^2 + \alpha^5 z + \alpha^2,\ r_0(z) = \alpha^7 z + \alpha^5.$$

Видим,что впервые неравенство $\deg r_k(z) \leq (t + l - 1)$ выполняется при $k = 0$, так как $t = 2,\ l = 2$. Согласно процедуре вычислений,

$$u_{-2}(z) = 0,\ u_{-1}(z) = 1,$$
$$u_0(z) = q_0(z)u_{-1}(z) + u_{-2}(z) = \alpha^3 z^2 + \alpha^5 z + \alpha^2.$$

Следовательно,

$$\sigma(z) = \chi u_0(z).$$

Для того, чтобы выполнялось условие $\sigma(0) = 1$, положим $\chi = \alpha^6$. Тогда

$$\sigma(z) = \alpha z^2 + \alpha^3 z + 1.$$

Корни этого многочлена суть $z_1 = 1,\ z_2 = \alpha^7$. Локаторы ошибок $X_1 = 1,\ X_2 = \alpha$.

Перейдём к нахождению значений ошибок и стираний. Напомним, что выше был введён в обращение многочлен

$$\tilde{\sigma}(z) = \prod_{i=1}^{i=t+l} (X_i z - 1) = \sum_{i=1}^{i=t+l} \sigma_i z^i, \qquad (7.10.72)$$

корнями которого являются величины, обратные локаторам ошибок и локаторам стираний. Объединим оба эти явления под названием "искажения" , а многочлен (7.10.72) назовём многочленом локаторов искажений и будем искать значения искажений. Решение этой задачи достигается посредством многочлена $-\tilde{\omega}(z) = \tilde{\sigma}(z)\tilde{S}(z)(\mathrm{mod} z^{d-1})$, который мы назовём многочленом значений искажений.

Возьмём производную от $\tilde{\sigma}(z)$:

$$\tilde{\sigma}'(z) = \sum_{j=1}^{j=t+l} X_j \prod_{i=1, i\neq j}^{i=t+l} (X_i z - 1) = \sum_{i=1}^{i=t+l} i\sigma_i z^{i-1}. \qquad (7.10.73)$$

Подставим в (7.9.59) и (7.10.73) любой корень X_j^{-1} многочлена локаторов искажений. Тогда получим

$$\tilde{\omega}(X_j^{-1}) = Y_i^{-1}\tilde{\sigma}'(X_j^{-1}). \qquad (7.10.74)$$

Действительно, в каждой из сумм (7.9.59) и (7.10.73) останется в точности по одному слагаемому, именно

$$X_j \prod_{i=1, i\neq j}^{i=t+l} (X_i X_j - 1) = \tilde{\sigma}'(X_j^{-1}) \qquad (7.10.75)$$

в сумме (7.10.73), и

$$Y_j X_j \prod_{i=1, i\neq j}^{i=t+l} (X_i X_j - 1) = \tilde{\omega}(X_j^{-1}) \qquad (7.10.76)$$

в сумме (7.9.59).

Остальные обращаются в нуль. Из (7.10.74), (7.10.75) следует

$$Y_i = \frac{\tilde{\omega}(X_i^{-1})}{\tilde{\sigma}'(X_i^{-1})}. \qquad (7.10.77)$$

Знаменатель последней дроби не обращается в нуль, так как многочлен $\tilde{\sigma}(z)$ не имеет кратных корней.

Продолжим пример 7.9.

$$\tilde{\sigma}(z) = \sigma(z)\nu(z) = (\alpha z^2 + \alpha^3 z + 1)(\alpha z^2 + z + 1) = \\ = \alpha^2 z^4 + \alpha^6 z^3 + \alpha^6 z^2 + \alpha^5 z + 1. \qquad (7.10.78)$$

$$\tilde{\sigma}'(z) = \alpha^2 z^3 + \alpha^2 z + \alpha^5. \qquad (7.10.79)$$

$$-\tilde{\omega}(z) = \tilde{\sigma}(z)\tilde{S}(z)(\mathrm{mod}\, z^{d-1}) = \\ = (\alpha^5 z^4 + \alpha^3 z^3 + \alpha^7 z^2 + \alpha^7 z + \alpha)(\alpha^3 z^2 + \alpha^5 z + \alpha^2)(\mathrm{mod}\, z^6) = \\ = \alpha^7 z^3 + \alpha^5 z^2 + \alpha z + \alpha.$$

Отсюда $\omega(z) = \alpha^3 z^3 + \alpha z^2 + \alpha^5 z + \alpha^5$.

Так как локаторы искажений

$$X_1 = 1,\ X_2 = \alpha,\ X_3 = \alpha^2,\ X_4 = \alpha^7,$$

то окончательно:

$$Y_1 = \frac{\tilde{\omega}(X_1^{-1})}{\tilde{\sigma}'(X_1^{-1})} = \frac{\alpha^6}{\alpha^4} = \alpha^2 = -\alpha^6,\ Y_2 = \frac{\tilde{\omega}(X_2^{-1})}{\tilde{\sigma}'(X_2^{-1})} = \frac{1}{\alpha^7} = \alpha = -\alpha^5,$$

$$Y_3 = \frac{\tilde{\omega}(X_3^{-1})}{\tilde{\sigma}'(X_3^{-1})} = \frac{\alpha^3}{\alpha^5} = -\alpha^2,\ Y_4 = \frac{\tilde{\omega}(X_4^{-1})}{\tilde{\sigma}'(X_4^{-1})} = \frac{0}{\alpha^4} = 0.$$

Остаётся сравнить полученный результат с парой векторов – отправленным вектором u и принятым вектором v.

Заметим, что замена стёртого символа c_i нулём означает вычитание $c_i - c_i$. Поэтому, восстановление данного стирания выполняется заменой символа стирания величиной c_i, в точности равной дроби (7.10.77) с обратным знаком.

Читателю предлагается довести до конца процедуру по вычислению значений ошибок примера 7.5., удалив сначала знак тильды в формулах, начиная с (7.9.59), так как упомянутый пример имеет дело только с ошибками, а не стираниями.

Декодирование посредством алгоритма Эвклида интересно не только из-за простоты понимания всей процедуры. Наибольший интерес представляет сама возможность применения алгоритма Эвклида. Еще раз подчеркнём замечание, приведённое в придисловии ко второму изданию. Фундаментальный алгебраический факт, известный с незапамятных времён, через века оказался как-будто специально приготовленным для решения ключевого уравнения, которое появилось совсем недавно ради вполне практической задачи: надёжной передачи инфомации.

7.11. Коды РС и построение каскадных кодов

Каскадные коды — это замечательная глава теории кодирования. Они занимают особое место в науке о построении эффективных кодов, лежащих в области наилучших параметров, и благодаря этому нашли достойное применение на практике. Сколько-нибудь серьезное и систематическое изложение всех достижений теории каскадных кодов не может уместиться в

рамках даже большой части общего руководства по теории кодирования. Поэтому здесь будет указан основной принцип построения каскадных кодов. Заинтересовавшийся читатель сможет обратиться к таким источникам, как [4, 7, 8].

Без потери общности и ради простоты ограничимся случаем $q = 2^m$. Пусть $a = (a_1, a_2, \ldots, a_k)$ есть вектор информационных символов над полем $GF(2^m)$. Он кодируется в вектор некоторого линейного кода с параметрами n_2, k_2, d_2. Этот код называют *внешним* кодом. Чаще всего внешним (n_2, k_2, d_2)-кодом бывает код РС, и потому $n_2 \leq 2^m - 1$. Представим каждый символ внешнего кода в виде вектора длины m над $GF(2)$, а затем, приняв эти m символов за информационные, их кодируют в вектор линейного кода с параметрами n_1, $k_1 = m$, d_1. Этот код называют *внутренним* кодом. Получится код над $GF(2)$ с параметрами $N = n_1 n_2$, $K = k_1 k_2$, $D = d_1 d_2$. По-видимому, первые два параметра очевидны. Покажем справедливость равенства $D = d_1 d_2$. Действительно, два вектора внешнего кода различаются не менее, чем в d_2 компонентах, а каждая пара символов внутреннего кода в каждой из этих компонент различается не менее, чем в d_1 компонентах. Полученный код называется *суперкодом.*

Возможно иное, эквивалентное, описание каскадного кода. Пусть имеется двоичная информационная последовательность длины $K = k_1 k_2$. Она разбивается на k_2 отрезков длины k_1. Эти отрезки рассматриваются как элементы поля $GF(2^{k_1})$. Последовательность длины k_2 над полем $GF(2^{k_1})$ будет ни чем иным, как информационной последовательностью, которая кодируется в кодовый вектор кода длины n_2 над тем же полем с кодовым расстоянием d_2. Пусть таким вектором будет $A = (\alpha_1, \alpha_2, \ldots, \alpha_{n_2})$, $\alpha_i \in GF(2^k)$. Кодирование заканчивается тем, что каждый элемент α_i кодируется внутренним кодом в (двоичный) вектор β_i длины n_1. Вектор $B = (\beta_1, \beta_2, \ldots, \beta_{n_2})$ отправляется в канал связи.

Как видим, снова параметрами суперкода будут

$$N = n_1 n_2,\ K = k_1 k_2,\ D = d_1 d_2.$$

Вспомним теперь (раздел 4.5) вывод границы Варшамова — Гилберта посредством построения проверочной матрицы линейного кода, а значит, и самого кода. Это построение велось методом перебора, и его объем выражался числом порядка 2^n, где n есть длина кода. Иначе говоря объем перебора при таком построении кода зависит экспоненциально от длины n кода.

Рассмотрим случай, когда $k_1/n_1 = 1/x$, и x не слишком велико. Иными словами, пусть скорость передачи внутреннего кода не слишком мала. Так как $n_2 = 2^m - 1 = 2^{k_1 - 1}$, то оказывается, что длина n_1 внутреннего кода есть величина порядка $x \log_2 n_2$. Это значит, что даже если внутренний код получается методом перебора, то его объем выражается величиной порядка $2^{n_1} \sim 2^{x \log_2 n_2} = n_2^x$. Такой перебор может считаться вполне приемлемым: его объем по порядку не превосходит небольшой степени длины внешнего кода (т.е. даже не суперкода). И вовсе не зависит от нее экспоненциально.

Если учесть, что внешний код является кодом с максимально достижимым кодовым расстоянием d_2, то становятся ясными весьма высокие качества каскадных кодов.

В литературе по теории кодирования показано, что класс всех каскадных кодов содержит коды, лежащие на границе Варшамова—Гилберта.

Еще более результативной является идея *обобщенных каскадных кодов*. Они представляют собой сочетание нескольких каскадных кодов, варьирование параметрами которых дает возможность получать коды высокого качества в широком диапазоне длин, скоростей и корректирующей способности.

Декодирование каскадного кода выполняется следующим образом. Принятая последовательность поступает на декодер внутреннего кода. Каждый отрезок длины n_1 один за другим рассматривается как, быть может искаженный, вектор внутреннего кода над $GF(2)$, и декодер исправляет в нем все ошибки кратности $(d_1 - 1)/2$ и менее.

Если в некотором отрезке произошло более, чем $(d_1 - 1)/2$ ошибок, то декодирование может оказаться неверным. Имеет место ошибка декодирования. После завершения работы внутреннего декодера принятая последовательность поступает на внешний декодер. Каждый отрезок длины n_1 рассматривается внешним декодером как элемент поля $GF(2^m)$. Ошибочно декодированные внутренним декодером отрезки длины n_1 представляют собой искаженные компоненты вектора внешнего кода . Если таких компонент окажется не более, чем $(d_2 - 1)/2$, то принятый вектор декодируется внешним декодером верно. Декодирование заканчивается.

7.12. Задачи к главе 7

7.1. Описать (15, 13)-код Рида — Соломона над полем $GF(2^4)$, определив его длину, порождающий многочлен и число исправ-

ляемых ошибок.

7.2. Процедура удлинения кода РС в разделе 7.4 основана на списке корней порождающего многочлена в (7.4.8). Как изменятся рассуждения об удлинении, если к списку корней будет принадлежать 1?

7.3. Передача ведется кодом РС над $GF(3^2)$, порождающий многочлен которого имеет своими корнями α^i, $i = 1, \ldots, 6$. Поле построено по модулю многочлена $x^2 + x + 2$. Принят вектор $v = (\alpha, \alpha^6, \alpha^4, \alpha^2, 0, \alpha^3, \alpha^7, \alpha)$. Исправить ошибки.

7.4. Доказать или опровергнуть, что для отыскания значений ошибок $Y_1, Y_2, \ldots, Y_t$ можно воспользоваться любыми t подряд идущими уравнениями (6.9.48).

7.5. Показать, что в векторе циклического (n, k)- кода любые k идущих подряд разрядов образуют информационную совокупность.

Глава 8.

Сводка границ

Приведем наиболее часто употребимые границы.

