РОССИЙСКАЯ АКАДЕМИЯ НАУК

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Журнал основан в январе 1965 г. **ISSN: 0555-2923** **Выходит 4 раза в год**

Том 58, 2022 **Вып. 3**

Июль–Август–Сентябрь

М о с к в а

С О Д Е Р Ж А Н И Е

CONTENTS

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391 : 519.724

© 2022 г. **М.В. Бурнашев**

О ФУНКЦИИ НАДЕЖНОСТИ ДСК С БЕСШУМНОЙ ОБРАТНОЙ СВЯЗЬЮ ПРИ НУЛЕВОЙ СКОРОСТИ[1]

Рассматривается передача неэкспоненциального числа сообщений по двоичному симметричному каналу с бесшумной обратной связью. Получена оценка сверху для наилучшей экспоненты вероятности ошибки декодирования. Вместе с известной подобной оценкой снизу это позволяет найти функцию надежности такого канала при нулевой скорости.

Ключевые слова: функция надежности, бесшумная обратная связь.

DOI: 10.31857/S0555292322030019, **EDN:** DZTOVB

§ 1. Введение и основные результаты

Рассматривается двоичный симметричный канал ДСК(p) с переходной вероятностью $0 < p < 1/2$, $q = 1-p$, и бесшумной обратной связью. Задано общее время передачи n, а также $M_n = 2^{Rn}$, $0 < R < 1$, равновероятных сообщений $\{\theta_1, \theta_2, \ldots, \theta_{M_n}\}$. После момента времени n приемник принимает решение $\widehat{\theta}$, какое из сообщений является истинным сообщением $\theta_{\rm true}$.

Определим минимально возможную вероятность ошибки декодирования

$$P_{\rm e}(M_n, n, p) = \min \frac{1}{M_n} \sum_{i=1}^{M_n} P(\mathrm{e} \,|\, \theta_i), \tag{1}$$

где $P(\mathrm{e} \,|\, \theta_i)$ – вероятность ошибки декодирования для используемого метода передачи при условии, что θ_i является истинным сообщением $\theta_{\rm true}$, а минимум берется по всем методам передачи длины n.

Обозначим через $F(R,p)$, $0 < R < 1$, наилучшую экспоненту вероятности ошибки декодирования для $M_n = e^{Rn}$ кодовых слов по каналу ДСК(p) с бесшумной обратной связью, т.е.

$$F(R,p) = \limsup_{n\to\infty} \frac{1}{n} \ln \frac{1}{P_{\rm e}(M_n, n, p)}, \quad M_n = e^{Rn}, \tag{2}$$

где $P_{\rm e}(M_n, n, p)$ определено в (1). Ясно, что функция $F(R,p)$ не возрастает по R.

Введем также предельную величину $F(0,p)$

$$F(0,p) = \lim_{R\to 0} F(R,p), \quad 0 < p < 1/2. \tag{3}$$

Предел в (3) существует, так как функция $F(R,p)$ ограничена и не возрастает по R.

[1] Исследование выполнено при частичной финансовой поддержке Российского фонда фундаментальных исследований (номер проекта 19-01-00364).

Эквивалентным образом функция $F(0,p)$ определяется соотношением (2), если число сообщений M_n таково, что $M_n \to \infty$, но $\log M_n = o(n)$ при $n \to \infty$.

Аналогично определим величину $F_K(p)$, $K = 2, 3, \ldots$, как наилучшую экспоненту вероятности ошибки декодирования для K кодовых слов в канале ДСК(p) с бесшумной обратной связью, т.е.

$$F_K(p) = \limsup_{n\to\infty} \frac{1}{n} \ln \frac{1}{P_{\mathrm{e}}(K,n,p)}, \tag{4}$$

где $P_{\mathrm{e}}(K,n,p)$ – минимально возможная вероятность ошибки декодирования (для всех методов передачи длины n). В работе [1] было показано, что

$$F_3(p) = F_4(p) = \ldots = F(0,p), \tag{5}$$

и поэтому для исследования величины $F(0,p)$ достаточно изучить величину $F_3(p)$.

Обозначим через $E_k(p)$, $k \geqslant 2$, наилучшую экспоненту вероятности ошибки декодирования для k кодовых слов по каналу ДСК(p) без обратной связи. Ясно, что

$$E_2(p) = F_2(p) = \frac{1}{2} \ln \frac{1}{4pq}.$$

Также понятно, что $E_3(p)$ определяется n-симплексным кодом $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ (т.е. кодом, в котором $d(\boldsymbol{x}_i, \boldsymbol{x}_j) \approx 2n/3$ для всех $i \neq j$), и поэтому

$$E_3(p) = \frac{1}{3} \ln \frac{1}{4pq}.$$

Ясно, что $E_3(p) \leqslant F_3(p) \leqslant E_2(p)$.

В работе [1] было доказано

Предложение. *Для $P_{\mathrm{e}}(3,n,p)$ справедлива оценка сверху* (*см.* (4), (5))

$$P_{\mathrm{e}}(3,n,p) \leqslant \Big(\frac{q}{p}\Big)^{1/3} \big(p^{1/3}q^{2/3} + p^{2/3}q^{1/3}\big)^n. \tag{6}$$

Из (3), (5) и (6) следует

$$F_3(p) \geqslant F_{\mathrm{fb}}(p), \tag{7}$$

где

$$F_{\mathrm{fb}}(p) = -\ln\big(p^{1/3}q^{2/3} + p^{2/3}q^{1/3}\big) = -\ln\big[p^{1/3}q^{2/3}(1+z^{1/3})\big] \geqslant 0 \tag{8}$$

и

$$p + q = 1, \quad z = z(p) = p/q.$$

Из последующих работ [2, 3] (где другими методами исследовалась вся функция надежности $F(R)$) также, в частности, следовала формула (7).

При этом в работе [1] утверждалось, что в формуле (7) выполняется также противоположное неравенство

$$F_3(p) \leqslant F_{\mathrm{fb}}(p), \tag{9}$$

и тогда в случае справедливости формулы (9) из (7) следовало бы равенство

$$F_3(p) = F_{\mathrm{fb}}(p). \tag{10}$$

Однако строгое доказательство формулы (9) в [1] отсутствовало. Позже в работе [4] была сделана еще одна попытка доказать формулу (9) (используя общее уравнение Беллмана), однако позже выяснилось, что доказательство также неверно.

Далее в статье доказывается формула (9), и поэтому справедлива формула (10).

Опишем возможные методы передачи одного из трех сообщений по каналу ДСК с бесшумной обратной связью. Заметим, что любая разумная стратегия передачи имеет следующий вид. В каждый момент k, $k = 1, \dots, n$, основываясь на полученных на выходе канала сигналах $\boldsymbol{y}^{k-1}$, приемник выделяет некоторое сообщение θ_{i_0} и задает передатчику вопрос, является ли θ_{i_0} истинным сообщением θ_{true}. Здесь важно наличие бесшумной обратной связи! Если истинное сообщение θ_{true} совпадает с θ_{i_0}, т.е. $\theta_{\text{true}} = \theta_{i_0}$, то передается сигнал $x_k = 0$. Если $\theta_{\text{true}} \neq \theta_{i_0}$, то передается сигнал $x_k = 1$. После момента n принимается решение в пользу наиболее вероятного сообщения θ_i.

Стратегия передачи, использованная в работах [1–3], вполне естественна: в каждый момент времени k в качестве $\theta_{i_0}(k)$ выбирается наиболее вероятное из сообщений θ_i при условии $\boldsymbol{y}^{k-1}$. Кажется, что такая стратегия передачи дает наилучшую экспоненту вероятности ошибки декодирования $F_3(p)$ (однако это необходимо доказать, что не было сделано в [1–3]).

Основной результат статьи представляет

Т е о р е м а 1. *Для $P_{\text{e}}(3, n, p)$ справедлива оценка снизу*

$$P_{\text{e}}(3, n, p) \geqslant \frac{1}{2}\big(p^{1/3}q^{2/3} + p^{2/3}q^{1/3}\big)^n = \frac{1}{2}\big[p^{1/3}q^{2/3}(1 + z^{1/3})\big]^n, \quad z = p/q, \tag{11}$$

и поэтому для $F_3(p)$ имеет место формула (*см.* (8))

$$F_3(p) = F(0, p) = F_{\text{fb}}(p). \tag{12}$$

Заметим, что

$$E_2(p) = F_2(p) > F_3(p) > E_3(p), \quad 0 < p < 1/2. \tag{13}$$

Выход ДСК будем обозначать через $\boldsymbol{y}^k = y_1^k = (y_1, \dots, y_k)$, $k = 1, \dots, n$, где $y_k \in \{0, 1\}$.

Замечание 1. Поясним, почему при трех сообщениях бесшумная обратная связь в принципе может помочь уменьшить вероятность ошибки декодирования. Действительно, предположим, что в момент времени i выполнено

$$p(\boldsymbol{y}^i \,|\, \boldsymbol{x}_1) \approx p(\boldsymbol{y}^i \,|\, \boldsymbol{x}_2) \gg p(\boldsymbol{y}^i \,|\, \boldsymbol{x}_3),$$

т.е. сообщение θ_3 значительно менее вероятно, чем сообщения θ_1, θ_2 (и в силу имеющейся бесшумной обратной связи об этом известно на передающем конце канала!). Тогда для моментов времени $t > i$ можно в основном проверять только оставшиеся сообщения θ_1 и θ_2 (например, используя для этого противоположные кодовые слова, как при двух сообщениях). Так как $E_2(p) > E_3(p)$ (см. (13)), то такое кодирование позволило бы уменьшить вероятность ошибки декодирования.

Замечание 2. Правая часть формулы (8) имеет следующую полезную интерпретацию (не отмечавшуюся ранее). Пусть $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$ – n-симплексный код (т.е. код, в котором $d(\boldsymbol{x}_i, \boldsymbol{x}_j) \approx 2n/3$ для всех $i \neq j$). Тогда справедлива формула (доказательство см. в Приложении)

$$\mathbf{P}\{\mathcal{E}_n\} \sim e^{-F_{\text{fb}}(p)n}, \tag{14}$$

где

$$\mathcal{E}_n = \{\boldsymbol{y}^n : p(\boldsymbol{y}^n \,|\, \theta_1) \approx p(\boldsymbol{y}^n \,|\, \theta_2) \approx p(\boldsymbol{y}^n \,|\, \theta_3)\}. \tag{15}$$

Иными словами, событие $\mathcal{E}_n$ определяет экспоненту вероятности ошибки декодирования.

Замечание 3. С точки зрения поведения функций надежности канал ДСК(p) и гауссовский канал $G(A)$ с ограничением на среднюю мощность A во многом похожи друг на друга [5–7]. Однако эти же каналы с бесшумной обратной связью показывают и существенную разницу. В частности, для ДСК(p) имеем $F_3(p) < E_2(p)$ (см. (13)), в то время как для гауссовского канала $G(A)$ справедливо $F_3(A) = E_2(A)$ (см. [7]). Эта разница обусловлена тем, что для гауссовского канала $G(A)$ в некоторые моменты можно передавать очень мощные сигналы, в то время как для канала ДСК(p) это невозможно.

Следующий результат описывает оптимальный метод передачи в случае трех сообщений.