8.1. Верхние границы

Граница Синглтона

$$d - 1 \leq n - k. \tag{8.1.1}$$

Граница Плоткина.

$$d \leq n\frac{q^{k-1}(q-1)}{q^k - 1}. \tag{8.1.2}$$

Граница Хэмминга на случай нечетного d

$$\frac{k}{n} \leq 1 - \frac{1}{n}\log_q \sum_{i=0}^{t} C_n^i (q-1)^i. \tag{8.1.3}$$

Граница Хэмминга на случай четного d

$$\frac{k}{n} \leq 1 - \frac{1}{n}\log_q \sum_{i=0}^{t} C_{n-1}^i (q-1)^i - \frac{1}{n}. \tag{8.1.4}$$

Граница Бассалыго—Элайеса. Пусть $A(n,\, d)-$ есть максимально возможное число кодовых векторов над $GF(q)$ кода длины n с минимальным расстоянием $d = 2t + 1$. Тогда

$$A(n,\, d) \leq \frac{q^n}{C_n^r (q-1)^r}, \tag{8.1.5}$$

где

$$r = \left(1 - \frac{1}{q}\right)\left(1 - \sqrt{1 - \frac{2qt}{(q-1)n}}\right) n. \tag{8.1.6}$$

8.2. Нижняя граница

Граница Варшамова—Гилберта
Если выполняется неравенство

$$\frac{k}{n} \le 1 - \frac{1}{n}\log_q \sum_{i=0}^{d-2} C_{n-1}^i (q-1)^i, \tag{8.2.7}$$

то существует линейный код с такими параметрами.

8.3. Асимптотические границы

Без труда находим, что при $n \to \infty$ граница Синглтона дает

$$\frac{d}{n} \le 1 - \frac{k}{n}.$$

Перейдем к выводу асимптотической границы Плоткина. Из (8.1.2) получается

$$q^{k-1}(qd + n - nq) \le d.$$

Отсюда при $qd + n - nq > 0$ имеем (см. задачу 5.14):

$$q^k = B(n, d) \le \frac{qd}{qd + n - nq}.$$

Для случая $n = (qd-1)/(q-1)$ (благодаря предположению о выполнении этого условия границу Плоткина называют границей кодов с большим кодовым расстоянием) получается

$$B\left(\frac{qd-1}{q-1}, d\right) \le qd. \tag{8.3.8}$$

Применяя a раз процедуру укорочения кода (см. задачу 5.14), получим

$$B(n,\, d) \le q^a B(n-a,\, d). \tag{8.3.9}$$

Из (8.3.8) и (8.3.9) при $n - a = (qd-1)/(q-1)$

$$B(n,\, d) \le q^a B\left(\frac{qd-1}{q-1},\, d\right) \le q^{n-(qd-1)/(q-1)} qd. \tag{8.3.10}$$

Вспомним, что для исходного кода с k информационными символами, т.е. до применения процедуры укорочения, должно быть $B(n,\, d) = q^k$. Подставляя это значение в неравенство (8.3.10) и логарифмируя последнее по q, получим

$$\frac{k}{n} \le 1 - \frac{qd-1}{n(q-1)} + \frac{1+\log_q d}{n}. \tag{8.3.11}$$

Пренебрегая в этом неравенстве при $n \to \infty$ последним слагаемым и удаляя -1 из числителя первой дроби в правой части неравенства, получим окончательно

$$\frac{k}{n} \le 1 - \frac{qd}{n(q-1)}.$$

Это и есть асимптотическа форма границы Плоткина.

Для получения асимптотики остальных границ следует оценить числа сочетаний и суммы чисел сочетаний.

В основе этих оценок лежат неравенства Стирлинга для факториала:

$$\sqrt{2\pi N}\left(\frac{N}{e}\right)^N \exp\left(\frac{1}{12N} - \frac{1}{360N^3}\right) < N! < \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \exp\frac{1}{12N}. \tag{8.3.12}$$

Пользуясь ими, получим

$$C_n^{\lambda n} = C_n^{\mu n} = \frac{n!}{(\lambda n)!(\mu n)!} <$$

$$< \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\frac{1}{12n}}{\sqrt{2\pi\lambda n}\left(\frac{\lambda n}{e}\right)^{\lambda n} \exp\left(\frac{1}{12\lambda n} - \frac{1}{360(\lambda n)^3}\right)\sqrt{2\pi\mu n}\left(\frac{\mu n}{e}\right)^{\mu n} \exp\left(\frac{1}{12\mu n} - \frac{1}{360(\mu n)^3}\right)} =$$

$$= \frac{1}{\sqrt{2\pi\lambda\mu n}\lambda^{\lambda n}\mu^{\mu n}} \exp A,$$

где

$$0 < \lambda, \mu < 1,\ \lambda + \mu = 1,$$

и

$$A = \frac{1}{12n} - \frac{1}{12\lambda n} - \frac{1}{12\mu n} + \frac{1}{360(\lambda n)^3} + \frac{1}{360(\mu n)^3}.$$

Все приведенные выше выражения симметричны относительно λ, μ. Пусть для определенности и без потери общности $\lambda \geq \mu$. Тогда

$$\frac{1}{360(\lambda n)^3} \leq \frac{1}{360(\mu n)^3} \leq \frac{1}{360\mu n},\ \frac{1}{12n} < \frac{1}{12\lambda n}.$$

Поэтому

$$A \leq \frac{1}{12n} - \frac{1}{12\lambda n} - \frac{1}{12\mu n} + \frac{1}{180\mu n} < 0.$$

Отсюда получается верхняя граница

$$C_n^{\lambda n} = \frac{1}{\sqrt{2\pi\lambda\mu n}\lambda^{\lambda n}\mu^{\mu n}} \exp A < \frac{1}{\sqrt{2\pi\lambda\mu n}\lambda^{\lambda n}\mu^{\mu n}}.$$

Воспользуемся теперь упрощенными неравенствами Стирлинга.

$$\sqrt{2\pi N}(\frac{N}{e})^N < N! < \sqrt{2\pi N}(\frac{N}{e})^N \exp \frac{1}{12N}. \tag{8.3.13}$$

$$C_n^{\lambda n} = C_n^{\mu n} = \frac{n!}{(\lambda n)!(\mu n)!} < \frac{\sqrt{2\pi n}(\frac{n}{e})^n \exp -(\frac{1}{12\lambda n} + \frac{1}{12\mu n})}{\sqrt{2\pi\lambda n}(\frac{\lambda n}{e})^{\lambda n}\sqrt{2\pi\mu n}(\frac{\mu n}{e})^{\mu n})}.$$

Положим $\mu n \geq 3,\ \lambda n \geq 1$, (или, наоборот, лишь бы одно из этих произведений было не меньше трех). Тогда

$$\frac{1}{12\lambda n} + \frac{1}{12\mu n} \leq \frac{1}{12} + \frac{1}{36} = \frac{1}{9},$$

и потому

$$\exp -(\frac{1}{12\lambda n}+\frac{1}{12\mu n})\geq \exp(-\frac{1}{9})=0.895\geq \frac{\sqrt{\pi}}{2}=0.887.$$

Собирая найденные результаты, получим нижнюю границу

$$C_n^{\lambda n}=C_n^{\mu n}>\frac{1}{\sqrt{8\lambda\mu n}\lambda^{\lambda n}\mu^{\mu n}}.$$

Примем теперь во внимание,что

$$\frac{1}{\lambda^{\lambda n}\mu^{\mu n}}=q^{-\log_q(\lambda^{\lambda n}\mu^{\mu n})}=q^{nH_q(\lambda)},$$

где

$$H_q(\lambda)=-\lambda\log_q\lambda-(1-\lambda)\log_q(1-\lambda). \tag{8.3.14}$$

(Иногда функцию $H_q(x)$ определяют, как

$$H_q(x)=x\log_q(q-1)-x\log_q x-(1-x)\log_q(1-x);$$

ее называют энтропией q-ичного симметричного канала с вероятностью ошибки x. Здесь же символ $H_q(x)$ употребляется в смысле (8.3.14)).

Окончательно имеем

$$\frac{1}{\sqrt{8\lambda\mu n}}q^{nH_q(\lambda)}<C_n^{\lambda n}=C_n^{\mu n}<\frac{1}{\sqrt{2\pi\lambda\mu n}}q^{nH_q(\lambda)} \tag{8.3.15}$$

Читателю предлагается убедиться, что неравенства справедливы и на случай λn, $\mu n<3$. Например, если $\lambda n=\mu n=1$, то $\lambda n+\mu n=2, (\lambda+\mu)n=2$, $n=2$, $\lambda=\mu=1/2$. Подставив эти значения в (8.3.15), получим $2=C_2^1<\frac{4}{\sqrt{\pi}}$.

Оценим теперь сумму $\sum_{i=\lambda n}^n C_n^i$, положив λn целым числом при условии $1/2<\lambda<1$.

Для этого умножим ее на $2^{ln\lambda}$, где l есть некоторое положительное число, и рассмотрим цепочку неравенств

$$2^{ln\lambda}\sum_{i=\lambda n}^n C_n^i\leq\sum_{i=\lambda n}^n 2^{li}C_n^i\leq\sum_{i=0}^n 2^{li}C_n^i=(1+2^l)^n. \tag{8.3.16}$$

Разделив выражение (8.3.16) на величину $2^{ln\lambda}$, а затем положив $l = \log_2(\lambda/(1-\lambda))$, получим

$$\begin{gathered}\sum_{i=\lambda n}^{n} C_n^i \le (2^{-l\lambda} + 2^{l(1-\lambda)})^n = \\ = (2^{-\lambda \log_2(\lambda/(1-\lambda))} + 2^{(1-\lambda)\log_2(\lambda/(1-\lambda)}))^n = \\ = ((\lambda/(1-\lambda))^{-\lambda} + (\lambda/(1-\lambda))^{1-\lambda})^n = \\ = ((\lambda/(1-\lambda))^{-\lambda}(1-1/(1-\lambda)))^n == (\lambda^{-\lambda}(1-\lambda)^{-(1-\lambda)})^n = \\ = q^{(-\lambda \log_q \lambda - (1-\lambda)\log_q(1-\lambda))n} = q^{nH_q(\lambda)}.\end{gathered}$$

Благодаря (8.3.16), при $0 < \mu < 1/2$ выполняется также и

$$\sum_{i=0}^{\mu n} C_\mu^i < q^{nH_q(\mu)}.$$

Тривиальным образом

$$\frac{1}{\sqrt{8(1-\mu)\mu n}} q^{nH_q(\mu)} < C_n^{\mu n} < \sum_{i=0}^{\mu n} C_\mu^i.$$

Окончательно

$$\frac{1}{\sqrt{8(1-\mu)\mu n}} q^{nH_q(\mu)} < \sum_{i=0}^{\mu n} C_\mu^i < q^{nH_q(\mu)}$$

Теперь асимптотические границы выглядят следующим образом:

Граница Хэмминга

$$\frac{k}{n} \le 1 - \frac{1}{n}\log_q \sum_{i=0}^{t} C_n^i (q-1)^i \le 1 - \frac{t}{n}\log_q(q-1) - \frac{1}{n}\log_q \sum_{i=0}^{t} C_n^i \le$$

$$\le 1 - \frac{t}{n}\log_q(q-1) - \frac{1}{n}\log_q q^{nH_q(\frac{t}{n})} = 1 - \frac{d}{2n}\log_q(q-1) - H_q\left(\frac{d}{2n}\right). \tag{8.3.17}$$

На случай четного d граница совпадает с (8.3.17).

Граница Бассалыго — Элайеса

$$\frac{k}{n} \le 1 - \left(1 - \frac{1}{q}\right)\left(1 - \sqrt{1 - \frac{qd}{(q-1)n}}\right)\log_q(q-1) -$$

$$-H_q\left(\left(1-\frac{1}{q}\right)\left(1-\sqrt{1-\frac{qd}{(q-1)n}}\right)\right).$$

Не сообщая вывода неасимптотической формы "границы четырех" (граница Мак-Элиса— Родемича — Рамсея—Велча), приведем ее асимптотическую форму[1]

$$\frac{k}{n} \le x\log_q(q-1) + H_q(x),$$

где

$$x = \frac{q-1-\frac{d}{n}(q-2)-2\sqrt{(q-1)\frac{d}{n}(1-\frac{d}{n})}}{q}.$$

Граница Варшамова—Гилберта.

Если выполняется условие

$$\frac{k}{n} \le 1-\frac{d}{n}\log_q(q-1) - H_q\left(\frac{d}{n}\right),$$

то существует код с параметрами $(n,\, k,\, d)$.

[1] С подробностями читатель может ознакомиться по книгам [7, 12].

Глава 9.

Регистры сдвига с линейными обратными связями

9.1. Элементарные устройства

Регистр сдвига действует в режиме дискретного времени и состоит из элементарных устройств трех видов:

1. Двухвходовый сумматор. Если на входы поступают величины a и b, то на выходе в тот же момент времени появляется сумма $a + b$ (рис. 9.1a).

2. Устройство умножения на константу поля $GF(q)$. Оно имеет один вход и один выход. Если устройство производит умножение на константу $\alpha \in GF(q)$, и на вход поступает элемент $\beta \in GF(q)$, то в тот же момент времени на выходе появляется произведение $\alpha\beta$. В случае поля $GF(2)$ умножение на константу 1 есть прямое соединение, а умножение на константу 0 есть отсутствие соединения, разрыв (рис. 9.1b.)

Сложение и умножение выполняются в поле $GF(q)$.

3. Элемент задержки на такт. Этот элемент как раз и осуществляет сдвиг. Значение на его выходе в момент времени $t+1$ равно значению на его входе в момент времени t, т.е. в предыдущий момент. Этот элемент есть также элемент памяти. Он "помнит" , что было на его входе, или что то же, что было его содержимым в предыдущий момент времени. (рис. 9.1 c.)

Нас совершенно не интересуют электрические характеристики перечисленных устройств. Их различные соединения открывают возможности графического выполнения чисто математических операций. Начнем с простейших.

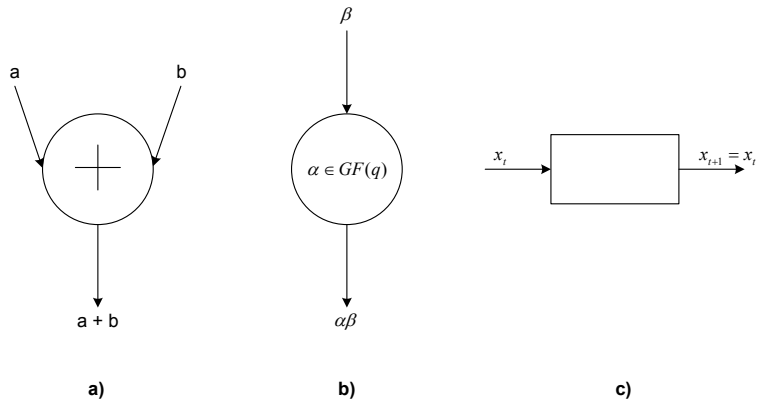

Рис. 9.1. a) сумматор; b) элемент умножения на константу поля; c) элемент задержки.

9.2. Вычисления в полях Галуа

Рассмотрим устройство, изображенное на рис. 9.2.

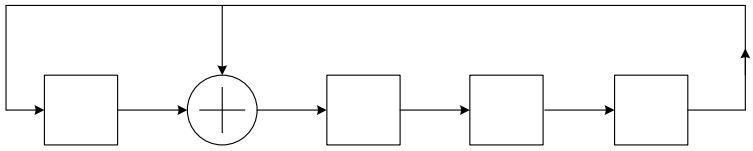

Рис. 9.2. Генератор поля $GF(2^4)$ по модулю многочлена $x^4 + x + 1$.

Во-первых, оно содержит четыре элемента задержки, которые составляют регистр сдвига. В каждом из элементов задержки (памяти) может содержаться одно из значений 0 или 1. В каждый момент времени содержимое этого регистра может рассматриваться как векторное изображение элемента поля $GF(2^4)$.

Поместим в регистр вектор (1000) и будем производить его последовательные сдвиги слева направо, как это показано стрелками. В результате первых трех сдвигов содержимое регистра будет последовательно меняться следующим образом: (0100), (0010), (0001). В результате следующего сдвига содержимое регистра станет (1100). Читатель уже вспомнил, что именно так

строилось поле $GF(2^4)$ (3.4.11) по модулю многочлена x^4+x+1. Как только происходит очередной сдвиг единицы, находящейся в последнем элементе задержки, она прибавляется к символам, находящимся в первом и втором элементах задержки. Это в точности соответствует формуле $x^4 = x+1$. Если $\alpha^4+\alpha+1 = 0$, т.е. α есть примитивный элемент поля, то излагаемый процесс работы регистра сдвига есть возведение α в последовательные степени и построение мультипликативной группы поля $GF(2^4)$.

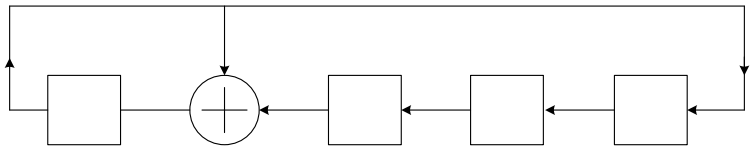

Рис. 9.3. Генератор поля $GF(2^4)$ по модулю многочлена $x^4 + x + 1$. Показатели степеней порождающего элемента α убывают.

Рассмотрим также устройство, изображенное на рис. 9.3. Как и выше, поместим в регистр вектор (1000). Будем производить его последовательные сдвиги справа налево, как это показано стрелками. В результате первого сдвига содержимым регистра будет вектор (1001), в котором читатель узнает элемент $\alpha^{14} = \alpha^{-1} \in GF(2^4)$. Устройство реализует построение поля $GF(2^4)$ в порядке убывания степеней элемента α.

Совместное применение обоих устройств может выполняться следующим образом.

Поместим в устройство рис. 9.2 элемент α^i, а в устройство рис. 9.3. — элемент α^j. Одновременные сдвиги обоих регистров приведут к тому, что в регистре рис. 9.2 показатели степеней α, будут расти, а в регистре рис. 9.3 — убывать. Когда в регистре рис. 9.3 наблюдатель увидит элемент $1=(1000)=\alpha^0$, он будет знать, что произошло в точности j сдвигов, а значит, содержимое регистра рис. 9.2 умножится на α^j и станет равным α^{i+j}. Так происходит умножение двух произвольных элементов поля Галуа.

Пусть теперь требуется разделить элемент α^i на элемент α^j. Поместим α^{-j} в регистр устройства рис. 9.2, а элемент α^i — в регистр рис. 9.3. Одновременные сдвиги обоих регистров приведут к тому, что, когда в регистре рис. 9.2 окажется вектор $1=(1000)=\alpha^0$, в регистре рис. 9.3 будет вектор α^{i-j}. Читатель без труда проведет самостоятельно соответствующее рассужде-

ние. Регистры сдвига с линейными обратными связями находят более значительное практическое применение.

9.3. Умножение и деление многочленов

Рассмотрим регистр сдвига, показанный на рис. 9.4.

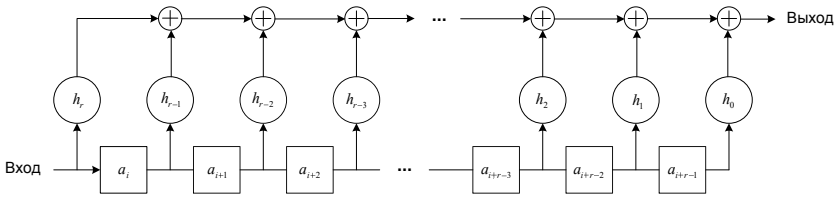

Рис. 9.4. Регистр сдвига для умножения на многочлен $h(x)$.

Он реализует умножение произвольного многочлена

$$a(x) = a_0 + a_1 x + \ldots + a_{k-1}x^{k-1} + a_k x^k \tag{9.3.1}$$

над $GF(q)$ на фиксированный многочлен

$$h(x) = h_0 + h_1 x + \ldots + h_{r-1}x^{r-1} + h_r x^r \tag{9.3.2}$$

с коэффициентами из того же поля. Все элементы умножения реализуют умножение на коэффициенты h_i, $i = k, k-1, \ldots, 0$ многочлена $h(x)$. Перед началом процесса умножения элементы задержки содержат нули. В начальный момент времени на вход регистра поступает коэффициент a_k, и на выходе тотчас появляется произведение $a_k h_r$, т.е. коэффициент при x^{k+r} многочлена $a(x)h(x)$. В следующий момент времени на вход поступает a_{k-1}, и на выходе появляется сумма $a_k h_{r-1} + a_{k-1} g_g$, т.е. коэффициент при x^{r+k-1}. Процесс продолжается до тех пор, пока последний коэффициент a_0 не появится на выходе последнего элемента задержки, а на выходе регистра — $a_0 h_0$, т.е. младший коэффициент произведения многочленов.

Вслед за последним коэффициентом a_0 многочлена $a(x)$ на вход регистра поступают нули, которые заполняют регистр к моменту появления на его выходе коэффициента $a_0 h_0$. Умножение закончено.

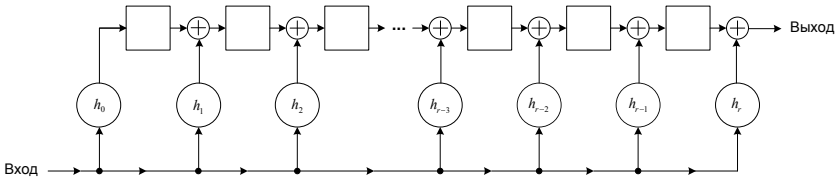

Рис. 9.5. Регистр сдвига для умножения на многочлен $h(x)$. Сумматоры встроены в регистр

Умножение произвольного многочлена $a(x)$ на фиксированный многочлен $h(x)$ можно выполнить и на другом регистре, который показан на рис. 9.5.

Его отличие от предыдущего состоит в том, что сумматоры вынесены из линии обратной связи и встроены в регистр между элементами задержки, а коэффициенты многочлена $h(x)$ следуют в порядке возрастания, а не убывания индексов.

В первый момент времени, когда содержимое регистра равно нулю, старший коэффициент a_k многочлена $a(x)$, поступивший на вход регистра, немедленно и одновременно умножается на все коэффициенты h_i. Но на выходе в этот момент времени появляется только старший коэффициент $a_k h_r$ произведения многочленов. После первого сдвига в элементах задержки содержатся произведения $a_k h_0$, $a_k h_1$, $\ldots$, $a_k h_{r-1}$, вход равен a_{k-1}, а на выходе появляется сумма $a_k h_{r-1} + a_{k-1} h_r$, т.е. второй коэффициент многочлена-произведения. Читатель без труда проследит процесс умножения до конца.

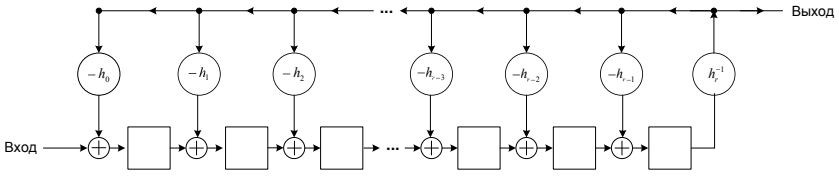

Рис. 9.6. Регистр сдвига для деления на многочлен $h(x)$.

На рис. 9.6 представлен регистр сдвига, предназначенный для деления произвольного многочлена

$$a(x) = a_0 + a_1 x + \ldots + a_n x^n \tag{9.3.3}$$

над $GF(q)$ на фиксированный многочлен

$$h(x) = h_0 + h_1 x + \ldots + h_{r-1} x^{r-1} + h_r x^r \qquad (9.3.4)$$

с коэффициентами из того же поля. Читатель заметил две основные черты схемы деления: все коэффициенты делителя, кроме старшего, взяты со знаком минус, а вместо старшего коэффициента делителя взят обратный ему элемент $h_r^{-1} \in GF(q)$.

Именно эти черты и характеризуют процесс деления.

Действительно, рассмотрим процесс деления "столбиком". На первом шаге деления старший член $a_n x^n$ делимого делится на старший член $h_r x^r$ делителя, т.е. умножается на h_r^{-1}.

Первое неполное частное есть $a_n h_r^{-1} x^{n-r}$.

(Так как показатели степеней определяются нижними индексами соответствующих коэффициентов, то упоминание о них опускается, что и усматривается непосредственно из строения регистра. Поэтому далее будем говорить только о делении коэффициентов).

Затем делитель умножается на неполное частное и вычитается из делимого. Но это равносильно тому, что на неполное частное умножается делитель, взятый с противоположным знаком, и складывается с делимым. Это обстоятельство и изображено на рис. 9.6. Так получается первый остаток. С ним происходит то же самое, что и с делимым, и т.д.

Перед началом работы регистра он заполнен нулями, и деление начинается, как только старший член делимого достигнет выхода последнего элемента задержки. Это случится на $(r+1)-$м шаге работы регистра. Процесс деления заканчивается, когда коэффициент a_0 поступит на вход регистра. В этот момент в последнем элементе задержки окажется коэффициент при x^{r-1}. Степень содержимого регистра меньше степени делителя. Деление прекращается. В регистре оказывается остаток. На рис. 9.7 представлен регистр сдвига для деления на многочлен $g(x) = 1 + x^3 + x^4 + x^5$ над $GF(2)$.

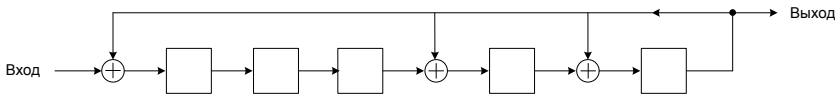

Рис. 9.7. Регистр сдвига для деления на многочлен $g(x) = 1 + x^3 + x^4 + x^5$ над $GF(2)$

Как уже наверняка догадался читатель, регистры сдвига для умножения и деления на многочлен заведомо могут быть употреблены для целей кодирования и декодирования.