Теорема 2. *В каждый момент k, $k = 1, \dots, n$, наилучшее разбиение сообщений $\{\theta_1, \theta_2, \theta_3\}$ (минимизирующее вероятность ошибки декодирования $\mathbf{P}_{\mathrm{e}}(n)$) имеет следующий вид: наиболее вероятное сообщение (при условии выхода $\boldsymbol{y}^{k-1}$) против двух оставшихся сообщений.*

Статья организована следующим образом. В § 2 для полноты приводится короткое и очень изящное доказательство формулы (6) из [1] (по-видимому, это доказательство имеется только в диссертации [1] и не публиковалось в более доступных источниках). В § 3 доказывается теорема 2. В § 4 вводится и описывается марковская диаграмма декодера для оптимальной стратегии передачи. В § 5 с помощью этой диаграммы доказывается теорема 1.

§ 2. Доказательство предложения 1

Через $d(\boldsymbol{y}, \boldsymbol{x})$ будем обозначать расстояние Хэмминга между векторами $\boldsymbol{y}$ и $\boldsymbol{x}$. Для каждого момента k, $k = 1, \dots, n$, через $\theta^{(1)}(k), \theta^{(2)}(k), \theta^{(3)}(k)$ будем обозначать упорядочение сообщений $\theta_1, \theta_2, \theta_3$ при условии $\boldsymbol{y}^k$, такое что

$$p\big(\boldsymbol{y}^k \,|\, \theta^{(1)}(k)\big) \geqslant p\big(\boldsymbol{y}^k \,|\, \theta^{(2)}(k)\big) \geqslant p\big(\boldsymbol{y}^k \,|\, \theta^{(3)}(k)\big). \tag{16}$$

Через $\boldsymbol{x}^{(1)}(k), \boldsymbol{x}^{(2)}(k), \boldsymbol{x}^{(3)}(k)$ будем обозначать соответствующее упорядочение использованных кодовых блоков. Тогда (16) эквивалентно упорядочению

$$d\big(\boldsymbol{y}^k, \boldsymbol{x}^{(1)}(k)\big) \leqslant d\big(\boldsymbol{y}^k, \boldsymbol{x}^{(2)}(k)\big) \leqslant d\big(\boldsymbol{y}^k, \boldsymbol{x}^{(3)}(k)\big).$$

Будем называть $d^{(i)}(k) = d(\boldsymbol{y}^k, \boldsymbol{x}^{(i)}(k))$ числом "отрицательных голосов" против $\theta^{(i)}(k)$ за время k. Обозначим также $d_i = d_i(n)$.

Обозначим через $d_{1,3}(k)$, $k = 1, \dots, n$, среднее число "отрицательных голосов" против всех сообщений за время k, т.е.

$$d_{1,3}(k) = \frac{1}{3} \sum_{i=1}^{3} d^{(i)}(k). \tag{17}$$

Используем стратегию, в которой в каждый момент времени k выделяется наиболее вероятное сообщение $\theta^{(1)}(k)$ и передатчик отвечает на вопрос, является ли $\theta^{(1)}(k)$ истинным сообщением θ_{true}. Если $\theta_{\mathrm{true}} = \theta^{(1)}(k)$, то передатчик передает сигнал $x_k = 0$, а если $\theta_{\mathrm{true}} \neq \theta^{(1)}(k)$, то передается сигнал $x_k = 1$.

Тогда если выходной сигнал – это $y_k = 1$, то сообщение $\theta^{(1)}(k)$ получает дополнительно один отрицательный голос, а оставшиеся два сообщения $\{\theta^{(2)}(k), \theta^{(3)}(k)\}$ отрицательных голосов не получают. Если же выходной сигнал – это $y_k = 0$, то сообщение $\theta^{(1)}(k)$ дополнительных отрицательных голосов не получает, а каждое из оставшихся сообщений $\theta^{(2)}(k)$ и $\theta^{(3)}(k)$ получает дополнительно по одному отрицательному голосу. В результате, если $y_k = 1$, то величина $d_{1,3}$ из (17) увеличивается на $1/3$. Если же $y_k = 0$, то величина $d_{1,3}$ увеличивается на $2/3$. Если за все время n было получено на выходе m нулей и $n - m$ единиц, то $d_{1,3}(n) = (n + m)/3$. Всего имеется $\binom{n}{m}$ возможностей разместить m нулей на n позициях.

Для любого момента k справедливы неравенства

$$d^{(1)}(k) \leqslant d^{(2)}(k) \leqslant d^{(3)}(k) \leqslant d^{(2)}(k) + 1. \tag{18}$$

В (18) следует пояснить только последнее неравенство. Действительно, оно выполняется при $k = 1$ (т.е. после получения выхода y_1). Далее при $k \geqslant 2$ для используемой стратегии сообщения $\theta^{(2)}(k)$ и $\theta^{(3)}(k)$ всегда попадают в одну группу, и поэтому условие $d^{(3)}(k) \leqslant d^{(2)}(k)+1$ сохраняется (хотя сами сообщения $\theta^{(2)}(k)$ и $\theta^{(3)}(k)$ могут меняться).

Из (17) и (18) следует неравенство

$$d^{(2)}(k) \geqslant d_{1,3}(k) - 1/3. \tag{19}$$

Заметим, что каждая реализация выхода $\boldsymbol{y}^n$ с e ошибками имеет вероятность $p^e q^{n-e}$. Так как истинное сообщение получает e отрицательных голосов, то для ошибки декодирования необходимо иметь $d^{(2)}(n) = e$ или $d^{(3)}(n) = e$. В любом случае в силу (19) требуется $e \geqslant d_{1,3}(n) - 1/3$, и поэтому необходимо

$$p^e q^{n-e} \leqslant \Big(\frac{q}{p}\Big)^{1/3} p^{d_{1,3}(n)} q^{n-d_{1,3}(n)}. \tag{20}$$

Условие (20) ограничивает вероятность любой ошибочной траектории через величину $d_{1,3}(n)$. Заметим, что если m – число нулей, полученных на выходе за все время n, то каждой ошибочной траектории соответствует величина $d_{1,3}(n) = (n+m)/3$, $m = 0, 1, \ldots, n$. Так как всего имеется $\binom{n}{m}$ возможностей разместить m нулей на n позициях, то в силу (20) получаем

$$P_{\mathrm{e}}(3, n, p) \leqslant \Big(\frac{q}{p}\Big)^{1/3} \sum_{m=0}^{n} \binom{n}{m} p^{(n+m)/3} q^{(2n-m)/3} = \Big(\frac{q}{p}\Big)^{1/3} \big(p^{1/3} q^{2/3} + p^{2/3} q^{1/3}\big)^n,$$

откуда следует (6). ▲

§ 3. Доказательство теоремы 2

Рассмотрим передачу трех равновероятных сообщений $\{\theta_1, \theta_2, \theta_3\}$. После каждого момента k будем находить апостериорные вероятности сообщений $\pi_i(k)$, $i = 1, 2, 3$, используя принятый блок $\boldsymbol{y}^k = y_1^k = (y_1, \ldots, y_k)$, $k = 1, \ldots, n$. Передача в момент $k + 1$ зависит только от этих апостериорных вероятностей $\{\pi_i(k)\}$ (так как они составляют достаточную статистику). Можно считать, что в момент $k+1$ мы начинаем передачу, но используя априорные вероятности $\{\pi_i(k)\}$.

Обозначим через $d_i(k) = d_i(\boldsymbol{y}^k) = d(\boldsymbol{y}^k, \boldsymbol{x}_i(k))$ общее число "отрицательных голосов" против θ_i за время $[1, k]$. Обозначим также $d_i = d_i(n)$.

Всю информацию, которую имеет декодер в момент k, $k = 1, \ldots, n$, после получения сигнала на выходе $\boldsymbol{y}^k$, составляют апостериорные вероятности $\pi_i(\boldsymbol{y}^k)$ сообщений θ_i, $i = 1, 2, 3$ (или, эквивалентным образом, набор расстояний $d_i(\boldsymbol{y}^k)$, $i = 1, 2, 3$). Обозначим через $i_0(\boldsymbol{y}^k) \in \{1, 2, 3\}$ номер, при котором достигается максимальное значение величины $\pi_i(\boldsymbol{y}^k)$ (или, эквивалентным образом, минимальное значение величины $d_i(\boldsymbol{y}^k)$), т.е.

$$\pi_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) = \max_i \pi_i(\boldsymbol{y}^k), \quad d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) = \min_i d_i(\boldsymbol{y}^k), \quad k = 1, \ldots, n. \tag{21}$$

С точки зрения декодера величина $\pi_i(\boldsymbol{y}^n)$ есть апостериорная вероятность события $\{\theta_i = \theta_{\text{true}}\}$. Поэтому наилучшим (с точки зрения вероятности ошибки декодирования) является принятие решения в пользу сообщения $\theta_{i_0(\boldsymbol{y}^n)}$ с максимальной апостериорной вероятностью $\pi_{i_0(\boldsymbol{y}^n)}(\boldsymbol{y}^n)$. Поэтому в силу (21) имеем

$$P_{\text{e}}(n) = \mathbf{P}\{\theta_{i_0(\boldsymbol{y}^n)} \neq \theta_{\text{true}}\} = 1 - \mathbf{E}\, I_{\{\theta_{i_0(\boldsymbol{y}^n)} = \theta_{\text{true}}\}} = 1 - \mathbf{E}\, \pi_{i_0(\boldsymbol{y}^n)}. \tag{22}$$

Через $\mathcal{A}_k \in \{\theta_1, \theta_2, \theta_3\}$, $k = 1, \ldots, n$, будем обозначать сообщение, выделяемое приемником в момент k, относительно которого он задает передатчику вопрос, является ли $\mathcal{A}_k$ истинным сообщением θ_{true}.

Рассмотрим изменение величины $\pi_{i_0(\boldsymbol{y}^k)}$ из (21), (22) в зависимости от выбора сообщения $\mathcal{A}_{k+1}$. Для этого достаточно рассмотреть изменение величины

$$\sum_{j \neq i_0} z^{d_j(k) - d_{i_0}(k)},$$

где $z = p/q$, т.е. изменение величины $1/\pi_{i_0(\boldsymbol{y}^k)}$ (см. формулы (37), (38)).

Возможны два случая:

1) Существует единственный номер $i_0(\boldsymbol{y}^k)$, такой что $d_j(\boldsymbol{y}^k) - d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) \geqslant 1$ для всех $j \neq i_0(\boldsymbol{y}^k)$. Тогда $i_0(\boldsymbol{y}^{k+1}) = i_0(\boldsymbol{y}^k)$ для всех y_{k+1}. В этом случае наиболее вероятное сообщение $\theta_{i_0(\boldsymbol{y}^k)}$ в момент k остается наиболее вероятным и в момент $k + 1$ для любого выхода y_{k+1}.
2) Есть два различных номера $i_0(\boldsymbol{y}^k)$ и $i_1(\boldsymbol{y}^k)$, таких что $d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) = d_{i_1(\boldsymbol{y}^k)}(\boldsymbol{y}^k)$ и $d_j(\boldsymbol{y}^k) - d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) \geqslant 1$ для третьего номера.

Ясно, что в третьем возможном случае (когда все расстояния $d_j(\boldsymbol{y}^k)$, $j = 1, 2, 3$, равны) в силу симметрии любой выбор сообщения $\mathcal{A}_{k+1}$ (т.е. любое разбиение сообщений) дает одинаковый результат.