9.4. Линенйные рекуррентные соотношения и генераторы с регистром сдвига

Построенные в предыдущем разделе схемы можно рассматривать как результат некоторой общей теории.

Пусть в выражении

$$\sum_{j=0}^{k} h_j a_{i+j} = 0 \tag{9.4.5}$$

$h_j \in GF(q)$, $h_0 \neq 0$, $h_k = 1$. Это равносильно соотношению

$$a_{i+k} = -\sum_{j=0}^{k-1} h_j a_{i+j} = 0. \tag{9.4.6}$$

Решением этих уравнений является последовательность

$$a_0, a_1, \ldots, a_{k-1}.$$

По любому комплекту k известных последовательных значений a^l вычисляется $(k+1)-$е. Линейная комбинация решений есть снова решение, и все решения образуют линейное векторное пространство размерности k.

Пусть величины h_j, $j = 0, 1, \ldots, k, h_0 \neq 0$, трактуются как коэффициенты нормированного многочлена

$$h(x) = \sum_{0}^{k} h_j x^j, \tag{9.4.7}$$

и условимся, что n это минимальное натуральное число, при котором $h(x)$ делит двучлен $x^n - 1$. Будем также интерпретировать решения уравнения (9.4.6), т.е. последовательности $a_0, a_1, \ldots, a_{n-1}$, как коэффициенты многочлена

$$a(x) = a_0 x^{n-1} + a_1 x^{n-2} + \ldots + a_{n-1}, \tag{9.4.8}$$

по модулю многочлена $x^n - 1$.

Тогда совокупность всех различных q^k решений есть не что иное, как идеал, порожденный многочленом $g(x) = (x^n - 1)/h(x)$.

Пусть теперь многочлен (9.4.7) есть примитивный многочлен степени $k = m$. Выходная последовательность начнет повторяться, как только начнут повторяться комплекты переменных в правой части (9.4.6). Длина n выходной последовательности в точности равна минимальному показателю степени, при которой двучлен $x^n - 1$ делится на $h(x)$. Число этих комплектов, включая и нулевой, в точности равно q^m, а потому $n = q^m - 1$, и этому же числу равно и количество самих ненулевых выходных последовательностей.

Согласно предыдущему рассуждению порождающим многочленом идеала, состоящего из всех возможных последовательностей, является многочлен $g(x) = (x^{q^m-1} - 1)/h(x)$. Построенный идеал при $q = 2$ есть код, двойственный циклическому двоичному коду Хэмминга, (см. разделы 5.6 и 6.2).

Если же многочлен $h(x)$ не является примитивным, то длина n последовательности окажется равной $q^{m_1} - 1$, где m_1 есть некоторый делитель числа m.

9.5. Схемы умножения на константу поля Галуа

Остается научиться строить схемы умножения на константу поля $GF(q)$. В регистрах сдвига они играют не последнюю роль.

Построение таких схем дается представлением поля Галуа матрицами (раздел 3.8) и теоремой 3.8.2. Здесь рассматривается только случай $q = 2^m$. Поэтому произвольный элемент поля $GF(2^m)$ представляется двоичным вектором длины m, а схема умножения, показанная на рис. 9.1*b*, имеет m двоичных входов $f_0, f_1, \ldots, f_{m-1}$ и m двоичных выходов $\varphi_0, \varphi_1, \ldots, \varphi_{m-1}$. Строение схемы умножения на фиксированный элемент $\alpha^i \in GF(2^m)$ полностью описывает

Теорема 9.5.1. *Значение выхода φ_u, $u = 0, 1, \ldots, m-1$, схемы получается в виде суммы по модулю 2 значений на тех ее входах, номера которых совпадают с номерами единичных разрядов u-го столбца матрицы B^i.*

Доказательство. Умножим элемент $\alpha^j \in GF(2^m)$ на фиксированный элемент α^i из того же поля. Из раздела 3.8 и теоремы 3.8.2 следует, что результат умножения α^{j+i} есть

в точности первая строка матрицы $B^{j+i} = B^j B^i$. По правилу умножения матриц значение выхода φ_u, $u = 0, 1, \ldots, m-1$, схемы получается как скалярное произведение вектора-строки α^j на u-й столбец матрицы B^i. Опустив нулевые члены этого столбца, получим требуемое.

На рис. 9.8 и 9.9 показаны схемы умножения соответственно на α^6 и α^{14} посредством матриц. Использовано поле (3.4.11).

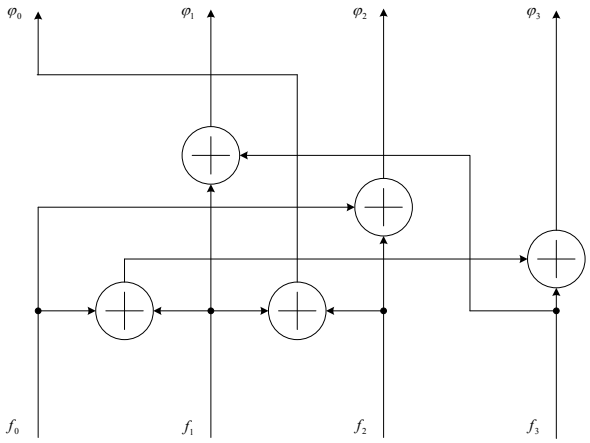

Рис. 9.8. Схема умножения на $\alpha^6 \in GF(2^4)$

На рис. 9.10 показан регистр сдвига для кодирования с помощью умножения на порождающий многочлен $g(x) = 1 + x + x^3$. 4 элемента задержки отвечают 4-м информационным символам циклического (7, 4)-кода.

9.6. Мажоритарное декодирование циклического кода

Эта глава заключается совместным рассмотрением трех факторов. Во-первых, циклический код, двойственный циклическому коду Хэмминга (см. раздел 6.2), допускает мажоритарное декодирование (теорема 4.8.1). Во-вторых, мажоритарное декодирование этих кодов реализует кодовое расстояние (теорема 4.8.2). Наконец, цикличность кода в сочетании с регистрами сдвигов открывает возможность построения весьма интересных схем декодирования.

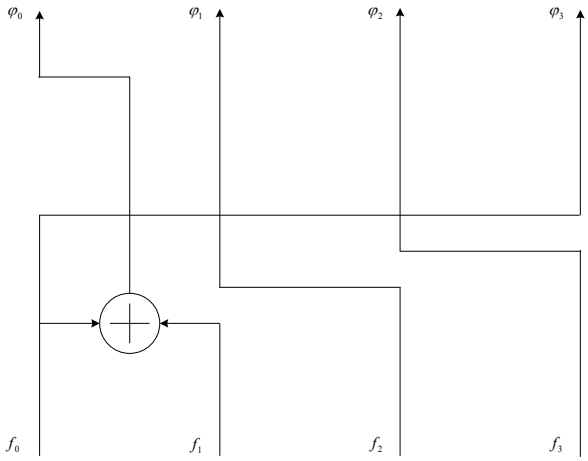

Рис. 9.9. Схема умножения на $\alpha^{14} \in GF(2^4)$

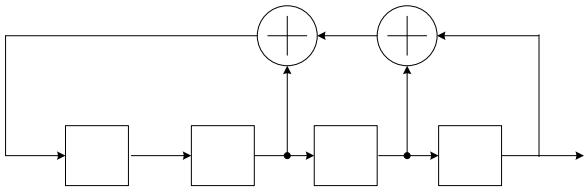

Рис. 9.10. Регистр сдвига для кодирования с помощью умножения на многочлен $g(x) = 1 + x + x^3$

Рассмотрим циклический код, порожденный многочленом $g(x) = (x+1)(x^3+x^2+1) = x^4+x^2+x+1$. (Он двойствен коду Хэмминга, порожденному многочленом $g(x) = (x^3 + x + 1)$).

Согласно процедуре построения канонических форм порождающей и проверочной матрицы кода, изложенной в разделе 5.4, получаем:

$$x^6 = (x^4 + x^2 + x + 1)(x^2 + 1) + x^3 + x + 1,$$
$$x^5 = (x^4 + x^2 + x + 1)x + x^3 + x^2 + x,$$
$$x^4 = (x^4 + x^2 + x + 1) + x^2 + x + 1.$$

Отсюда порождающая матрица кода будет

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

и согласно соотношению между каноническими формами порождающей и проверочной матриц

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

Вспоминая, как из проверочной матрицы проверочные символы выражаются через информационные, получим

$$\begin{aligned} a_4 &= a_1 + a_2, \\ a_5 &= a_2 + a_3, \\ a_6 &= a_1 + a_2 + a_3, \\ a_7 &= a_1 + a_3. \end{aligned}$$

Сложив второе равенство с третьим и разрешив все равенства относительно a_1, а затем, добавив к трем равенствам тривиальное $a_1 = a_1$, окончательно найдем:

$$\begin{aligned} a_1 &= a_2 + a_4, \\ a_1 &= a_5 + a_6, \\ a_1 &= a_3 + a_7, \\ a_1 &= a_1. \end{aligned} \tag{9.6.9}$$

Это система разделенных проверок для информационного символа a_1. Вследствие цикличности кода аналогичные системы проверок имеют место для всех, а не только информационных, символов кодового вектора. С помощью этой системы исправляется любая одиночная ошибка. Вообще, как знает читатель, для любого циклического $(2^m - 1,\ m,\ 2^{m-1})$-кода, двойственного коду Хэмминга, можно построить систему из 2^{m-1} разделенных проверок, а число исправляемых ошибок есть $2^{m-2} - 1$. Здесь код имеет длину 7 только ради удобства демонстрации схемы декодирования. Схема декодирования рассмотренного циклического (7, 3)-кода показана на рис. 9.11.

В первые 7 тактов работы в регистр сдвига вводится принятый вектор $(a_1,\ a_2,\ a_3,\ a_4,\ a_5,\ a_6,\ a_7)$. В момент, когда последний символ a_7 займет свое место в регистре, ключ на входе

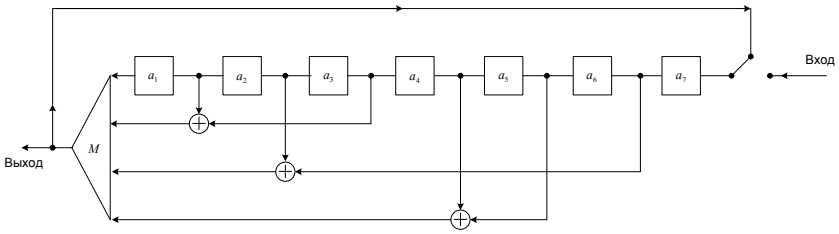

Рис. 9.11. Схема мажоритарного декодирования циклического (7, 3)-кода с порождающим многочленом $g(x) = (x^3 + x^2 + 1)(x + 1) = x^4 + x^2 + x + 1$

занимает положение, показанное на рис. 9.11, и начинается процесс декодирования. Вычисленные посредством сумматоров значения a_1 на каждом из равенств системы (9.6.9) поступают на входы мажоритарного элемента M. Истинное (если число ошибок не превосходит $2^{m-2} - 1$, т.е. в данном случае — одной) значение символа a_1 поступает и на выход схемы, и по линии обратной связи в последний элемент задержки, который освободился от символа a_7 благодаря циклическому сдвигу содержимого регистра на один шаг (влево).

Происходит декодирование символа a_2.

Процесс декодирования продолжается пока не будут декодированы либо только информационные символы, либо, по желанию пользователя, все символы принятого вектора.

9.7. Задачи к главе 9

9.1. Построить схемы умножения на константы α^5, α^7. α^9 поля $GF(2^4)$.
9.2. Построить схемы умножения на константы α^3, α^6, α^8 поля $GF(2^8)$.

Глава 10.

Указания к решению задач

10.1. К главе 1

1.2. Числа $a\xi$ пробегают приведенную систему вычетов по модулю m. Расположим правильные несократимые дроби $\{(a\xi)/m\}$ в порядке возрастания их числителей: $(i_1/m), (i_2/m), \ldots (i_c/m)$, где $c = \varphi(m)$, а затем в порядке убывания. Каждая сума двух числителей, равноудаленных от концов, равна m, а всего таких сумм будет $\varphi(m)$. Таким образом, суммируется в точности $\varphi(m)$ единиц. Но просуммированы два ряда дробей, откуда и получается коэффициент $1/2$.

1.3. Имеем $x = 10y + 3$, $x \equiv 15(\mathrm{mod}17)$, $x \equiv 11(\mathrm{mod}13)$. К любой части сравнения можно прибавить любое целое, кратное модуля: $x \equiv -2(\mathrm{mod}17)$, $x \equiv -2(\mathrm{mod}13)$. Если сравнение имеет место по нескольким модулям, то оно имеет место и по модулю, равному наименьшему общему кратному модулей: $x \equiv -2(\mathrm{mod}221)$. Отсюда:$10y + 3 \equiv -2(\mathrm{mod}221)$, $10y + 5 \equiv 0(\mathrm{mod}221)$, $10y + 5 = t221$. Наименьшее $t = 5, 10y + 5 = 1105$. $x = 1103$.

1.6. Пусть

$$as_0 + bt_0 \tag{10.1.1}$$

есть минимальное из положительных чисел вида

$$as + bt, \tag{10.1.2}$$

где a и b не равны нулю одновременно. Покажем, что число $as_0 + bt_0$ делит все числа $as + bt$. Действительно, $as + bt = (as_0 + bt_0)q + r$, откуда $r = as + bt - (as_0 + bt_0)q = a(s - qs_0) + b(t - ty_0))$. Иначе говоря, остаток r имеет вид $as' + bt'$, и он меньше делителя (10.1.1). Но по условию число (10.1.1) есть минимальное из положительных чисел вида (10.1.2). Поэтому $r = 0$.

При $s = 1,\ t = 0\ as + bt = a$, и $as_0 + bt_0$ делит a. При $s = 0,\ t = 1\ as + bt = b$, и $as_0 + bt_0$ делит b. Это значит, что число (10.1.1) делит одновременно и a и b. Отсюда следует, что любой общий делитель чисел a и b, в том числе и их наибольший общий делитель, делит число (10.1.1), а потому число (10.1.1) есть наибольший общий делитель чисел a и b. $as_0 + bt_0 = (a, b)$.

Другое решение немедленно следует из выражения (1.9.18) в формулировке леммы о мультипликативности функции Эйлера.

1.7. При $p = 2$ решение тривиально. Пусть $p > 2$. Выразим элементы приведенной системы вычетов по модулю p через первообразный корень g, который существует. $(p-1)! = g^0 g^1 \cdots g^{p-2} = g^{(p-1)(p-2)/2}$. По теореме Ферма $g^{p-1} - 1 \equiv 0(\mod p)$, $g^{p-1} - 1 = (g^{(p-1)/2} - 1)(g^{(p-1)/2} + 1) \equiv 0(\mathrm{mod} p)$. Только одна из скобок слева делится на p. В противном случае их разность, равная 2, делилась бы на $p > 2$. Не может быть $(g^{(p-1)/2} - 1) \equiv 0(\mod p)$, так как g первообразный корень. Следовательно $g^{(p-1)/2} \equiv -1(\mathrm{mod} p)$. И так как $p - 2$ нечётно, то $g^{(p-1)(p-2)/2} \equiv -1(\mathrm{mod} p)$.

1.8. Если q — простое нечетное, и q делит $a^p - 1$, т.е. $a^p \equiv 1(\mathrm{mod} q)$, то a по модулю q может принадлежать только одному из показателей $\delta = 1; p$. Действительно, если a по модулю q принадлежит показателю δ, т.е. одновременно $a^p \equiv 1(\mathrm{mod} q)$, и $a^\delta \equiv 1(\mathrm{mod} q)$, то $a^p \equiv a^\delta(\mathrm{mod} q)$, и, следовательно, $p \equiv \delta(\mathrm{mod}\delta)$. Это значит, что δ делит p, и так как p простое, то $\delta = 1; p$. При $\delta = 1$ имеем $a \equiv 1(\mod q)$, и q делит $a - 1$. Если же $\delta = p$, то p является делителем числа $\varphi(q) = q - 1$, а потому $q - 1 = 2px$, так как $q - 1$ четное.

1.9. Если q простое нечетное, и $a^p + 1 \equiv 0(\mathrm{mod} q)$, то $a^p \equiv -1(\mathrm{mod} q)$, и $a^{2p} \equiv 1(\mathrm{mod} q)$. Пусть δ есть показатель, которому a принадлежит по модулю q, т.е. $a^\delta \equiv 1(\mathrm{mod} q)$. Если два числа сравнимы с третьим, то они сравнимы между собой, значит, $a^\delta \equiv a^{2p}(\mathrm{mod} q)$. Но по теореме 1.11.3 отсюда $\delta \equiv 2p(\mathrm{mod}\delta)$. Значит, δ делит $2p$. Поэтому $\delta = 1, 2, p, 2p$. Но случаи $\delta = 1, p$, невозможны. Действительно, $a^p \equiv 1(\mod q)$ не может быть, так как по условию $a^p \equiv -1(\mathrm{mod} q)$. Если бы было $a \equiv 1(\mathrm{mod} q)$, то и $a^p \equiv 1(\mathrm{mod} q)$, что невозможно по доказанному выше.

Теперь, если $\delta = 2$, то $a^2 \equiv 1(\mathrm{mod} q)$, $a^2 - 1 \equiv 0(\mathrm{mod} q)$, и $a + 1 \equiv 0(\mathrm{mod} q)$, так как $a - 1 \not\equiv 0(\mathrm{mod} q)$, Если же $\delta = 2p$, то

$\varphi(q) = q-1$ делится на $2p$, и потому $q-1 = 2px$, или $q = 2px+1$.
1.10. Используя формулу для полиномиальных коэффициентов, найдём

$$(x_1 + x_2 + \ldots + x_a)^p = \sum_{i_1, i_2, \ldots i_a} \frac{p!}{i_1! i_2! \ldots i_a!} x_1^{i_1} x_2^{i_2} \ldots x_a^{i_a},$$

где

$0 \le i_j \le p,\ i_1 + i_2 + \ldots + i_a = p.$

Отсюда, полагая последовательно $i_j = p$, все остальные члены суммы можно объединить членом pQ, т.е.,

$$(x_1 + x_2 + \ldots + x_a)^p = (x_1^p + x_2^p + \ldots + x_a^p) + pQ.$$

Положим $x_1 = x_2 = \ldots = x_a = 1$. Получим

$$a^p = a + pQ \equiv a(\mathrm{mod} p).$$

1.11. Имеем

$$(a,\, p) = 1.\ a^{p-1} \equiv 1(\mathrm{mod} p),$$

$$a^{p-1} = 1 + pt_1,$$

$$a^{p(p-1)} = (1 + pt_1)^p = (1 + C_p^1 pt_1 + ...) = 1 + p^2 t_2,$$

$$a^{p^2(p-1)} = (1 + p^2 t_2)^p = (1 + C_p^1 p^2 t_3 + ...) = 1 + p^3 t_3,$$

$$\ldots,$$

$$a^{p^{\alpha-1}(p-1)} = 1 + p^\alpha t_\alpha.$$

Но

$$p^{\alpha-1}(p-1) = \varphi(p^\alpha),$$

и потому

$$a^{\varphi(p^\alpha)} \equiv 1(\mathrm{mod} p^\alpha).$$

Пусть

$$m = p_1^{\alpha^1} p_2^{\alpha^2} \cdots p_k^{\alpha^k}$$

Тогда

$a^{\varphi(p_1^{\alpha_1})} \equiv 1(\mathrm{\ mod\ } p_1^{\alpha_1}),\ a^{\varphi(p_2^{\alpha_2})} \equiv 1(\mathrm{\ mod\ } p_2^{\alpha_2}), \ldots, a^{\varphi(p_k^{\alpha_k})} \equiv 1(\mathrm{\ mod\ } p_k^{\alpha_k}).$

Возведем каждое сравнение в $k-1$ степеней так, чтобы каждый показатель в левых частях сравнений был равен

$$\varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2})\cdots\varphi(p_k^{\alpha_k}).$$

А это есть ни что иное, как $\varphi(m)$. Таким образом, имеем одно и то же сравнение

$$a^{\varphi(m)} \equiv 1(\mathrm{mod} p_1^{\alpha_1}),\ a^{\varphi(m)} \equiv 1(\mathrm{mod} p_2^{\alpha_2}),\ a^{\varphi(m)} \equiv 1(\mathrm{mod} p_k^{\alpha_k})$$

по нескольким модулям. Значит, оно имеет место по модулю, равному наименьшему общему кратному модулей:

$$a^{\varphi(m)} \equiv 1(\mathrm{mod} p_1^{\alpha_1} p_2^{\alpha_2}\cdots p_k^{\alpha_k}).$$

Иначе говоря,

$$a^{\varphi(m)} \equiv 1(\mathrm{mod} m).$$

10.2. К главе 2

2.1. В качестве примера рассмотреть приведённые системы вычетов по модулям $m = 5$ и $m = 8$.
2.8. Все степени элемента x составляют циклическую группу порядка pq. Все степени элемента x^p, составляют ее подгруппу порядка q, а все степени элемента x^q составляют ее подгруппу порядка p. Пусть $u = (x^p)^m, \quad v = (x^q)^n$. Покажем, что можно найти такие m и n, чтобы было $x = uv = vu$. Последнее означает, что $x = (x^p)^m \times (x^q)^n = (x^q)^n \times (x^p)^m$.

Иначе говоря, $x = x^{pm+qn}$. Это равенство будет выполнено, если $pm + qn = 1 + tpq$, где t целое. Или, что то же,

$$pm + qn - tpq = pm + q(n - pt) = pm + qn^{'} = 1. \qquad (10.2.3)$$

Так как $(p, q) = 1$, то существуют такие m и $n^{'}$, а значит и m и n, что (10.2.3) выполняется.

Легко вдеть также, что

$$u^q = ((x^p)^m)^q = (x^{pq})^m = 1, \quad v^p = ((x^q)^n)^p = (x^{pq})^n = 1,$$

2.10. Рассмотреть разложение группы H_1 по подгруппе D, которая есть пересечение подгрупп H_1 и H_2. Затем умножить это разложение на H_2.

В качестве примера можно рассмотреть циклическую группу G порядка $n = 63$ в разделе 2.8 (пример 2.3). $H_3 \cap H_7 = H_{21}$ порядка $d = 3$. или $H_7 \cap H_9 = I$ порядка $d = 1$. $H_3H_7 = H_7H_9 = G$.

2.15. Проверка замкнутости. Если $a, b \in H$, или $a, b \in N$, то замкнутость очевидна. Если $a \in H, b \in N$, то и в этом случае $ab \in HN$. Наличие обратного элемента. Для $a \in H, b \in N$ наличие обратных элементов очевидно, так как H и N группы, и обе они содержатся в HN. Найдём обратный элемент для ab. Имеем $(ab)^{-1} = b^{-1}a^{-1}$. Но $b^{-1} \in N$, $a^{-1} \in H$. Далее $a^{-1}N \in HN$, $a^{-1}N = Na^{-1}$, так как N нормальный делитель. $Na^{-1} \in HN$, $b^{-1} \in N$, так как N группа. $b^{-1}a^{-1} \in HN$.

2.19. Приведённая система вычетов по простому модулю есть группа чётного порядка. Для каждого её элемента существует ему обратный. Их произведение равно единице. Сами себе обратными являются 1 и $p-1$. Отсюда $(p-1)! \equiv (p-1)(\mathrm{mod}\,)p \equiv -1$.

Переведём изложенное решение на язык решения 1.7. Объединим в непересекающиеся пары $g^i, g^j,\ 0 < i, j \leq p-2$, степени $g^0, g^1, \ldots, g^{p-2}$ первообразного корня g так, чтобы было $i + j = p - 1$. Произведение этих степеней в паре обратится в единицу. Иначе говоря, сомножители в паре — это взаимно обратные элементы в группе элементов приведённой системы вычетов по модулю p. Нет пары для элемента $g^0 = 1$, но он сам себе обратный, и элемента $g^{(p-1)/2}$, который, как было показано, равен -1. Он тоже сам себе обратный: $(g^{(p-1)2})^2 = g^{p-1} \equiv 1(\mathrm{mod}p)$.

10.3. К главе 3

3.4. Построить поле $GF(2^5)$ и найти комплекты сопряженных элементов.

3.5. Пары взаимно обратных элементов в поле $GF(11)$ являются 2, 6; 3, 4; 5, 9; 7, 8; 1 и 10 обратны сами себе. Только эти пары могли бы быть корнями двучлена x^2+1, так как произведения внутри пар дают свободный член двучлена. Однако ни в одной паре сумма её членов не равна 0, т.е. коэффициенту при x. Аналогичное рассуждение следует провести на случай многшочлена $x^2 + x + 4$.

3.6. Получить $a^3 = b^3$, и, используя ограничение на степень расширения, показать, что $a = b = 0$.

3.11. Корни любого неприводимого над $GF(q)$ квадратного многочлена лежат в $GF(q^2)$.

3.12. $x^8 + x^7 + x^3 + x + 1 = (x^7 + 1)(x + 1) + (x + 1) + x^3 + x + 1$. Многочлен $x^3 + x + 1$ неприводим, он делит двучлен $x^7 + 1$ и не делит двучлен $x + 1$.

3.14. Найдется такое i, что, если α есть корень многочлена над $GF(q)$, то $\alpha^{q^i} = \alpha^{-1}$. Далее использовать факт примитивности многочлена.

3.15. Воспользоваться выражением для показателя неприводимого самодвойственного многочлена.

3.16. –**3.17.** Воспользоваться процедурой построения поля Галуа и показать, что, если α корень многочлена, то его 3,7,9 и 21 степени в первом случае, и 3,5,15,17,51 и 85 степени — во втором не равны 1.