Рассмотрим сначала случай 1). Обозначим для краткости (где $z = p/q$)

$$\begin{aligned}
&a_j = a_j(\boldsymbol{y}^k) = z^{d_j(\boldsymbol{y}^k) - d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k)},\\
&\delta_j = \delta_j(y_{k+1}) = d_j(y_{k+1}) - d_{i_0(\boldsymbol{y}^k)}(y_{k+1}),\\
&B(k, \boldsymbol{y}^k) = \sum_{j \neq i_0} z^{d_j(\boldsymbol{y}^k) - d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k)} = \sum_{j \neq i_0} a_j,\\
&B(k+1) = \sum_{j \neq i_0} z^{d_j(\boldsymbol{y}^{k+1}) - d_{i_0(\boldsymbol{y}^{k+1})}(\boldsymbol{y}^{k+1})} = \sum_{j \neq i_0} z^{d_j(\boldsymbol{y}^{k+1}) - d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^{k+1})} =\\
&= \sum_{j \neq i_0} a_j z^{\delta_j(y_{k+1})}.
\end{aligned} \tag{23}$$

Обозначим также

$$B_j(k+1) = B(k+1), \quad \text{если } \mathcal{A}_{k+1} = \theta_j, \quad j = 1, 2, 3. \tag{24}$$

Заметим, что величины δ_j, $j = 1, 2, 3$, принимают только значения 0, 1 и -1. Без потери общности можно считать, что $i_0(\boldsymbol{y}^k) = 1$, и поэтому $i_0(\boldsymbol{y}^{k+1}) = 1$. Тогда имеем $\delta_1(y_{k+1}) = 0$ и

$$\begin{aligned} &B(k) = a_2 + a_3, \quad B(k+1) = a_2 z^{\delta_2(y_{k+1})} + a_3 z^{\delta_3(y_{k+1})}, \\ &\pi_1(k) = \frac{1}{1 + B(k)}, \quad \pi_1(k+1) = \frac{1}{1 + B(k+1)}. \end{aligned} \tag{25}$$

Рассмотрим распределения случайных величин $B_j(k+1)$, $j = 1, 2, 3$, при условии $i_0(\boldsymbol{y}^k) = 1$. Для $j = 1$, т.е. если $\mathcal{A}_{k+1} = \theta_1$, имеем

$$\delta_2 = \delta_3 = \begin{cases} 1 & \text{с вероятностью } \pi_1(k)q + (1 - \pi_1(k))p = p + (q-p)\pi_1(k), \\ -1 & \text{с вероятностью } q - (q-p)\pi_1(k), \end{cases} \tag{26}$$

и тогда

$$B_1(k+1) = \begin{cases} (a_2 + a_3)z = B(k)z & \text{с вероятностью } p + (q-p)\pi_1(k), \\ (a_2 + a_3)/z = B(k)/z & \text{с вероятностью } q - (q-p)\pi_1(k). \end{cases} \tag{27}$$

Для $j = 2$, т.е. если $\mathcal{A}_{k+1} = \theta_2$, имеем

$$\delta_3 = 0, \quad \delta_2 = \begin{cases} 1 & \text{с вероятностью } \pi_1(k)q + (1 - \pi_1(k))p = p + (q-p)\pi_1(k), \\ -1 & \text{с вероятностью } q - (q-p)\pi_1(k), \end{cases} \tag{28}$$

и поэтому

$$B_2(k+1) = \begin{cases} a_2 z + a_3 & \text{с вероятностью } p + (q-p)\pi_1(k), \\ a_2/z + a_3 & \text{с вероятностью } q - (q-p)\pi_1(k). \end{cases} \tag{29}$$

Аналогично для $j = 3$, т.е. если $\mathcal{A}_{k+1} = \theta_3$, имеем

$$\delta_2 = 0, \quad \delta_3 = \begin{cases} 1 & \text{с вероятностью } p + (q-p)\pi_1(k), \\ -1 & \text{с вероятностью } q - (q-p)\pi_1(k), \end{cases} \tag{30}$$

и поэтому

$$B_3(k+1) = \begin{cases} a_3 z + a_2 & \text{с вероятностью } p + (q-p)\pi_1(k), \\ a_3/z + a_2 & \text{с вероятностью } q - (q-p)\pi_1(k). \end{cases} \tag{31}$$

В результате при $i_0(\boldsymbol{y}^k) = 1$ имеем

$$\begin{aligned} E_1 &= \mathbf{E}\big[\pi_{i_0}(k+1) \,|\, \boldsymbol{y}^k, \mathcal{A}_{k+1} = \theta_1\big] = \mathbf{E}\left[\frac{1}{1 + B_1(k+1)} \;\Big|\; \boldsymbol{y}^k, \mathcal{A}_{k+1} = \theta_1\right] = \\ &= \frac{p + (q-p)\pi_1(k)}{1 + (a_2 + a_3)z} + \frac{q - (q-p)\pi_1(k)}{1 + (a_2 + a_3)/z}. \end{aligned} \tag{32}$$

Аналогично имеем

$$\begin{aligned} E_2 &= \mathbf{E}\big[\pi_{i_0}(k+1) \,|\, \boldsymbol{y}^k, \mathcal{A}_{k+1} = \theta_2\big] = \mathbf{E}\left[\frac{1}{1 + B_2(k+1)} \;\Big|\; \boldsymbol{y}^k, \mathcal{A}_{k+1} = \theta_2\right] = \\ &= \frac{p + (q-p)\pi_1(k)}{1 + a_2 z + a_3} + \frac{q - (q-p)\pi_1(k)}{1 + a_2/z + a_3} \end{aligned} \tag{33}$$

и

$$E_3 = \mathbf{E}\big[\pi_{i_0}(k+1)\,|\,\boldsymbol{y}^k, \mathcal{A}_{k+1}=\theta_3\big] = \mathbf{E}\left[\frac{1}{1+B_3(k+1)}\;\Big|\;\boldsymbol{y}^k, \mathcal{A}_{k+1}=\theta_3\right] =$$
$$= \frac{p+(q-p)\pi_1(k)}{1+a_2+a_3z} + \frac{q-(q-p)\pi_1(k)}{1+a_2+a_3/z}. \tag{34}$$

Покажем, что $E_1 \geqslant \max\{E_2, E_3\}$, т.е. что наилучшим является использование $\mathcal{A}_{k+1}=\theta_1=\theta_{i_0(\boldsymbol{y}^k)}$. В силу симметрии достаточно показать, что $E_1 \geqslant E_2$. Действительно, в силу (32) и (33) имеем

$$E_1 - E_2 = [p+(q-p)\pi_1(k)]\left[\frac{1}{1+(a_2+a_3)z} - \frac{1}{1+a_2z+a_3}\right] +$$
$$+ [q-(q-p)\pi_1(k)]\left[\frac{1}{1+(a_2+a_3)/z} - \frac{1}{1+a_2/z+a_3}\right] =$$
$$= \frac{[p+(q-p)\pi_1(k)]a_3(1-z)}{[1+(a_2+a_3)z][1+a_2z+a_3]} + \frac{[q-(q-p)\pi_1(k)]a_3(1-1/z)}{[1+(a_2+a_3)/z][1+a_2/z+a_3]} =$$
$$= a_3(1-z)\left\{\frac{p+(q-p)\pi_1(k)}{[1+(a_2+a_3)z][1+a_2z+a_3]} - \right.$$
$$\left. - \frac{q-(q-p)\pi_1(k)}{z[1+(a_2+a_3)/z][1+a_2/z+a_3]}\right\} = qa_3(1-z)\times$$
$$\times\left\{\frac{z+(1-z)\pi_1(k)}{[1+(a_2+a_3)z](1+a_2z+a_3)} - \frac{1-(1-z)\pi_1(k)}{(z+a_2+a_3)(1+a_2/z+a_3)}\right\}. \tag{35}$$

Достаточно показать, что

$$\frac{z+(1-z)\pi_1(k)}{[1+(a_2+a_3)z](1+a_2z+a_3)} - \frac{1-(1-z)\pi_1(k)}{(z+a_2+a_3)(1+a_2/z+a_3)} \geqslant 0,$$

или, эквивалентным образом (после стандартных преобразований с использованием формулы $\pi_1(k) = 1/(1+a_2+a_3)$),

$$a_2(1-z) \geqslant 0. \tag{36}$$

Формула (36) справедлива, если $z \leqslant 1$ (т.е. если $p \leqslant 1/2$). В силу (35) и (36) получаем $E_1 \geqslant E_2$. Аналогично получаем $E_1 \geqslant E_3$. Поэтому $E_1 \geqslant \max\{E_2, E_3\}$, а это значит, что наилучшим является использование $\mathcal{A}_{k+1}=\theta_1=\theta_{i_0(\boldsymbol{y}^k)}$. Это завершает рассмотрение случая 1).

Рассмотрим теперь случай 2), когда есть два различных номера $i_0(\boldsymbol{y}^k)$ и $i_1(\boldsymbol{y}^k)$, таких что $d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) = d_{i_1(\boldsymbol{y}^k)}(\boldsymbol{y}^k)$ и $d_j(\boldsymbol{y}^k) - d_{i_0(\boldsymbol{y}^k)}(\boldsymbol{y}^k) \geqslant 1$ для третьего номера. Без ограничения общности можно считать, что $i_0(\boldsymbol{y}^k)=1$ и $i_1(\boldsymbol{y}^k)=3$. Тогда $E_1 = E_3$, и остается показать, что $E_1 \geqslant E_2$, и тогда наилучшим является использование $\mathcal{A}_{k+1}=\theta_1=\theta_{i_0(\boldsymbol{y}^k)}$ (или $\mathcal{A}_{k+1}=\theta_3=\theta_{i_1(\boldsymbol{y}^k)}$). Заметим, что для любого y_{k+1} одно из расстояний $d_{i_0(\boldsymbol{y}^{k+1})}(\boldsymbol{y}^{k+1})$ или $d_{i_1(\boldsymbol{y}^{k+1})}(\boldsymbol{y}^{k+1})$ остается таким же, как и ранее в момент k. Оставшиеся вычисления по существу совпадают с (23)–(36) (в действительности они даже проще), и мы их опускаем. Это завершает доказательство теоремы 2. ▲

§ 4. Марковская диаграмма декодера для оптимальной стратегии

Введем марковскую диаграмму, описывающую эволюцию декодера во времени. Обозначим через $d_i(k) = d(\boldsymbol{y}^k, \boldsymbol{x}_i(k))$ общее число "отрицательных голосов" против θ_i

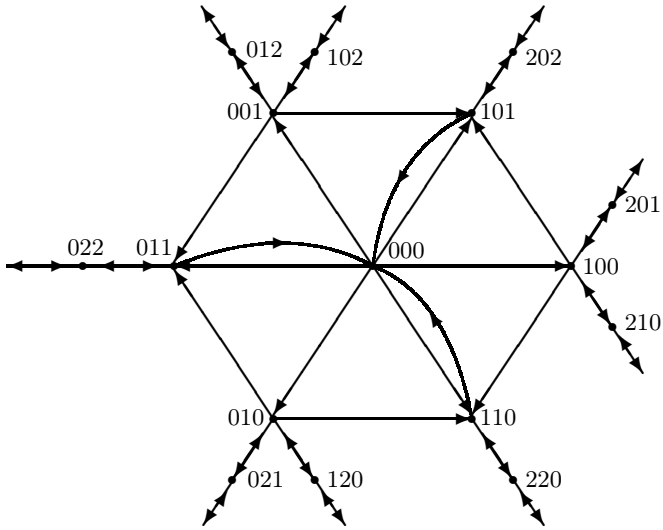

Рис. 1. Марковская диаграмма декодера

за время $[1, k]$. Обозначим также $d_i = d_i(n)$. Тогда ($z = p/q < 1$)

$$\pi_i(k) = \frac{z^{d_i(k)}}{\sum\limits_{j=1}^{3} z^{d_j(k)}} = \frac{1}{1 + \sum\limits_{j \neq i} z^{d_j(k) - d_i(k)}}, \quad \pi_i(n) = \frac{1}{1 + \sum\limits_{j \neq i} z^{d_j(n) - d_i(n)}}. \tag{37}$$

Заметим, что

$$\frac{\pi_i(k)}{1 - \pi_i(k)} = \frac{z^{d_i(k)}}{\sum\limits_{j \neq i} z^{d_j(k)}} = \frac{1}{\sum\limits_{j \neq i} z^{d_j(k) - d_i(k)}}. \tag{38}$$

Для каждого момента k и каждого сигнала на выходе $\boldsymbol{y}^k$ определим для сообщения θ_i метрику $m_i(k, \boldsymbol{y}^k)$ следующим образом:

$$m_i(k, \boldsymbol{y}^k) = d\big(\boldsymbol{y}^k, \boldsymbol{x}_i(k)\big) - \min_j d\big(\boldsymbol{y}^k, \boldsymbol{x}_j(k)\big) = d_i(k) - \min_j d_j(k), \quad i = 1, 2, 3. \tag{39}$$

Ясно, что $m_i(k, \boldsymbol{y}^k) \geqslant 0$ и $\min\limits_i m_i(k, \boldsymbol{y}^k) = 0$. Набор $\{m_i(k, \boldsymbol{y}^k)\}$ является достаточной статистикой, поскольку он определяет апостериорные вероятности $\{\pi_i(k)\}$ (см. (37)–(39)).

Обозначим через $S_{ij\ell} = S_{ijl}(k) = S_{ijl}(k, \boldsymbol{y}^k)$ состояние диаграммы с $i = m_1(k, \boldsymbol{y}^k)$, $j = m_2(k, \boldsymbol{y}^k)$, $\ell = m_3(k, \boldsymbol{y}^k)$.

В результате вся диаграмма выглядит как “осьминог” с девятью “щупальцами” (см. рис. 1). Например, одним из этих “щупалец” является $(S_{011}, S_{022}, S_{033}, \ldots)$.

Будем называть S_{000} *главным* состоянием, а шесть состояний $\{S_{011}, S_{100}, S_{101}, S_{010}, S_{110}, S_{001}\}$ – *основными* состояниями. Оставшиеся состояния расположены на “щупальцах”.

Тогда для вероятности ошибки декодирования $P_{\mathrm{e}}(n)$ имеем

$$P_{\mathrm{e}}(n) \geqslant \frac{2}{3} P_0(n), \tag{40}$$

где

$$P_0(n) = P\{S_{000}(0) \Rightarrow S_{000}(n)\}. \tag{41}$$

Опишем переходы между состояниями для оптимальной стратегии. Без ограничения общности можно считать, что $\theta_{\text{true}} = \theta_1$.

Если в момент k декодер находится в состоянии $S_{000}(k)$, то в этом случае множество $\mathcal{A}(k+1)$ выбирается равновероятно между тремя возможными вариантами. В результате для следующего возможного состояния $S(k+1)$ получаем

$$S_{000}(k) \to \begin{cases} S_{011}(k+1) & \text{с вероятностью } q/3, \\ S_{100}(k+1) & \text{с вероятностью } p/3, \\ S_{101}(k+1) & \text{с вероятностью } p/3, \\ S_{010}(k+1) & \text{с вероятностью } q/3, \\ S_{110}(k+1) & \text{с вероятностью } p/3, \\ S_{001}(k+1) & \text{с вероятностью } q/3. \end{cases} \tag{42}$$

Действительно, в момент k каждое сообщение θ_i имеет вероятность $\pi_i(k) = 1/3$, $i = 1, 2, 3$. Поэтому с вероятностью $1/3$ имеем $\mathcal{A}(k+1) = \theta_1$. Так как мы предположили, что $\theta_{\text{true}} = \theta_1$, то с вероятностью $q/3$ получаем $S(k+1) = S_{011}(k+1)$, а с вероятностью $p/3$ получаем $S(k+1) = S_{100}(k+1)$. Аналогично получаются остальные строки в (42).

Проще всего описываются переходы из тех состояний, для которых множество $\mathcal{A}(k+1)$ определяется однозначно, без рандомизации (т.е. когда есть только одно наиболее вероятное сообщение). Это состояния $S_{011}(k), S_{101}(k), S_{110}(k), \ldots$ Для таких состояний получаем

$$S_{011}(k) \to \begin{cases} S_{000}(k+1) & \text{с вероятностью } p, \\ S_{022}(k+1) & \text{с вероятностью } q, \end{cases} \tag{43}$$

$$S_{101}(k) \to \begin{cases} S_{000}(k+1) & \text{с вероятностью } q, \\ S_{202}(k+1) & \text{с вероятностью } p \end{cases} \tag{44}$$

$$S_{110}(k) \to \begin{cases} S_{000}(k+1) & \text{с вероятностью } q, \\ S_{220}(k+1) & \text{с вероятностью } p. \end{cases} \tag{45}$$

Аналогично описываются переходы из подобных же состояний $S_{022}(k), S_{202}(k)$, $S_{220}(k), \ldots$ Переходы из оставшихся состояний $S_{100}(k), S_{010}(k), S_{001}(k)$ описываются аналогично (42):

$$S_{100}(k) \to \begin{cases} S_{101}(k+1) & \text{с вероятностью } q/2, \\ S_{210}(k+1) & \text{с вероятностью } p/2, \\ S_{110}(k+1) & \text{с вероятностью } q/2, \\ S_{201}(k+1) & \text{с вероятностью } p/2, \end{cases} \tag{46}$$

$$S_{010}(k) \to \begin{cases} S_{021}(k+1) & \text{с вероятностью } q/2, \\ S_{110}(k+1) & \text{с вероятностью } p/2, \\ S_{120}(k+1) & \text{с вероятностью } p/2, \\ S_{011}(k+1) & \text{с вероятностью } q/2, \end{cases} \tag{47}$$

$$S_{001}(k) \to \begin{cases} S_{012}(k+1) & \text{с вероятностью } q/2, \\ S_{101}(k+1) & \text{с вероятностью } p/2, \\ S_{102}(k+1) & \text{с вероятностью } p/2, \\ S_{011}(k+1) & \text{с вероятностью } q/2. \end{cases} \tag{48}$$

§ 5. Доказательство теоремы 1

В силу (40), (41) достаточно оценить величину $P_0(n) = P\{S_{000}(0) \Rightarrow S_{000}(n)\}$ снизу. Ясно, что

$$P_0(n) = \sum_{t_n} \mathbf{P}\{t_n\}, \tag{49}$$

где сумма берется по всем путям t_n длины n вида $S_{000}(0) \Rightarrow S_{000}(n)$.

Будем называть 3-путем любой путь длины 3 вида $S_{000}(k) \Rightarrow S_{000}(k+3)$. Будем также называть 2-путем любой путь длины 2 вида $S_{000}(k) \Rightarrow S_{000}(k+2)$.

Ограничимся сначала в сумме в правой части (49) путями t_n, проходящими только через главное и основные состояния (т.е. не выходящие на щупальца). Тогда нетрудно видеть, что любой такой путь t_n состоит из 3-путей и 2-путей.

Всего имеется шесть 3-путей:

$$\begin{aligned}
&S_{000} \to S_{100} \to S_{101} \to S_{000} \quad \text{с вероятностью } pq^2/6, \\
&S_{000} \to S_{100} \to S_{110} \to S_{000} \quad \text{с вероятностью } pq^2/6, \\
&S_{000} \to S_{010} \to S_{011} \to S_{000} \quad \text{с вероятностью } pq^2/6, \\
&S_{000} \to S_{010} \to S_{110} \to S_{000} \quad \text{с вероятностью } pq^2/6, \\
&S_{000} \to S_{001} \to S_{101} \to S_{000} \quad \text{с вероятностью } pq^2/6, \\
&S_{000} \to S_{001} \to S_{011} \to S_{000} \quad \text{с вероятностью } pq^2/6.
\end{aligned}$$

Поэтому

$$P\{S_{000}(k) \to S_{000}(k+3)\} = pq^2. \tag{50}$$

Всего имеется три 2-пути:

$$\begin{aligned}
&S_{000} \to S_{011} \to S_{000} \quad \text{с вероятностью } qp/3, \\
&S_{000} \to S_{101} \to S_{000} \quad \text{с вероятностью } qp/3, \\
&S_{000} \to S_{110} \to S_{000} \quad \text{с вероятностью } qp/3.
\end{aligned} \tag{51}$$

Поэтому

$$P\{S_{000}(k) \to S_{000}(k+2)\} = pq. \tag{52}$$

Оценим теперь величину $P_0(n)$ из (49), используя (50)–(52). Любой путь t_n, ограниченный основными состояниями, состоит из некоторого числа 2-путей n_2 и некоторого числа 3-путей n_3. При этом $2n_2 + 3n_3 = n$, $0 \leqslant n_2 \leqslant n/2$, а общее число путей равно

$$m = n_2 + n_3 = \frac{n + n_2}{3}.$$

Имеется $\binom{m}{n_2}$ способов разместить n_2 2-путей. Оставшиеся $m-n_2$ мест занимают n_3 3-путей. Поэтому имеем ($z=p/q$)

$$P_0(n)=\sum_{n_2=0}^{n/2}\binom{(n+n_2)/3}{n_2}(pq)^{n_2}(pq^2)^{n_3}=$$

$$=\sum_{n_2=0}^{n/2}\binom{(n+n_2)/3}{n_2}(pq)^{n_2}(pq^2)^{(n-2n_2)/3}=(pq^2)^{n/3}\sum_{n_2=0}^{n/2}\binom{(n+n_2)/3}{n_2}z^{n_2/3}. \quad (53)$$

Оценим снизу сумму в правой части (53). Максимум величины $\binom{(n+n_2)/3}{n_2}z^{n_2/3}$ по n_2 достигается при $n_2=a_0n$, где величина a_0 будет найдена далее. Тогда

$$\sum_{n_2=0}^{n/2}\binom{(n+n_2)/3}{n_2}z^{n_2/3}\geqslant\binom{n(1+a_0)/3}{a_0n}z^{a_0n/3}\geqslant$$

$$\geqslant\frac{1}{n}\sum_{n_2=0}^{n(1+a_0)/3}\binom{n(1+a_0)/3}{n_2}z^{n_2/3}=\frac{1}{n}(1+z^{1/3})^{n(1+a_0)/3}. \quad (54)$$

Для аккуратности оценим также сверху сумму в правой части (53). Имеем

$$\sum_{n_2=0}^{n}\binom{(n+n_2)/3}{n_2}z^{n_2/3}\leqslant n\binom{n(1+a_0)/3}{a_0n}z^{a_0n/3}\leqslant$$

$$\leqslant n\sum_{n_2=0}^{n(1+a_0)/3}\binom{n(1+a_0)/3}{n_2}z^{n_2/3}=n(1+z^{1/3})^{n(1+a_0)/3}.$$

В результате из (53) и (54) получаем

$$P_0(n)\geqslant\frac{1}{n}(pq^2)^{n/3}(1+z^{1/3})^{n(1+a_0)/3}. \quad (55)$$

Найдем теперь величину a_0 в (54), (55). Так как

$$\ln\binom{(n+n_2)/3}{n_2}\approx\frac{(n+n_2)}{3}h\Big(\frac{3n_2}{n+n_2}\Big),$$