3.18. Показать, что если $\alpha^5 = \alpha - 1$, то в $GF(3^5)$ $\alpha^2, \alpha^{11}, \alpha^{22}, \alpha^{121} \neq 1$.

3.21. Число $q^m - 1$ может быть простым только при $q = 2$ и только при m простом.

3.22. Необходимость Пусть $a = b^2$, т.е b есть корень квадратный из a. Тогда $a^{(q-1)/2} = (b^2)^{(q-1)/2} = b^{(q-1)} = 1$. Достаточность. Пусть c есть первообразный элемент мультипликативной группы $GF^*(q)$ Тогда c^2 порождает (циклическую) подгруппу G порядка $(q-1)/2$ группы $GF^*(q)$. Из условия задачи, т.е., из того, что $a^{(q-1)/2} = 1$, следует, что $a \in G$. Но эта подгруппа состоит из вторых степеней, т.е., $a = c^{2i}$.

3.23. Необходимость. Пусть a является $k-$й степенью, т.е., $a = b^k$. Тогда $a^{(q-1)/d} = (b^k)^{(q-1)/d} = b^{(q-1)k/d} = b^{(q-1)k_1} = 1$, где $k_1 = k/d$, так как по условию d делит k.

Достаточность. Пусть $a^{(q-1)/d} = 1$, где $d = ((q-1), k)$. Все a, которые удовлетворяют этому условию, являются $d-$ми степенями. Действительно, d является делителем $q-1$ т.е.,$q-1 = ld$. Для каждого делителя l порядка мультипликативной группы поля есть ее подгруппа порядка l . Её образующим элементом является $c^{(q-1)/l} = c^d$, где c есть порождающий элемент мультипликативной группы поля. Так как $d = ((q-1), k)$, то $k = dk_1$ и $(k_1, l) = 1$. Поэтому $(c^d)^{k_1} = c^{dk_1} = c^k$ – также образующий элемент. Значит, все $a = c^{id}$ имеют вид $a = c^{ik}$, что и требовлось.

3.24. Если в мультипликативной группе поля есть элемент второго порядка, то он только один.
3.25. Так как многочлен самодвойственный, то для каждого его корня α найдется такое $j \neq m$, что $\alpha^{p^j} = \alpha^{-1}, \alpha^{p^j+1} = 1$. С другой стороны, так как α^{-1} есть корень, то найдется такое $i \neq m$, что $(\alpha^{-1})^{p^i} = \alpha$, и $\alpha^{p^i} = \alpha^{-1}$. Поэтому $\alpha^{p^i} = \alpha^{p^j}, p^i = p^j, i = j$. Далее, из $(\alpha^{-1})^{p^i} = \alpha$ и $\alpha^{-1} = \alpha^{p^j}$ следует $(\alpha^{p^j})^{p^i} = \alpha^{p^{i+j}} = \alpha = \alpha^{p^m}$. Отсюда $i + j = m$ и $j = i = m/2, \alpha^{p^j} = \alpha^{p^{m/2}} = \alpha^{-1}, \alpha^{p^{m/2}+1} = 1$. Значит, любой корень самодвойственного многочлена является корнем двучлена $H_m = x^{p^{m/2}+1} - 1$, который, таким оразом, делится на любой самодвойственный неприводимый многочлен степени m, что и требовалось.
3.28. Рассмотреть свободный член двучлена $x^{p-1} - 1$.

10.4. К главе 4

4.1. Показать, что в противном случае строки порождающей матрицы будут линейно зависимы.
4.3. Пусть вектор $v \in V$ представляет собой базис подпространства $V' \subset V$. Его размерность равна 1, и размерность его ортогонального подпространства равна $n - 1$. Легко проверяется, что все векторы, ортогональные данному, образуют подпространство.
4.5. Если некоторая линейная комбинация w столбцов проверочной матрицы равна S, то соответствующий вектор веса w содержится в смежном классе с синдромом S. Если вектор веса w содержится в смежном классе с синдромом S, то это означает,что линейная комбинация некоторых w столбцов проверочной матрицы равна S.
4.6. Рассмотреть разложение кодового подпространства A по некоторому его подходящему подпространству B.
4.7. Векторы четного веса образуют подгруппу. Векторы нечётного веса образуют смежный класс.
4.9. Скалярные произведения принятого вектора на строки проверочной матрицы образуют систему линейных уравнений с неизвестными x_i, номера которых есть номера компонент, где произошли стирания.
4.10. Совокупность векторов, веса которых имеют одинаковую чётность, образует код, заведомо обнаруживающий любую одиночную ошибку. Сравнить в произвольном коде количества векторов четного и нечетного весов.

4.11. Рассмотреть процесс и очерёдность выбора базисных векторов.
4.12. Уяснить отличие данной задачи от предыдущей.
4.13. Воспользоваться результатом задачи 4.6.
4.14. Применить операцию укорочения кода.
4.15. Применить границу Плоткина (см. задачу 4.13).
4.16. Согласно задаче 4.15, число кодовых векторов не может превосходить 10. Но код линейный. Значит, число кодовых векторов может быть равно 4 и 8. Код мощности 4 построить легко: 000000000, 000011111, 111110000 и сумма двух последних 111101111. Пусть мощность кода равна 8. Так как d=5, то имеются векторы нечетного веса, и таких векторов должно быть четыре. Вектор веса 9 отсутствует, так как в противном случае минимальный вес окажется равным 4, что невозможно. Вектор веса 7 может быть только один. В противном случае минимальный вес будет не более, чем 4. Таким образом, при наличии вектора веса 7 остальные три вектора нечетного веса должны иметь иметь вес 5. Но разместить три вектора веса 5 на длине 9, так, чтобы их сумма имела вес не мене пяти, невозможно. Это означает, что векторов нечетного веса имеется только два. Значит, и четного — только два, значит всего их четыре, а не восемь, что и требовалось. Кроме приведенного выше кода, имеется и такой: 000000000, 000011111, 111111100, 111100011.
4.19. Вычислить сумму

$$n \geq d_0 + d_1 + \ldots + d_{k-1}$$

и воспользоваться границей Плоткина (4.16.53)

10.5. К главе 5

5.1. Опуская тривиальные случаи, когда $g(x) = x,\ x-1$, заметим, что, если бы минимальный вес циклического кода был равен 2, то идеал содержал бы многочлен $x^{n_1}-1,\ n_1 < n$, который обязан делиться на $g(x)$, а это противоречит условию задачи.
5.2. Рассмотреть корни самодвойственного многочлена.
5.3. Пусть многочлен $g(x)$ не делится на $(x-1)$. Так как x^n-1 делится на $g(x)$, и $\frac{x^n-1}{(x-1)} = (x^{n-1}+x^{n-2}+\ldots+1)$, то многочлен $(x^{n-1}+x^{n-2}+\ldots+1)$ делится на $g(x)$, т.е. принадлежит коду. Наоборот, если сплошь единичный вектор принадлежит коду,

и n не делится на p, то $\underbrace{1+1+\ldots+1}_{n \text{ слагаемых}} \neq 0$, т.е. 1 не является корнем многочлена $x^{n-1}+x^{n-2}+\ldots+1$, а потому и не является корнем многочлена $g(x)$, который, таким образом, не делится на $x-1$.

5.4. Разложить на множители порождающий многочлен и найти его корни.

5.5. Наличие в коде векторов только чётного веса означает, что порождающий многочлен имеет корнем 1, а потому он делится на $x-1$. Согласно задаче 5.3., код не имеет сплошь единичного вектора. Наоборот, если в коде есть сплошь единичный вектор, то порождающий многочлен не имеет корнем 1, а потому не все векторы имеют чётный вес.

5.6. При $n|q^m-1$ имеет место соотношение $(x^n-1)|(x^{q^m}-1)$. Для каждого делителя n порядка q^m-1 циклической группы существует её подгруппа порядка n. Элементами этой группы являются корни двучлена x^n-1, и их порядки делят n.

5.7. Многочлен $x^{n_1}-1$ не имеет кратных корней.

5.8. Представить пачку ошибок в виде многочлена и выяснить, какова степень порождающего многочлена.

10.6. К главе 6

6.1. Определить длину кода.

6.2. Разложить совокупность чисел 0, 1, ... , 79 по модулю 80 на циклотомические классы.

6.3. Если β есть корень самодвойственного многочлена x^2+x+1, то $\beta^3=1$. Если $\beta \in GF(2^3,)$ то самодвойственный порождающий многочлен содержит либо два кодовых вектора и имеет расстояние 7, либо содержит только один, именно, нулевой, кодовый вектор. Все корни самодвойственного многочлена $x^4+x^3+x^2+x+1$ имеют порядок 5, но код длины $n=5$ не может иметь расстояние 6 ни при каких условиях. Рассмотрение кодов бóльших длин не встречает особенностей. Учитывая, что $255=15\times 17$, найти две подходящие последовательности корней порядка 17 самодвойственного многочлена $g(x)$ и вставить в каждую из них 1.

6.4. Если элементы α и α' одного порядка, то они принадлежат одной и той же подгруппе мультипликативной группы поля. Более того они являются порождающими элементами этой группы, и длины обоих кодов совпадают. Порядок следования

корней порождающего многочлена по возрастанию показателей степеней также сохраняется.

10.7. К главе 7

7.5. Воспользоваться приведением порождающей матрицы циклического кода к каноническому виду.

Канонические разложения некоторых чисел

Различные параметры циклических кодов, в том числе длины кодов связаны с порядками элементов мультипликативных групп конечных полей. Как известно, эти порядки есть делители порядков указанных групп. В таблице П.1 представлены канонические разложения чисел $2^m - 1$, $m = 1, 2 \ldots, 34$.

Таблица П.1

$2^3 - 1 = 7$	$2^{19} - 1 = 524287$
$2^4 - 1 = 3 \times 5$	$2^{20} - 1 = 3 \times 5^2 \times 11 \times 31 \times 41$
$2^5 - 1 = 31$	$2^{21} - 1 = 7^2 \times 127 \times 337$
$2^6 - 1 = 3^2 \times 7$	$2^{22} - 1 = 3 \times 23 \times 89 \times 683$
$2^7 - 1 = 127$	$2^{23} - 1 = 47 \times 178481$
$2^8 - 1 = 3 \times 5 \times 17$	$2^{24} - 1 = 3^2 \times 5 \times 7 \times 13 \times 17 \times 241$
$2^9 - 1 = 7 \times 73$	$2^{25} - 1 = 31 \times 601 \times 1801$
$2^{10} - 1 = 3 \times 11 \times 31$	$2^{26} - 1 = 3 \times 2731 \times 8191$
$2^{11} - 1 = 23 \times 89$	$2^{27} - 1 = 7 \times 73 \times 262657$
$2^{12} - 1 = 3^2 \times 5 \times 7 \times 13$	$2^{28} - 1 = 3 \times \times 29 \times 43 \times 113 \times 127$
$2^{13} - 1 = 8191$	$2^{29} - 1 = 233 \times 1103 \times 2089$
$2^{14} - 1 = 3 \times 43 \times 127$	$2^{30} - 1 = 3^2 \times 7 \times 11 \times 31 \times 151 \times 331$
$2^{15} - 1 = 7 \times 31 \times 151$	$2^{31} - 1 = 2147483647$
$2^{16} - 1 = 3 \times 5 \times 17 \times 257$	$2^{32} - 1 = 3 \times 5 \times 17 \times 257 \times 65537$
$2^{17} - 1 = 131071$	$2^{33} - 1 = 7 \times 23 \times 89 \times 599479$
$2^{18} - 1 = 3^3 \times 7 \times 19 \times 73$	$2^{34} - 1 = 3 \times 43691 \times 131071$

Неприводимые многочлены

Задание циклического кода посредством корней порождающего многочлена требует нахождения их минимальных многочленов. Разумеется, зная хотя бы один неприводимый многочлен данной степени, можно по модулю этого многочлена построить соответствующее поле Галуа. После этого для каждого элемента поля находят его сопряженные элементы, т.е. все корни его минимального многочлена. а по ним и сам минимальный многочлен. Практические цели, однако, потребовали исключить этот процесс из реального обращения. Существуют таблицы неприводимых многочленов, освобождающие и исследователей и инженеров от рутинных вычислений (хотя можно предположить, что создатели этих таблиц отказались бы от такой благородной затеи – создать таблицы –, знай они наперед возможности современных вычислительных машин).