то обозначая $n_2=an$, $0\leqslant a\leqslant 1/2$, введем функцию

$$f_1(p,a)=(1+a)h\Big(\frac{3a}{1+a}\Big)-a\ln(q/p),\quad 0\leqslant a\leqslant 1/2.$$

Величина a_0 максимизирует функцию $f_1(p,a)$ по $0\leqslant a\leqslant 1/2$. Заметим, что

$$f_1(p,a)=(1+a)\ln(1+a)-3a\ln(3a)-(1-2a)\ln(1-2a)-a\ln(q/p),$$

$$(f_1(p,a))'_a=\ln\frac{p(1+a)(1-2a)^2}{27qa^3},\quad (f_1(p,a))''_{aa}<0,$$

$$(f_1(p,a))'_{a=0}=\infty,\quad (f_1(p,a))'_{a=1/2}=-\infty.$$

Поэтому $a_0(p)$ является единственным корнем уравнения

$$27qa^3-p(1+a)(1-2a)^2=0=(27-31p)a^3+3pa-p.$$

Для этого корня имеем [8, п. 1.8-3]

$$a_0(p) = \left[\frac{p}{2(27-31p)}\right]^{1/3} \left\{ \left[1 + \sqrt{\frac{27(1-p)}{27-31p}}\right]^{1/3} + \left[1 - \sqrt{\frac{27(1-p)}{27-31p}}\right]^{1/3} \right\}.$$

При малых p имеем $3a_0(p) \approx p^{1/3}$. Так как $a_0 < 2$, то оценка (55) уступает оценке сверху (6)–(8) для $P_{\mathrm{e}}(3, n, p)$. Однако оценка (55) показывает, что при исследовании величины $P_0(n)$ нельзя ограничиваться только основными состояниями, а следует учитывать также и состояния на щупальцах.

Усилим теперь оценку (55), принимая во внимание также и состояния на щупальцах. Будем называть 2-петлей любой путь длины 2 с одинаковыми начальным и конечным состояниями (не обязательно состояниями S_{000}). Помимо 2-путей из (51) примерами 2-петель являются также

$$\begin{aligned}
&S_{011} \to S_{022} \to S_{011} \quad \text{с вероятностью } qp,\\
&S_{100} \to S_{201} \to S_{100} \quad \text{с вероятностью } qp/2,\\
&S_{100} \to S_{210} \to S_{100} \quad \text{с вероятностью } qp/2,\\
&S_{101} \to S_{202} \to S_{101} \quad \text{с вероятностью } qp, \quad \ldots
\end{aligned}$$

Такие 2-петли выходят на щупальца.

Рассмотрим пути t_n, состоящие из некоторого числа 3-путей n_3 и некоторого числа 2-петель k_2. Пусть мы разместили n_3 3-путей на $[1, n]$. После этого вставим k_2 2-петель в любые различные k_2 моментов на $[1, n]$. Если такая 2-петля попадает на начальное состояние 3-пути, то этот 3-путь просто полностью сдвигается вправо на два шага. Если такая 2-петля попадает на внутреннее состояние 3-пути, то часть этого 3-пути сдвигается вправо на два шага, чтобы вместить эту 2-петлю. Аналогичным образом 2-петли можно также вставлять в другие 2-петли.

Так как необходимо, чтобы $n = 3n_3 + 2k_2$, то

$$P_0(n) \geqslant \sum_{k_2=0}^{n/2} \binom{n}{k_2} (pq)^{k_2} (pq^2)^{n_3} = \sum_{k_2=0}^{n/2} \binom{n}{k_2} (pq)^{k_2} (pq^2)^{(n-2k_2)/3} =$$

$$= (pq^2)^{n/3} \sum_{k_2=0}^{n/2} \binom{n}{k_2} z^{k_2/3}. \tag{56}$$

Заметим, что

$$\binom{n}{k_2} z^{k_2/3} + \binom{n}{n-k_2} z^{(n-k_2)/3} \leqslant 2\binom{n}{k_2} z^{k_2/3}, \quad k_2 \leqslant n/2, \quad z < 1.$$

Тогда (56) можно продолжить следующим образом:

$$P_0(n) \geqslant \frac{1}{2}(pq^2)^{n/3} \sum_{k_2=0}^{n} \binom{n}{k_2} z^{k_2/3} = \frac{1}{2}(pq^2)^{n/3}(1 + z^{1/3})^n. \tag{57}$$

Поэтому из (57), (40) и (41) получаем

$$P_{\mathrm{e}}(n) \geqslant \frac{2}{3} P_0(n) \geqslant \frac{1}{3}(pq^2)^{n/3}(1 + z^{1/3})^n. \tag{58}$$

Из (58) следует (8) и теорема 1 (формулы (11), (12)). ▲

Доказательство формулы (14). Рассмотрим n-симплексный код вида $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$, где $\boldsymbol{x}_1$ имеет вначале $n/3$ единиц и затем $2n/3$ нулей, $\boldsymbol{x}_2$ имеет вначале $n/3$ нулей, затем $n/3$ единиц и потом $n/3$ нулей, а $\boldsymbol{x}_3$ имеет вначале $2n/3$ нулей и затем $n/3$ единиц. Тогда $w(\boldsymbol{x}_1) = w(\boldsymbol{x}_2) = w(\boldsymbol{x}_3) = n/3$ и $d_{12} = d_{13} = d_{23} = 2n/3$.

Пусть выход $\boldsymbol{y}$ имеет $u_1 n/3$ единиц на первых $n/3$ позициях, $u_2 n/3$ единиц на следующих $n/3$ позициях и $u_3 n/3$ единиц на последних $n/3$ позициях. Тогда

$$
\begin{aligned}
&d(\boldsymbol{x}_1, \boldsymbol{y})/n = (1 - u_1 + u_2 + u_3)/3,\\
&d(\boldsymbol{x}_2, \boldsymbol{y})/n = (1 + u_1 - u_2 + u_3)/3,\\
&d(\boldsymbol{x}_3, \boldsymbol{y})/n = (1 + u_1 + u_2 - u_3)/3.
\end{aligned}
$$

Так как $d(\boldsymbol{x}_1, \boldsymbol{y}) = d(\boldsymbol{x}_2, \boldsymbol{y}) = d(\boldsymbol{x}_3, \boldsymbol{y})$, то получаем $u_1 = u_2 = u_3$ и

$$
p(\boldsymbol{y}^n \,|\, \boldsymbol{x}_1) = p^{d(\boldsymbol{x}_1, \boldsymbol{y})} q^{n - d(\boldsymbol{x}_1, \boldsymbol{y})} = q^n z^{d(\boldsymbol{x}_1, \boldsymbol{y})} = q^n z^{(1+u)n/3}, \quad z = p/q < 1.
$$

Поэтому

$$
\begin{aligned}
&\mathbf{P}\big\{p(\boldsymbol{y}^n \,|\, \boldsymbol{x}_1) \approx p(\boldsymbol{y}^n \,|\, \boldsymbol{x}_2) \approx p(\boldsymbol{y}^n \,|\, \boldsymbol{x}_3)\big\} \sim \max_{0 \leqslant u \leqslant 1} \mathbf{P}\big\{p(\boldsymbol{y}^n \,|\, \boldsymbol{x}_1) \approx q^n z^{(1+u)n/3}\big\} \sim\\
&\sim \max_{0 \leqslant u \leqslant 1} \left\{ \binom{n}{un} p^{(1+u)n/3} q^{(2-u)n/3} \right\} \sim q^n \max_{0 \leqslant u \leqslant 1} \left\{ \binom{n}{un} z^{(1+u)n/3} \right\},
\end{aligned}
$$

и

$$
\frac{1}{n} \max_{0 \leqslant u \leqslant 1} \ln \mathbf{P}\{p(\boldsymbol{y}^n \,|\, \boldsymbol{x}_1)\} = \ln q + \max_{0 \leqslant u \leqslant 1} g(u), \tag{59}
$$

где

$$
g(u) = h(u) + (1 + u) \ln(z^{1/3}), \quad g'(u) = \ln \frac{1 - u}{u} + \ln(z^{1/3}), \quad g''(u) < 0.
$$

Для максимизирующего u_0 получаем

$$
u_0 = \frac{1}{1 + z^{-1/3}} = \frac{p^{1/3}}{p^{1/3} + q^{1/3}},
$$

и после несложных преобразований

$$
\ln q + g(u_0) = \ln\big(p^{1/3} q^{2/3} + p^{2/3} q^{1/3}\big). \tag{60}
$$

Из (59) и (60) следуют формулы (14) и (15). ▲

Автор благодарит Л.А. Бассалыго и Г.А. Кабатянского за полезные обсуждения и конструктивные критические замечания, улучшившие статью.

СПИСОК ЛИТЕРАТУРЫ

1. *Berlekamp E.R.* Block Coding with Noiseless Feedback. PhD Thesis. MIT, Cambridge, USA, 1964. Available at `http://hdl.handle.net/1721.1/14783`.
2. *Зигангиров К.Ш.* Верхние оценки вероятности ошибки для каналов с обратной связью // Пробл. передачи информ. 1970. Т. 6. № 2. С. 87–92. `http://mi.mathnet.ru/ppi1740`
3. *Бурнашев М.В.* О функции надежности двоичного симметричного канала с обратной связью // Пробл. передачи информ. 1988. Т. 24. № 1. С. 3–10. `http://mi.mathnet.ru/ppi681`

4. *Зигангиров К.Ш.* Оптимальная передача сообщений по двоичному симетричному каналу с обратной связью при нулевой скорости // Probl. Control Inform. Theory. 1978. V. 7. № 3. P. 183–198.
5. *Shannon C.E.* Probability of Error for Optimal Codes in a Gaussian Channel // Bell Syst. Tech. J. 1959. V. 38. № 3. P. 611–656. https://doi.org/10.1002/j.1538-7305.1959.tb03905.x
6. *Галлагер Р.* Теория информации и надежная связь. М.: Сов. радио, 1974.
7. *Пинскер М.С.* Вероятность ошибки при блоковой передаче по гауссовскому каналу без памяти с обратной связью // Пробл. передачи информ. 1968. Т. 4. № 4. С. 3–19. http://mi.mathnet.ru/ppi1868
8. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. М.: Наука, 1974.

Бурнашев Марат Валиевич
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
burn@iitp.ru

Поступила в редакцию
07.05.2022
После доработки
14.06.2022
Принята к публикации
21.06.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391 : 519.72

© 2022 г. **В.В. Прелов**

ОДНА ЭКСТРЕМАЛЬНАЯ ЗАДАЧА ДЛЯ ВЗАИМНОЙ ИНФОРМАЦИИ

Рассматривается задача о нахождении максимума взаимной информации $I(X;Y)$ двух случайных величин X и Y с конечным числом значений при условии, что задана лишь величина их склеивания, т.е. вероятность $\Pr\{X = Y\}$. Получены явные нижние и верхние границы для указанного максимума, являющиеся в некоторых случаях оптимальными.

Ключевые слова: взаимная информация, склеивание дискретных распределений вероятностей, вероятность ошибки.

DOI: 10.31857/S0555292322030020, **EDN:** DZXJIO

Посвящается памяти проф. Э.М. Габидулина

Рассмотрению различных экстремальных задач, связанных с взаимной информацией $I(X;Y)$ случайных величин X и Y, посвящено много работ. В частности, классическая задача о нахождении минимума $I(X;Y)$ дискретных случайных величин X и Y при условии, что заданы распределение вероятностей случайной величины X и вероятность ошибки $\Pr\{X \neq Y\} = \varepsilon$, известная как задача об ε-энтропии X, рассматривалась в [1–3]. Ряд других подобных экстремальных задач рассматривался, например, в [4–7], где также можно найти библиографию по этой тематике.

В настоящей статье рассматривается задача о получении явных нижних и верхних границ, а в некоторых случаях и точных значений для максимума взаимной информации $I(X;Y)$ двух случайных величин X и Y с конечным числом значений при условии, что задана лишь вероятность их склеивания, т.е. вероятность $\alpha = \Pr\{X = Y\}$ (или, что эквивалентно, задана вероятность ошибки $\varepsilon = \Pr\{X \neq Y\} = 1 - \alpha$).

Для формулировки полученных результатов введем некоторые определения и обозначения. Будем всегда предполагать, что рассматриваемые ниже случайные величины X и Y принимают n различных значений в некотором множестве $\mathcal{N}$, $|\mathcal{N}| = n$, $n \geqslant 2$. Для заданного параметра α, $0 \leqslant \alpha \leqslant 1$, положим

$$I_{\max}(\alpha) = \max_{(X,Y):\, \Pr\{X=Y\}=\alpha} I(X;Y), \tag{1}$$

где максимум берется по всевозможным совместным распределениям случайных величин X и Y со значениями в множестве $\mathcal{N}$, таким что $\Pr\{X = Y\} = \alpha$. Введем также функции

$$H_k(x) = -x \ln \frac{x}{k} - (1-x) \ln \frac{1-x}{n-k}, \quad 0 \leqslant x \leqslant 1, \quad k = 1, 2, \ldots, n-1, \tag{2}$$

$$J(x) = (1+x) \ln n - (1+x) \ln(1+x) + x \ln x, \quad x \geqslant 0, \tag{3}$$

$$\varphi(x,y) = (x+y) \ln(x+y) - x \ln x - y \ln y, \quad 0 \leqslant x, y \leqslant 1. \tag{4}$$

Заметим, что $H_k(p)$ – это энтропия распределения вероятностей $P = \{p_i,\ i \in \mathcal{N}\}$, где

$$p_i = \begin{cases} \dfrac{p}{k} & \text{при} \quad i = 1, 2, \ldots, k, \\ \dfrac{1-p}{n-k} & \text{при} \quad i = k+1, k+2, \ldots, n. \end{cases}$$

Сформулируем теперь основные результаты статьи.

Предложение 1. *Величина $I_{\max}(\alpha)$ принимает свое максимальное значение $\ln n$ тогда и только тогда, когда*

$$\alpha = \frac{k}{n}, \quad k = 0, 1, 2, \ldots, n-2, n$$

(т.е. k – любое целое между 0 и n, кроме $k = n-1$).

Доказательства этого и нижеследующих предложений приведены в Приложении. В следующем предложении рассматриваются малые значения α.

Предложение 2. *Для любого α, $0 \leqslant \alpha \leqslant \dfrac{1}{n}$, справедливы следующие утверждения*:

- *Для $I_{\max}(\alpha)$ справедлива нижняя граница*

$$I_{\max}(\alpha) \geqslant \begin{cases} J(\alpha), & \text{если } 0 \leqslant \alpha \leqslant \alpha^*, \\ H_1(\alpha), & \text{если } \alpha^* \leqslant \alpha \leqslant \dfrac{1}{n}, \end{cases} \tag{5}$$

  где $H_1(\alpha)$ и $J(\alpha)$ определены в (2) и (3), а α^*, $0 < \alpha^* < \dfrac{1}{3n-1}$, – единственное решение уравнения $H_1(\alpha^*) = J(\alpha^*)$ на указанном интервале;
- *Нижняя граница в (5) является оптимальной на отрезке $\alpha \in [0, \alpha_*]$, т.е.*

$$I_{\max}(\alpha) = J(\alpha), \quad \text{если } 0 \leqslant \alpha \leqslant \alpha_*, \tag{6}$$

  где α_*, $0 < \alpha_* < \alpha^*$, – единственное решение уравнения $H_1(2\alpha_*) = J(\alpha_*)$ на указанном интервале;
- *Для $I_{\max}(\alpha)$ справедлива верхняя граница*

$$I_{\max}(\alpha) \leqslant H_1(2\alpha), \quad \text{если } \alpha_* \leqslant \alpha \leqslant \frac{1}{3n-1}. \tag{7}$$

Замечание 1. Естественной нижней границей для $I_{\max}(\alpha)$ может служить величина $\max\limits_{(X,Y):\, \Pr\{X=Y\}=\alpha} I(X;Y)$ при условии, что X имеет равномерное распределение на множестве $\mathcal{N}$. Из результатов [7, следствие 2] следует, что такая граница имеет вид

$$I_{\max}(\alpha) \geqslant \begin{cases} \ln n - \varphi\Big(\dfrac{1}{n}, \alpha\Big), & \text{если } 0 \leqslant \alpha \leqslant \dfrac{1}{2n}, \\ \ln n - \varphi\Big(\dfrac{1}{n}, \dfrac{1}{n} - \alpha\Big), & \text{если } \dfrac{1}{2n} \leqslant \alpha \leqslant \dfrac{1}{n}, \end{cases} \tag{8}$$

где функция $\varphi(\cdot, \cdot)$ определена в (4). Однако эта нижняя граница (8) слабее, чем нижняя граница (5). Действительно, для этого следует лишь убедиться, что при любом расположении параметра α^* внутри отрезка $\left[0, \dfrac{1}{n}\right]$ (хотя при доказательстве формулы (5) будет показано, что на самом деле всегда имеет место неравен-

ство $\alpha^* < \dfrac{1}{3n-1}$) справедливы неравенства

$$J(\alpha) \geqslant \ln n - \varphi\Big(\frac{1}{n}, \alpha\Big), \quad \text{если } 0 \leqslant \alpha \leqslant \frac{1}{2n}$$

и

$$H_1(\alpha) \geqslant \ln n - \varphi\Big(\frac{1}{n}, \frac{1}{n} - \alpha\Big), \quad \text{если } \frac{1}{2n} \leqslant \alpha \leqslant \frac{1}{n}.$$

Первое из них следует из того, что разность

$$J(\alpha) - \ln n + \varphi\Big(\frac{1}{n}, \alpha\Big)$$

возрастает по α, а при $\alpha = 0$ она равна нулю, а второе – из того, что разность

$$H_1(\alpha) - \ln n + \varphi\Big(\frac{1}{n}, \frac{1}{n} - \alpha\Big),$$

как нетрудно проверить, убывает по α, а при $\alpha = \dfrac{1}{n}$ она тоже равна нулю.

При этом заметим, что в формуле (22) следствия 2 в [7] имеются опечатки, поэтому приведем здесь исправленный вид этой формулы, которая понадобится нам и в дальнейшем. А именно, утверждение, содержащее эту формулу, должно выглядеть следующим образом:

Если X имеет равномерное распределение на множестве из n элементов и $1 - \varepsilon = \dfrac{m}{n} + \beta$, $0 \leqslant \beta \leqslant \dfrac{1}{n}$, $m = 0, 1, 2, \ldots, n-1$, то

$$H^{(2)}_{\min}(P, \varepsilon) = \begin{cases} \varphi\Big(\dfrac{1}{n}, \varepsilon\Big), & \text{если } m = n-1, \\ \varphi\Big(\dfrac{1}{n}, \beta\Big), & \text{если } m = n-2 \text{ или } m < n-2 \text{ и } 0 \leqslant \beta \leqslant \dfrac{1}{2n}, \\ \varphi\Big(\dfrac{1}{n}, \dfrac{1}{n} - \beta\Big), & \text{если } m < n-2 \text{ и } \dfrac{1}{2n} \leqslant \beta \leqslant \dfrac{1}{n}. \end{cases} \tag{9}$$

Напомним также, что в [7] использовалось определение

$$H^{(2)}_{\min}(P, \varepsilon) = \min_{P_{Y|X}\colon \Pr\{Y \neq X\} = \varepsilon} H(X \,|\, Y), \tag{10}$$

где P – заданное распределение вероятностей случайной величины X, а $H(X \,|\, Y)$ – условная энтропия X при заданном Y.

Большим значениям α посвящено

Предложение 3. *Для любого α, $1 - \dfrac{1}{n} \leqslant \alpha \leqslant 1$, справедливы следующие утверждения*:

- *Для $I_{\max}(\alpha)$ и любого натурального n, $3 \leqslant n \leqslant 14$, справедлива нижняя граница*

$$I_{\max}(\alpha) \geqslant \begin{cases} J(1-\alpha), & \text{если } 1 - \dfrac{1}{2n-1} \leqslant \alpha \leqslant 1, \\ H_1(1-\alpha) - \varphi(1-\alpha, 1-\alpha), & \text{если } \widehat{\alpha} \leqslant \alpha \leqslant 1 - \dfrac{1}{2n-1}, \\ H_2(1-\alpha), & \text{если } 1 - \dfrac{1}{n} \leqslant \alpha \leqslant \widehat{\alpha}, \end{cases} \tag{11}$$

где $\widehat{\alpha}$, $1-\dfrac{1}{n}<\widehat{\alpha}<1-\dfrac{1}{2n-1}$, – *единственное решение уравнения*

$$H_1(1-\widehat{\alpha})-\varphi(1-\widehat{\alpha},1-\widehat{\alpha})=H_2(1-\widehat{\alpha})$$

на указанном интервале;

- *Для* $I_{\max}(\alpha)$ *и любого натурального* n, $n\geqslant 15$, *справедлива нижняя граница*

$$I_{\max}(\alpha)\geqslant\begin{cases}J(1-\alpha), & \text{если } \widetilde{\alpha}\leqslant\alpha\leqslant 1,\\ H_2(1-\alpha), & \text{если } 1-\dfrac{1}{n}\leqslant\alpha\leqslant\widetilde{\alpha},\end{cases}\tag{12}$$

где $\widetilde{\alpha}$, $1-\dfrac{1}{2n-1}<\widetilde{\alpha}<1$, – *единственное решение уравнения*

$$J(1-\widetilde{\alpha})=H_2(1-\widetilde{\alpha})$$

на указанном интервале;

- *Нижние границы в* (11) *и* (12) *являются оптимальными на отрезке* $\alpha\in[\overline{\alpha},1]$, *т.е.*

$$I_{\max}(\alpha)=J(1-\alpha),\quad \text{если } \overline{\alpha}\leqslant\alpha\leqslant 1,\tag{13}$$

где $\overline{\alpha}$, $\widetilde{\alpha}<\overline{\alpha}<1$, – *единственное решение уравнения*

$$H_1(1-\overline{\alpha})=J(1-\overline{\alpha})$$

на указанном интервале (*заметим, что* $\overline{\alpha}=1-\alpha^*$ – *см. предложение* 2);

- *Для* $I_{\max}(\alpha)$ *и любого натурального* n, $n\geqslant 3$, *справедлива верхняя граница*

$$I_{\max}(\alpha)\leqslant H_1(1-\alpha),\quad \text{если } 1-\frac{1}{2n-1}\leqslant\alpha\leqslant\overline{\alpha}.\tag{14}$$

Замечание 2. Как и в случае малых значений α, представляется естественным сравнить нижние границы (11) и (12) для $I_{\max}(\alpha)$ при больших значениях α со следующей нижней границей:

$$I_{\max}(\alpha)\geqslant\ln n-\varphi\Big(\frac{1}{n},1-\alpha\Big),\quad \text{если } 1-\frac{1}{n}\leqslant\alpha\leqslant 1,\tag{15}$$