В этом приложении помещена таблица П.2. В ней представлены все неприводимые многочлены степеней до $m = 12$ над полем $GF(2)$. [1]

Каждый неприводимый многочлен представлен после знака (-) набором цифр от 0 до 7 (т.е. в восьмиричном виде). Представление каждой из цифр этого набора ее двоичным эквивалентом, дает последовательность коэффициентов неприводимого многочлена. Например, рассмотрим в разделе таблицы "СТЕПЕНЬ 8" самую первую запись 1-435E. Набор 435 в двоичном эквиваленте его цифр имеет вид 100011101, что соответствует многочлену $x^8+x^4+x^3+x^2+1$. Перед набором 435 помещено число 1. Это означает, что корнем многочлена является элемент α, и таким образом, поле $GF(2^8)$ построено по модулю именно этого многочлена. Вообще, число i, стоящее в начале каждой записи, означает, что корнем соответствующего многочлена является α^i, и α есть корень многочлена, находящегося в самой первой записи. Ясно, что, найдя некоторый многочлен,

[1] Перепечатка из книги У.У. Питерсона

тем самым находят и двойственный ему. Поэтому в таблице из каждой пары двойственных многочленов помещен только один из них. Поле $GF(2^m)$ является полем разложения многочлена $x^{2^m}-x$, который делится на все минимальные многочлены своих корней, т.е., на все минимальные многочлены всех элементов поля.

Каждая запись сопровождается буквой. Смысл букв таков:

A,B,C,D–Многочлен непримитивный,

E,F,G,H–Многочлен примитивный,

A,B,E,F–Корни линейно зависимы,

C,D,G,H–Корни линейно независимы,

A,C,E,G–Корни двойственного многочлена линейно зависимы,

B,D,F,H–Корни двойственного многочлена линейно независимы.

Некоторые сведения о многочлене можно извлечь и непосредственно. Например, если число, стоящее перед записью, встречается в разложениях чисел 2^m-1, то многочлен непримитивный.

Далее, линейная зависимость корней неприводимого многочлена обнаруживается немедленно по его второму коэффициенту. Второй коэффициент, как известно, равен суме корней, и если он равен нулю, то корни линейно зависимы.

Если число m не является простым, то для каждого делителя m_1 числа m поле содержит нетривиальное подполе $GF(2^{m_1})$. Поэтому многочлен степени $m_1<m$ заведомо помещен в разделе "СТЕПЕНЬ m_1" и там снабжен соответствующей буквой. Но он помещен и в разделе "СТЕПЕНЬ m", так как делит многочлен $x^{2^m-1}-1$. И здесь он уже буквой не снабжается.

Для примера вернемся к хорошо изученному полю $GF(2^4)$, т.е., к разделу "СТЕПЕНЬ 4". Здесь представлены знакомые многочлены x^4+x+1, $x^4+x^3+x^2+x+1$, x^2+x+1. Последний помещен в виде 5-07 и не снабжен буквой, так как он в виде 1-7H уже представлен в разделе "СТЕПЕНЬ 2" и там снабжен буквой H, а многочлен x^4+x^3+1 отсутствует, будучи двойственным многочлену 1-23F. Заметим, что многочлен 7 содержится в разделах степеней 2, 4, 6, 8, 10 и 12 в виде, соответственно, 7H, 07, 007, 0007 и 00007, а многочлен 37 содержится в разделах степеней 4, 8, и 12 в виде, соответственно, 37D, 037 и 00037.

Таблица П.2

СТЕПЕНЬ 2 1—7H.

СТЕПЕНЬ 3 1—13F.

СТЕПЕНЬ 4 1—23F 3—37D 5—07.

СТЕПЕНЬ 5 1—45E 3—75G 5—67H.

СТЕПЕНЬ 6 1—103F 3—127B 5—147H 7—111A 9—015 11—155E 21—007.

СТЕПЕНЬ 7

$1-211E$ $3-217E$ $5-235E$ $7-367H$ $9-277E$ $11-325G$
$13-203F$ $19-313H$ $21-345G$.

СТЕПЕНЬ 8

$1-435E$	$3-567B$	$5-763D$	$7-551E$	$9-675C$
$11-747H$	$13-453F$	$15-727D$	$17-023$	$19-545E$
$21-613D$	$23-543F$	$25-433B$	$27-477B$	$37-537F$
$43-703H$	$45-471A$	$51-037$	$85-007$.	

СТЕПЕНЬ 9

$1-1021E$	$3-1131E$	$5-1461G$	$7-1231A$	$9-1423G$
$11-1055E$	$13-1167F$	$15-1541E$	$17-1333F$	$19-1605G$
$21-1027A$	$23-1751E$	$25-1743H$	$27-1617H$	$29-1553H$
$35-1401C$	$37-1157F$	$39-1715E$	$41-1563H$	$43-1713H$
$45-1175E$	$51-1725G$	$53-1225E$	$55-1275E$	$73-0013$
$75-1773G$	$77-1511C$	$83-1425G$	$85-1267E$.	

СТЕПЕНЬ 10

$1-2011E$	$3-2017B$	$5-2415E$	$7-3771G$	$9-2257B$
$11-2065A$	$13-2157F$	$15-2653B$	$17-3515G$	$19-2773F$
$21-3753D$	$23-2033F$	$25-2443F$	$27-3573D$	$29-2461E$
$31-3043D$	$33-0075C$	$35-3023H$	$37-3543F$	$39-2107B$
$41-2745E$	$43-2431E$	$45-3061C$	$47-3177H$	$49-3525G$
$51-2547B$	$53-2617F$	$55-3453D$	$57-3121C$	$59-3471G$
$69-2701A$	$71-3323H$	$73-3507H$	$75-2437B$	$77-2413B$
$83-3623H$	$85-2707E$	$87-2311A$	$89-2327F$	$91-3265G$
$93-3777D$	$99-0067$	$101-2055E$	$103-3575G$	$105-3607C$
$107-3171G$	$109-2047F$	$147-2355A$	$149-3025G$	$155-2251A$
$165-0051$	$171-3315C$	$173-3337H$	$179-3211G$	$341-0007$.

СТЕПЕНЬ 11

$1-4005E$	$3-4445E$	$5-4215E$	$7-4055E$	$9-6015G$
$11-7413H$	$13-4143F$	$15-4563F$	$17-4053F$	$19-5023F$
$21-5623F$	$23-4757B$	$25-4577F$	$27-6233H$	$29-6673H$
$31-7237H$	$33-7335G$	$35-4505E$	$37-5337F$	$39-5263F$
$41-5361E$	$43-5171E$	$45-6637H$	$47-7173H$	$49-5711E$
$51-5221E$	$53-6307H$	$55-6211G$	$57-5747F$	$59-4533F$
$61-4341E$	$67-6711G$	$69-6777D$	$71-7715G$	$73-6343H$
$75-6227H$	$77-6263H$	$79-5235E$	$81-7431G$	
$83-6455G$	$85-5247F$	$87-5265E$	$89-5343B$	$91-4767F$
$93-5607F$	$99-4603F$	$101-6561G$	$103-107H$	$105-7041G$
$107-4251E$	$109-5675E$	$111-4173F$	$113-4707F$	$115-7311C$
$117-5463F$	$119-5755E$	$137-6675G$	$139-7655G$	$141-5531E$
$147-7243H$	$149-7621G$	$151-7161G$	$153-4731E$	$155-445IE$
$157-6557H$	$163-7745G$	$165-7317H$	$167-5205E$	$169-4565E$
$171-6765G$	$173-7535G$	$179-4653F$	$181-5411E$	$183-5545E$
$185-7565G$	$199-6543H$	$201-5613F$	$203-6013H$	$205-7647H$
$211-6507H$	$213-6037H$	$215-7363H$	$217-7201G$	$219-7273H$
$293-7723H$	$299-4303B$	$301-5007F$	$307-7555G$	$309-4261E$
$331-6447H$	$333-5141E$	$339-7461G$	$341-5253F.$	

СТЕПЕНЬ 12

$1-10123F$	$3-12133B$	$5-10115A$	$7-12153B$
$9-11765A$	$11-15447E$	$13-12513B$	$15-13077B$
$17-16533H$	$19-16047H$	$21-10065A$	$23-11015E$
$25-13377B$	$27-14405A$	$29-14127H$	$31-17673H$
$33-13311A$	$35-10377E$	$37-13565E$	$39-13321A$
$41-15341G$	$43-15053H$	$45-15173C$	$47-15621E$
$49-17703C$	$51-10355A$	$53-15321G$	$55-10201A$
$57-12331A$	$59-11417E$	$61-13505E$	$63-10761A$
$65-00141$	$67-13275E$	$69-16663C$	$71-11471E$
$73-16237E$	$75-16267D$	$77-15115G$	$79-12515E$
$81-17545C$	$83-12255E$	$85-11673B$	$87-17361A$
$89-11271E$	$91-10011A$	$93-14755C$	$95-17705A$
$97-17121G$	$99-17323D$	$101-14227H$	$103-12117E$
$105-13617A$	$107-14135G$	$109-14711G$	$111-15415C$
$113-13131E$	$115-13223A$	$117-16475C$	$119-14315C$
$121-16521E$	$123-13475A$	$133-11433B$	$135-10571A$
$137-15437G$	$139-12067F$	$141-13571A$	$143-12111A$
$145-16535C$	$147-17657D$	$149-12147F$	$151-14717F$
$153-13517B$	$155-14241C$	$157-14675G$	$163-10663F$
$165-10621A$	$167-16115G$	$169-16547C$	$171-10213B$
$173-12247E$	$175-16757D$	$177-16017C$	$179-17675E$
$181-10151E$	$183-14111A$	$185-14037A$	$187-14613H$

$189-13535A$	$195-00165$	$197-11441E$	$199-10321E$
$201-14067D$	$203-13157B$	$205-14513D$	$207-10603A$
$209-11067F$	$211-14433F$	$213-16457D$	$215-10653B$
$217-13563B$	$219-11657B$	$221-17513C$	$227-12753F$
$229-13431E$	$231-10167B$	$233-11313F$	$235-11411A$
$237-13737B$	$239-13425E$	$273-00023$	$275-14601C$
$277-16021G$	$279-16137D$	$281-17025G$	$283-15723F$
$285-17141A$	$291-15775A$	$293-11477F$	$295-11463B$
$297-17073C$	$299-16401C$	$301-12315A$	$307-14221E$
$309-11763B$	$311-12705E$	$313-14357F$	$315-17777D$
$325-00163$	$327-17233D$	$329-11637B$	$331-16407F$
$333-11703A$	$339-16003C$	$341-11561E$	$343-12673B$
$345-14537D$	$347-17711G$	$349-13701E$	$355-10467B$
$357-15347C$	$359-11075E$	$361-16363F$	$363-11045A$
$365-11265A$	$371-14043D$	$397-12727F$	$403-14373D$
$405-13003B$	$407-17057G$	$409-10437F$	$411-10077B$
$421-14271G$	$423-14313D$	$425-14155C$	$427-10245A$
$429-11073B$	$435-10743B$	$437-12623F$	$439-12007F$
$441-15353D$	$455-00111$	$585-00013$	$587-14545G$
$589-16311G$	$595-13413A$	$597-12265A$	$603-14411C$
$613-15413H$	$619-17147F$	$661-10605E$	$683-10737F$
$685-16355C$	$691-15701G$	$693-12345A$	$715-00133$
$717-16571C$	$819-00037$	$1365-00007.$	

Таблица П.3

Некоторые неприводимые многочлены над полями

$$GF(3),\ GF(5),\ GF(7).$$

1. Над $GF(3)$: $x+1$, x^2+x+2, x^3+2x+1, x^4+x+2, x^5+2x+1, x^6+x+2.

2. Над $GF(5)$: $x+1$, x^2+x+2, x^3+3x+2, x^4+x^2+2x+2.

3. Над $GF(7)$: $x+1$, x^2+x+3, x^3+3x+2.

Литература

[1] *Питерсон У.У.* Коды, исправляющие ошибки. М.: Мир. 1964. 264 с.

[2] *Колесник В.Д., Мирончиков Е.Т.* Декодирование циклических кодов. М.: Связь. 1968. 252 с.

[3] *Берлекэмп Э.* Алгебраическая теория кодирования. М.: Мир. 1971. 477 с.

[4] *Блох Э.Л., Зяблов В.В.* Обобщенные каскадные коды. М.: Связь.1976. 240 с.

[5] *Питерсон У.У., Уэлдон Э.Дж.* Коды, исправляющие ошибки. М.: Мир. 1976. 594 с.

[6] *Касами Т., Токура Н., Ивадари Е., Инагаки Я.* Теория кодирования. М.: Мир. 1978. 576 с.

[7] *Мак-Вильямс Ф.Дж., Н.Дж.А.Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки. М.: Связь. 1979. 744 с.

[8] *Блох Э.Л., Зяблов В.В.* Линейные каскадные коды. М.: Наука. 1982. 230 с.

[9] *Афанасьев В.Б., Габидулин Э.М.* Кодирование в радиоэлектронике. М.: Радио и связь. 1986. 176 с.

[10] *Блэйхут Р.* Теория и практика кодов, контролирующих ошибки. М.: Мир. 1986. 576 с.

[11] *Лидл Р., Нидеррайтер Г.* Конечные поля. М.: Мир. 1988. Т. I, Т. II. 818 с.

[12] *Влэдуц С.Г., Ногин Д.Ю., Цфасман М.А.* Алгеброгеометрические коды (Основные понятия). М.: Московский центр непрерывного математического образования. 2003. 504 с.