полученной при условии, что X имеет равномерное распределение, поскольку в этом случае известно точное значение для $\min\limits_{Y\colon \Pr\{X=Y\}=\alpha} H(X\,|\,Y)$ (см. (9)). Нетрудно показать, что нижняя граница (15) слабее нижних границ (11) и (12). Для этого следует вначале заметить, что при доказательстве предложения 3 будет доказана формула (П.22). Поэтому, чтобы доказать, что (15) слабее (11) и (12), достаточно показать, что при любом $n\geqslant 3$ справедливы неравенства

$$J(1-\alpha)\geqslant\ln n-\varphi\Big(\frac{1}{n},1-\alpha\Big),\quad \text{если } 1-\frac{1}{2n-1}\leqslant\alpha\leqslant 1,\tag{16}$$

и

$$\begin{aligned}&H_1(1-\alpha)-\varphi(1-\alpha,1-\alpha)\geqslant\ln n-\varphi\Big(\frac{1}{n},1-\alpha\Big),\\ &\text{если } 1-\frac{1}{n}\leqslant\alpha\leqslant 1-\frac{1}{2n-1}.\end{aligned}\tag{17}$$

Неравенство (16) следует из того, что разность его левой и правой частей убывает по α, а при $\alpha=1$ она равна нулю. Для доказательства (17) следует заметить, что

разность его левой и правой частей есть вогнутая функция α, при $\alpha = 1 - \frac{1}{n}$ она равна нулю, а при $\alpha = 1 - \frac{1}{2n-1}$ положительна, так как

$$H_1\Big(\frac{1}{2n-1}\Big) - \varphi\Big(\frac{1}{2n-1}, \frac{1}{2n-1}\Big) = J\Big(\frac{1}{2n-1}\Big) > \ln n - \varphi\Big(\frac{1}{n}, \frac{1}{2n-1}\Big)$$

согласно (16).

В следующем предложении рассматриваются средние значения параметра α.

Предложение 4. *Для любого* α, $\frac{1}{n} < \alpha < 1 - \frac{1}{n}$, *справедливы следующие нижние границы*:

$$I_{\max}(\alpha) \geqslant \begin{cases} H_k(\alpha), & \text{если } \dfrac{k}{n} < \alpha \leqslant \widehat{\alpha}(k) \text{ и } k \leqslant n-3, \\ H_{k+1}(\alpha), & \text{если } \widehat{\alpha}(k) \leqslant \alpha \leqslant \dfrac{k+1}{n} \text{ и } k \leqslant n-3, \\ H_{n-2}(\alpha), & \text{если } \dfrac{n-2}{n} < \alpha < \dfrac{n-1}{n}, \end{cases} \tag{18}$$

где $\widehat{\alpha}(k) = \ln\Big(1 + \frac{1}{n-k-1}\Big) \,/\, \ln\Big(1 + \frac{n}{k(n-k-1)}\Big)$, *и*

$$I_{\max}(\alpha) \geqslant \begin{cases} \ln n - \varphi\Big(\frac{1}{n}, \alpha - \frac{k}{n}\Big), & \text{если } \dfrac{k}{n} < \alpha \leqslant \dfrac{2k+1}{2n} \text{ и } k \leqslant n-3, \\ \ln n - \varphi\Big(\frac{1}{n}, \frac{k+1}{n} - \alpha\Big), & \text{если } \dfrac{2k+1}{2n} \leqslant \alpha \leqslant \dfrac{k+1}{n} \text{ и } k \leqslant n-3, \\ \ln n - \varphi\Big(\frac{1}{n}, \alpha - \frac{n-2}{n}\Big), & \text{если } \dfrac{n-2}{n} < \alpha < \dfrac{n-1}{n}. \end{cases} \tag{19}$$

При этом нижняя граница (18) *лучше нижней границы* (19), *полученной при условии, что* X *имеет равномерное распределение* (*см.* (9)). *Более того, нижняя граница* (18) *является оптимальной, т.е. в* (18) *имеет место знак равенства, если* $\alpha = \frac{k}{n}$, $k = 0, 1, 2, \ldots, n-2, n$ (*см. также предложение* 1).

Замечание 3. Отметим также еще одну естественную и простую нижнюю границу для $I_{\max}(\alpha)$, справедливую для всех α, $0 \leqslant \alpha \leqslant 1$. А именно, справедливо неравенство

$$I_{\max}(\alpha) \geqslant \ln n - (1-\alpha)\ln(n-1) - h(\alpha), \tag{20}$$

где

$$h(x) = -x \ln x - (1-x)\ln(1-x), \quad 0 \leqslant x \leqslant 1,$$

– двоичная энтропия. Неравенство (20) является прямым следствием известного неравенства Фано

$$H(X \,|\, Y) \leqslant P_e \ln(n-1) + h(P_e), \quad \text{где} \quad P_e = \Pr\{X \neq Y\}. \tag{21}$$

Однако граница (20) слабее приведенных выше нижних границ для $I_{\max}(\alpha)$ при равномерном распределении X, так как в последнем случае известно точное значение для минимума условной энтропии $H(X \,|\, Y)$ (см. (9)), в то время как неравенство Фано дает лишь оценку сверху для $H(X \,|\, Y)$, не зависящую от распределения X. А нижние границы для $I_{\max}(\alpha)$ при равномерном распределении X в свою очередь слабее наших нижних границ (5), (11), (12) и (18) для $I_{\max}(\alpha)$, как указано в замечаниях 1 и 2 и предложении 4.

Доказательство предложения 1. Утверждение этого предложения почти очевидно. Действительно, информация $I(X;Y)$ принимает свое максимальное значение $\ln n$ тогда и только тогда, когда X имеет равномерное распределение, а условная энтропия $H(X\,|\,Y)$ равна нулю, что означает, что X должно быть детерминированной функцией Y. Поэтому $I_{\max}(\alpha) = \ln n$ тогда и только тогда, когда в $(n \times n)$-матрице совместного распределения X и Y можно расставить n чисел $\dfrac{1}{n}$ таким образом, чтобы в каждой строке и в каждом столбце этой матрицы стояло ровно по одному элементу $\dfrac{1}{n}$, а сумма диагональных элементов была равна α (а все остальные элементы этой матрицы должны быть равны нулю). Ясно, что это можно сделать, только если $\alpha = \dfrac{k}{n}$, где k – любое целое от 0 до n, кроме $k = n-1$. ▲

Доказательство предложения 2. Прежде всего отметим, что всюду в дальнейшем, когда речь идет о распределении вероятностей $P_X = P = \{p_i,\, i \in \mathcal{N}\}$ случайной величины X, будем для удобства считать, что компоненты p_i этого распределения упорядочены по убыванию, так что $p_1 \geqslant p_2 \geqslant \ldots \geqslant p_n > 0$. В частности, всегда будем считать, что $p_{\min} = \min\limits_{i \in \mathcal{N}} p_i = p_n$.

Для доказательства нижней границы (5) воспользуемся следующим результатом из [7, следствие 2]: для заданного распределения вероятностей $P = \{p_i,\, i \in \mathcal{N}\}$ случайной величины X справедливо равенство

$$\min_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y) = \begin{cases} \varphi(\alpha, p_n), & \text{если } 0 \leqslant \alpha \leqslant \alpha_0, \\ \varphi(p_n - \alpha, p_{n-1}), & \text{если } \alpha_0 \leqslant \alpha \leqslant p_n, \end{cases} \tag{П.1}$$

где α_0 – решение уравнения

$$\varphi(p_n - \alpha_0, p_{n-1}) = \varphi(\alpha_0, p_n), \tag{П.2}$$

а функция $\varphi(x,y)$ определена в (4). Так как $\varphi(x,y)$ возрастает по каждому из своих аргументов, а мы рассматриваем минимум по всем распределениям X и Y, таким что $\Pr\{Y = X\} = \alpha$ и задана только минимальная компонента p_n распределения P, то из (П.1) и (П.2) следует, что указанный минимум достигается на любом P, у которого $p_{n-1} = p_n$, а тогда получаем,что

$$\min_{(X,Y)\colon \Pr\{Y=X\}=\alpha,\ p_{\min}=p_n} H(X\,|\,Y) = \varphi(\alpha, p_n), \qquad \text{если } 0 \leqslant \alpha \leqslant \frac{p_n}{2}, \tag{П.3}$$

$$\min_{(X,Y)\colon \Pr\{Y=X\}=\alpha,\ p_{\min}=p_n} H(X\,|\,Y) = \varphi(p_n - \alpha, p_n) \leqslant$$

$$\leqslant \varphi(\alpha, p_n), \qquad \text{если } \frac{p_n}{2} \leqslant \alpha \leqslant p_n. \tag{П.4}$$

Теперь, учитывая (П.3) и (П.4), имеем

$$I_{\max}(\alpha) \geqslant \max_{p_n\colon \alpha \leqslant p_n \leqslant \frac{1}{n}} \max_{X\colon p_{\min}=p_n} \Big[H(X) - \min_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y)\Big] \geqslant$$

$$\geqslant \max_{p_n\colon \alpha \leqslant p_n \leqslant \frac{1}{n}} \big[H_1(p_n) - \varphi(\alpha, p_n)\big], \tag{П.5}$$

где $H_1(\cdot)$ определена в (2). При выводе (П.5) мы также воспользовались тем, что $\max\limits_{X\colon p_{\min}=p_n} H(X) = H_1(p_n)$. Нетрудно убедиться, что $H_1(p_n) - \varphi(\alpha, p_n)$ является во-

гнутой функцией от p_n, и ее максимум на отрезке $p_n \in \left[\alpha, \frac{1}{n}\right]$ достигается в точке

$$p_n = p_n^*(\alpha) = \frac{1}{n} - \frac{n-1}{n}\alpha,$$

если $p_n^*(\alpha) \in \left[\alpha, \frac{1}{n}\right]$, т.е. если $0 \leqslant \alpha \leqslant \frac{1}{2n-1}$. Если же $\frac{1}{2n-1} \leqslant \alpha \leqslant \frac{1}{n}$, то максимум $H_1(p_n) - \varphi(\alpha, p_n)$ достигается при $p_n = \alpha$. Нетрудно убедиться, что

$$H_1(p_n^*(\alpha)) - \varphi(\alpha, p_n^*(\alpha)) = J(\alpha),$$

где $J(\cdot)$ определено в (3). Таким образом, из (П.5) теперь следует, что

$$I_{\max}(\alpha) \geqslant \begin{cases} J(\alpha), & \text{если } 0 \leqslant \alpha \leqslant \frac{1}{2n-1}, \\ H_1(\alpha) - \varphi(\alpha, \alpha), & \text{если } \frac{1}{2n-1} < \alpha \leqslant \frac{1}{n}. \end{cases} \tag{П.6}$$

Теперь заметим, что для любых $\alpha \in \left[0, \frac{1}{n}\right]$ справедлива другая нижняя граница:

$$I_{\max}(\alpha) \geqslant H_1(\alpha), \quad \text{если } \alpha \in \left[0, \frac{1}{n}\right]. \tag{П.7}$$

Действительно, $I(X;Y) = H_1(\alpha)$, если компоненты p_{ij} матрицы $M = \|p_{ij}\|_{i,j=1}^n$ совместного распределения X и Y имеют вид

$$p_{ij} = \begin{cases} \alpha, & \text{если } i = j = 1, \\ \frac{1-\alpha}{n-1}, & \text{если } j = i+1,\ i = 2, 3, \ldots, n-1, \text{ и если } i = n,\ j = 2, \\ 0 & \text{в остальных случаях,} \end{cases}$$

поскольку в этом случае $H(X) = H_1(\alpha)$, а $H(X \,|\, Y) = 0$.