Предметный указатель

Вандермонда
  определитель, 188

Ньютона
  тождества, 204, 205

Стирлинга
  неравенства, 259

Эвклида алгоритм , 22

автоморфизм
  группы, 67
автоморфизм поля, 107
алгоритм Эвклида
  для многочленов, 238
алгоритмом Питерсона — Горенстейна — Цирлера, 219
атака, 45

базис, 76
вектор, 8
  кодовый, 123
  ошибка, 8
вектор-ошибка, 19
вес вектора, 123
взаимно однозначное соответствие, 106
вложение кодов РС, 228
выбрасывание, 133
выкалывание, 133
выкалывание кода МДР, 224
вычет, 29
  наименьший неотрицательный, 29

генератор мультипликативной группы поля, 266
гомоморфный
  образ, 68
  прообраз, 68
граница
  Бассалыго-Элайеса, 257
    асимптотическая , 262
  Варшамова-Гилберта, 125, 258, 263
    асимптотическая, 263
  Гилберта, 15
  Плоткина, 166, 257, 258
    асимптотическая, 258
  Синглтона, 124, 257, 258
  Хэмминга, 130, 257, 262
    для случая четного расстояния, 134
граница Синглтона, 221
группа
  абелева, 49
  автоморфизмов, 107
  аддитивная, 53
  вращений, 55
  изоморфная, 65
  конечная, 53
  корней из единицы, 101
  мультипликативная, 53
  перестановок, 49
  показатель, 55
  порядок, 53
  симметрическая, 49
  циклическая, 40, 57, 84

декодер, 126
декодирование, 9
  мажоритарное, 134
  синдромное, 128
деление на фиксированный многочлен, 269

деление произвольных
элементов поля, 266
делитель нуля, 71
дешифрование, 43
длина вектора, 10

знакопеременная, 56
значение ошибки, 214, 218
значения искажений, 252

идеал, 73, 271
главный, 75
нулевой, 74
целый, 74
изоморфизм
полей Галуа, 105
индекс
группы по подгруппе, 63
информационная совокупность, 221
искажение, 246
искажения, 251
исправление ошибок и стираний, 220
исчерпание, 15

канал, 7
двоичный
стирающий, 16
двоичный симметричный, 8
с аддитивной помехой, 9
симметричный
q-ичный, 114
квадратное уравнение, 200
класс
смежный
левосторонний, 61
правосторонний, 61
циклотомический, 104
ключ
открытый, 42
секретный, 42
ключевое уравнение, 242
для ошибок и стираний, 247
код, 11
БЧХ, 186
Голея, 130
циклический, 192
МДР, 221
Рида — Маллера, 139
сложность
декодирования, 157
Рида—Маллера
кодирование, 148
Рида—Соломона
декодирования, 234
Рида–Маллера
минимальное расстояние, 145
Рида-Маллера
r-го порядка, 141
кодирование, 152
порождающая матрица, 141
сложность кодирования, 149
Рида-Соломона, 224
1-удлинение, 230
2-удлинение, 231
3-удлинение, 233
кодирование, 226
Хэмминга
циклический, 182
групповой, 114
двойственный
коду Хэмминга, 136
каскадный, 253
квазисовершенный, 130
линейный, 113
ортогональный, 118
систематический, 120
совершенный, 129
циклический, 167
двойственный коду Хэмминга, 190
примитиивный, 180
кодирование, 9
вектора, 116
кольцо, 70, 71, 79
без делителей нуля, 71
без единицы, 72
главных идеалов, 75
с делителем нуля, 71
с единицей, 72
корень
из единицы, 101
многочлена, 84
первообразный, 39
примитивный, 100
уравнения, 84

корни порождающего многочлена, 177
криптоаналитик, 42
криптограф, 42

лидер смежного класса, 128
линейная зависимость, 76
линейная независимость, 76
локаторы ошибок, 203, 218
локаторы стираний, 219

матрица
  Адамара, 161
  каноническая форма, 120
  порождающая, 116
    метрические свойства, 116
    приведенно-ступенчатая форма, 119
    циклический код, 170
  порождающая циклического кода
    приведённо-ступенчатая форма, 174
  проверочная, 118
    метрические свойства, 124
    циклического кода, 172
  проверочная циклического кода
    приведенно-ступенчатая форма, 175
  ранг, 124
  сопровождающая многочлена, 108
метод исчерпания, 14
минимальная функция, 94
минимальный многочлен, 93
многочлен
  двойственный, 111
  информационный, 226
  круговой, 101
  локаторов ошибок, 204
  локаторов стираний, 247, 250
  неприводимый, 78
  нормированный, 93, 169
    примитивный, 111
  порождающий
    идеал, 169
    циклический код, 169
  примитивный, 100
  проверочный, 173
  самодвойственный, 111
многочлен значений ошибок, 242
многочлен локаторов ошибок, 242
множество
  с операцией, 48
  замкнутое, 48
модуль, 24, 69
  простой, 73
  составной, 73
мультипликативная группа, 80

неприводимый многочлен, 78
нормальный базис, 198
нормальный делитель, 64

область целостности, 72, 79
операция, 48
  ассоциативная, 49
  групповая, 53
  коммутативная, 48
  некоммутативная, 48
  обратная, 48
отказ от декодирования, 12
ошибка
  декодирования, 10
  исправление, 12
  исправление и обнаружение, 13
  обнаружение, 12

пачка ошибок, 185
подгруппа, 55
  истинная, 56
  несобственная, 56
  циклической группы, 57
подполе, 95, 96
подпространства
  ортогональные, 118
подпространство, 76
показатель
  многочлена, 100
  элемента, 37
поле, 72
  конечное, 78
  конечной характеристики, 82
  простое , 95
  разложения, 82
поле Галуа, 86

группа
аддитивная, 86
мультипликативная, 86
пополнение, 133
порядок
группы, 64
корней неприводимого многочлена, 100
элемента, 64
последовательность, 8
присоединение корня, 84
проверка
разделенные, 136
проверочная сумма, 152
пропускная способность, 10
пространство, 75
линейное векторное, 76

разложение пространства по подпространству, 126
размерность пространства, 76
расстояние
гарантированное, 192
кодовое, 11
реализуется, 139
конструктивное, 192
минимальное, 12
расширение кода, 131
расширение поля, 81
решение ключевого уравнения, 243

символы
информационные, 120
проверочные, 120
синдром, 126
синдромный многочлен, 242
модифицированный, 248
система
вычетов, 29
наименьших абсолютных, 29
наименьших неотрицательных, 29
приведенная
вычетов, 32
система передачи, 7
скалярное произведение, 18
скорость передачи, 9
след, 199
слово, 8
собственная, 56
сравнение, 24
стандартное расположение, 128
степенные суммы, 204
стирание, 16
исправление, 16
сумматор двухвходовой, 264
схема умножения на константу поля, 271

теорема
Кэли, 66
Лагранжа, 63
Ферма, 33
Эйлера, 33
о гомоморфизме, 68

удлинение кода, 133
укорочение кода, 133
укорочение кода МДР, 224
умножение на фиксированный многочлен, 267
умножение произвольных элементов поля, 266
устройство умножения на константу поля, 264

фактор-группа, 65
фактор-модуль, 70
функция
Эйлера, 31
мультипликативность, 34
минимальная, 93
элементарная
симметрическая, 204
энтропии, 10

характеристика поля, 82

частичная сумма, 158
числа
сравнимые, 41

шар, 15
шифрование, 42

элемент
образующий, 84
первообразный, 58
порождающий, 58, 83

порядок, 54
сопряженный, 100
элемент задержки на такт, 264
энтропия $q-$ичного
симметричного канала,
261

ядро гомоморфизма , 69

Оглавление

Предисловие **3**

Предисловие ко второму изданию **6**

Введение **7**
0.1 Система передачи информации 7
0.2 Кодовое расстояние 11
0.3 Скорость передачи и расстояние 14
0.4 Код Хэмминга 18
0.5 Задачи к введению 19

1 Начальные сведения из теории чисел **21**
1.1 Предварительные замечания 21
1.2 Наибольший общий делитель. Алгоритм Эвклида 22
1.3 Сравнения 24
1.4 Свойства сравнений 25
1.5 Дальнейшие свойства сравнений 27
1.6 Полная система вычетов 29
1.7 Приведённая система вычетов 31
1.8 Теоремы Эйлера и Ферма 33
1.9 Функция Эйлера мультыпликативна 34
1.10 Вычисление функции Эйлера 35
1.11 Первообразные корни 36
1.12 Индексы 40
1.13 Приложения к криптографии 41
1.14 Задачи к главе 1 46

2 Элементы теории групп, колец и полей **48**
2.1 Множество с операцией 48
2.2 Обратная операция 48
2.3 Группа 49

2.4 Порядок группы и порядок элемента группы . . 53
2.5 Примеры групп 55
2.6 Подгруппы 55
2.7 Циклические группы 56
2.8 Подгруппы циклической группы 57
2.9 Разложение группы по подгруппе 61
2.10 Нормальные делители 64
2.11 Изоморфизм групп 65
2.12 Гомоморфизм групп 67
2.13 Несколько замечаний 69
2.14 Кольцо 70
2.15 Поле 72
2.16 Идеал 73
2.17 Линейное векторное пространство 75
2.18 Задачи к главе 2 76

3 Конечные поля **78**
3.1 Множество классов-вычетов 78
3.2 Поле разложения многочлена $x^{p^m} - x$ 81
3.3 Цикличность мультипликативной группа поля . 82
3.4 Задание поля корнем неприводимого многочлена 84
3.5 Строение конечных полей. 91
3.6 Изоморфизм полей Галуа 105
3.7 Автоморфизм поля Галуа 106
3.8 Представление поля Галуа матрицами 108
3.9 Задачи к главе 3 110

4 Линейные коды **113**
4.1 Код как линейное подпространство 113
4.2 Порождающая матрица кода 114
4.3 Проверочная матрица кода 117
4.4 Каноническая форма базисных матриц 118
4.5 Проверочная матрица и расстояние 122
4.6 Декодирование линейного кода 125
4.7 Операции над кодами 131
4.8 Мажоритарное декодирование 134
4.9 Коды Рида—Маллера 139
4.10 Кодирование кода Рида—Маллера 148
4.11 Сложность кодирования 149
4.12 Декодирование кода Рида—Маллера 152
4.13 Сложность декодирования 157
4.14 Матрицы Адамара 161
4.15 Заключение 164
4.16 Задачи к главе 4 164

5 Циклические коды **167**
5.1 Циклический код как идеал 167
5.2 Порождающая матрица циклического кода . . . 170
5.3 Проверочная матрица циклического кода 172
5.4 Каноническая форма базисных матриц 174
5.5 Многочлен с заданными свойствами 177
5.6 Циклический код Хэмминга 182
5.7 Векторы всех циклических кодов 182
5.8 Задачи к главе 5 184

6 Коды Боуза—Чоудхури—Хоквингема **186**
6.1 Важнейший класс циклических кодов 186
6.2 Коды, двойственные кодам Хэмминга 190
6.3 Параметры кодов БЧХ 191
6.4 Декодирование кодов БЧХ 194
6.5 Декодирование двоичных кодов с исправлением двух ошибок 195
6.6 Нормальный базис и след элемента поля 198
6.7 Квадратное уравнение над $GF(2^m)$ 200
6.8 Общий случай декодирования двоичных кодов БЧХ 203
6.9 Общий случай декодирования q-ичных кодов БЧХ 214
6.10 Коды БЧХ и исправление стираний 219
6.11 Задачи к главе 6 220

7 Коды МДР **221**
7.1 Коды на границе Синглтона 221
7.2 Коды Рида—Соломона 224
7.3 Кодирование кода РС 226
7.4 Удлинение кодов РС 229
7.5 Декодирование кодов РС 234
7.6 Алгоритм Эвклида для многочленов 238
7.7 Вывод ключевого уравнения 241
7.8 Решение ключевого уравнения 243
7.9 Вывод ключевого уравнения на случай ошибок и стираний 246
7.10 Решение ключевого уравнения на случай ошибок и стираний 248
7.11 Коды РС и построение каскадных кодов 253
7.12 Задачи к главе 7 255

8 Сводка границ **257**
8.1 Верхние границы 257
8.2 Нижняя граница 258
8.3 Асимптотические границы 258

9 Регистры сдвига **264**
9.1 Элементарные устройства 264
9.2 Вычисления в полях Галуа 265
9.3 Умножение и деление многочленов 267
9.4 Линейные рекуррентные соотношения 270
9.5 Схемы умножения на константу поля Галуа . . . 271
9.6 Мажоритарное декодирование циклического кода 272
9.7 Задачи к главе 9 275

10 Указания к решению задач **276**
10.1 К главе 1 276
10.2 К главе 2 279
10.3 К главе 3 280
10.4 К главе 4 282
10.5 К главе 5 283
10.6 К главе 6 284
10.7 К главе 7 285

Канонические разложения некоторых чисел **286**

Неприводимые многочлены **287**

Литература **292**

Предметный указатель **294**