Очевидно, что нижняя граница (П.7) лучше (П.6), если $\frac{1}{2n-1} < \alpha \leqslant \frac{1}{n}$, так как $\varphi(\alpha, \alpha) > 0$ при $\alpha > 0$. Поэтому необходимо сравнить эти границы также и при $\alpha \in \left[0, \frac{1}{2n-1}\right]$. Имеем

$$[J(\alpha) - H_1(\alpha)]'_\alpha = \ln \frac{n(n-1)\alpha^2}{1-\alpha^2} < 0 \quad \text{при} \quad \alpha \in \left[0, \frac{1}{2n-1}\right],$$

так как $\frac{n(n-1)\alpha^2}{1-\alpha^2} \leqslant \frac{1}{4}$. Значит, функция $J(\alpha) - H_1(\alpha)$ убывает по α, а при $\alpha = 0$ она положительна. Покажем теперь, что

$$J\left(\frac{1}{3n-1}\right) - H_1\left(\frac{1}{3n-1}\right) < 0.$$

Действительно, нетрудно убедиться, что

$$J\left(\frac{1}{3n-1}\right) = \ln(3n-1) - \frac{3n-2}{3n-1}\ln 3 - \frac{2\ln 3}{3n-1}, \tag{П.8}$$

$$H_1\left(\frac{1}{3n-1}\right) = \ln(3n-1) - \frac{3n-2}{3n-1}\ln 3 - \frac{3n-2}{3n-1}\ln\left(1 + \frac{1}{3n-3}\right). \tag{П.9}$$

Неравенство $J\Big(\dfrac{1}{3n-1}\Big) - H_1\Big(\dfrac{1}{3n-1}\Big) < 0$ следует из того, что

$$\frac{3n-2}{3n-1}\ln\left(1+\frac{1}{3n-3}\right) < \frac{2\ln 3}{3n-1}.$$

Таким образом, имеем

$$J(\alpha) > H_1(\alpha), \quad \text{если } 0 \leqslant \alpha \leqslant \alpha^*,$$
$$J(\alpha) < H_1(\alpha), \quad \text{если } \alpha^* < \alpha < \frac{1}{3n-1},$$

где α^* – единственное решение уравнения $J(\alpha^*) = H_1(\alpha^*)$ на интервале $0 \leqslant \alpha^{(*)} < \dfrac{1}{3n-1}$. Нижняя граница (5) доказана.

Докажем теперь равенство (6) и верхнюю границу (7). Для этого заметим, что

$$I_{\max}(\alpha) = \max\{J_1(\alpha), J_2(\alpha)\}, \quad \text{если } 0 \leqslant \alpha \leqslant \frac{1}{2n}, \tag{П.10}$$

где

$$J_1(\alpha) = \max_{p_n:\, 2\alpha \leqslant p_n \leqslant \frac{1}{n}} \max_{X:\, p_{\min}=p_n} \Big[H(X) - \min_{Y:\, \Pr\{Y=X\}=\alpha} H(X\,|\,Y)\Big], \tag{П.11}$$

$$J_2(\alpha) = \max_{p_n:\, p_n \leqslant 2\alpha} \max_{X:\, p_{\min}=p_n} \Big[H(X) - \min_{Y:\, \Pr\{Y=X\}=\alpha} H(X\,|\,Y)\Big]. \tag{П.12}$$

Согласно (П.3) имеем

$$\min_{Y:\, \Pr\{Y=X\}=\alpha} H(X\,|\,Y) = \varphi(\alpha, p_n), \quad \text{если } p_n \geqslant 2\alpha,$$

и мы видели, что

$$\max_{p_n:\, 2\alpha \leqslant p_n \leqslant \frac{1}{n}} \max_{X:\, p_{\min}=p_n} [H(X) - \varphi(\alpha, p_n)] = J(\alpha)$$

и достигается в точке

$$p_n = p_n^*(\alpha) = \frac{1}{n} - \frac{n-1}{n}\alpha.$$

Однако последнее равенство для максимумов верно, лишь если $p_n^*(\alpha)$ удовлетворяет нашему ограничению $p_n^*(\alpha) \geqslant 2\alpha$ (ранее у нас было другое ограничение $p_n^*(\alpha) \geqslant \alpha$), что эквивалентно условию $0 \leqslant \alpha \leqslant \dfrac{1}{3n-1}$. Если же $\dfrac{1}{3n-1} \leqslant \alpha \leqslant \dfrac{1}{2n}$, то

$$\max_{p_n:\, 2\alpha \leqslant p_n \leqslant \frac{1}{n}} \max_{X:\, p_{\min}=p_n} [H(X) - \varphi(\alpha, p_n)]$$

достигается в точке $p_n = 2\alpha$. Таким образом, справедливо равенство

$$J_1(\alpha) = \begin{cases} J(\alpha), & \text{если } 0 \leqslant \alpha \leqslant \dfrac{1}{3n-1}, \\ H_1(2\alpha) - \varphi(\alpha, 2\alpha), & \text{если } \dfrac{1}{3n-1} < \alpha \leqslant \dfrac{1}{2n}. \end{cases} \tag{П.13}$$

В случае, когда $p_n \leqslant 2\alpha$, точное значение

$$\max_{p_n:\, p_n \leqslant 2\alpha} \max_{X:\, p_{\min}=p_n} \Big[H(X) - \min_{Y:\, \Pr\{Y=X\}=\alpha} H(X\,|\,Y)\Big]$$

неизвестно, и поэтому можно лишь утверждать, что

$$J_2(\alpha) \leqslant \max_{X:\, p_n \leqslant 2\alpha} H(X) = H_1(2\alpha), \quad 0 \leqslant \alpha \leqslant \frac{1}{2n}. \tag{П.14}$$

Поэтому из (П.10)–(П.14) теперь следует, что

$$I_{\max}(\alpha) = J(\alpha), \quad \text{если } J(\alpha) \geqslant H_1(2\alpha). \tag{П.15}$$

Найдем ограничения на α, при которых $J(\alpha) \geqslant H_1(2\alpha)$. Нетрудно убедиться, что разность $J(\alpha) - H_1(2\alpha)$ убывает по α, при $\alpha = 0$ она положительна, а при $\alpha = \dfrac{1}{3n-1}$ отрицательна. Действительно, имеем

$$J\Big(\frac{1}{3n-1}\Big) = \ln(3n-1) - \frac{3n}{3n-1}\ln 3 \quad \text{(см. (П.8))},$$
$$H_1\Big(\frac{2}{3n-1}\Big) = \ln(3n-1) - \frac{3n}{3n-1}\ln 3 + \frac{1}{3n-1}\ln\frac{27}{4} > J\Big(\frac{1}{3n-1}\Big).$$

Поэтому $J(\alpha) \geqslant H_1(2\alpha)$ при $0 \leqslant \alpha \leqslant \alpha_*$ и $J(\alpha) \leqslant H_1(2\alpha)$ при $\alpha_* < \alpha < \dfrac{1}{3n-1}$, где α_* – решение уравнения $J(\alpha_*) = H_1(2\alpha_*)$, причем очевидно, что $\alpha_* < \alpha^* < \dfrac{1}{3n-1}$, где α^* – решение уравнения $J(\alpha^*) = H_1(\alpha^*)$. Теперь из (П.13)–(П.15) следует равенство (6) и верхняя граница (7).

В заключение отметим, что второе из равенств в (П.13) приводит к нижней границе

$$I_{\max}(\alpha) \geqslant H_1(2\alpha) - \varphi(\alpha, 2\alpha), \quad \text{если } \frac{1}{3n-1} \leqslant \alpha \leqslant \frac{1}{2n}, \tag{П.16}$$

которая, однако, слабее доказанной выше нижней границы (5). Для этого следует лишь убедиться, что

$$H_1(\alpha) > H_1(2\alpha) - \varphi(\alpha, 2\alpha), \quad \text{если } \frac{1}{3n-1} \leqslant \alpha \leqslant \frac{1}{2n}.$$

Нетрудно проверить, что разность $H_1(\alpha) - [H_1(2\alpha) - \varphi(\alpha, 2\alpha)]$ возрастает по α на отрезке $\alpha \in \big[\dfrac{1}{3n-1}, \dfrac{1}{2n}\big]$, а при $\alpha = \dfrac{1}{3n-1}$ она положительна. Действительно, легко убедиться, что на этом отрезке

$$[H_1(\alpha) - H_1(2\alpha) + \varphi(\alpha, 2\alpha)]'_\alpha = \ln\frac{27(n-1)\alpha(1-\alpha)}{(1-2\alpha)^2} > 0,$$

а

$$H_1\Big(\frac{1}{3n-1}\Big) - H_1\Big(\frac{2}{3n-1}\Big) + \varphi\Big(\frac{1}{3n-1}, \frac{2}{3n-1}\Big) =$$
$$= -\frac{3n-2}{3n-1}\ln\Big(1 + \frac{1}{3n-3}\Big) + \frac{2\ln 3}{3n-1} \geqslant 0.$$

На этом доказательство предложения 2 заканчивается. ▲

Доказательство предложения 3. Отметим сразу, что хотя доказательство этого предложения в идейном плане вполне аналогично вышеприведенному доказательству предложения 2, но оно имеет ряд существенных отличий. В частности, в рассматриваемом сейчас случае в отличие от предложения 2 нижние границы для $I_{\max}(\alpha)$ существенно различаются в зависимости от величины параметра n, когда $n \leqslant 14$ или когда $n \geqslant 15$.

Вначале снова воспользуемся следующим результатом из [7, следствие 2]: для заданного распределения вероятностей $P = \{p_i,\ i \in \mathcal{N}\}$ случайной величины X справедливо равенство

$$\min_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y) = \varphi(1-\alpha, p_n), \quad \text{если } 1-p_n \leqslant \alpha \leqslant 1. \tag{П.17}$$

Воспользовавшись (П.17), получаем

$$I_{\max}(\alpha) = \max\{J_1'(\alpha), J_2'(\alpha)\}, \tag{П.18}$$

где

$$J_1'(\alpha) = \max_{p_n\colon 1-\alpha\leqslant p_n\leqslant \frac{1}{n}}\ \max_{X\colon p_{\min}=p_n}\left[H(X) - \min_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y)\right] =$$
$$= \max_{p_n\colon 1-\alpha\leqslant p_n\leqslant \frac{1}{n}} [H_1(p_n) - \varphi(1-\alpha, p_n)], \tag{П.19}$$

$$J_2'(\alpha) = \max_{p_n\colon p_n\leqslant 1-\alpha}\ \max_{X\colon p_{\min}=p_n}\left[H(X) - \min_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y)\right] \leqslant$$
$$\leqslant \max_{p_n\colon p_n\leqslant 1-\alpha} H(X) = H_1(1-\alpha). \tag{П.20}$$

При доказательстве предложения 2 было показано, что

$$\max_{p_n\colon \alpha\leqslant p_n\leqslant \frac{1}{n}} [H_1(p_n) - \varphi(\alpha, p_n)] =$$
$$= \begin{cases} J(\alpha), & \text{если } 0 \leqslant \alpha \leqslant \dfrac{1}{2n-1}, \\ H_1(\alpha) - \varphi(\alpha,\alpha), & \text{если } \dfrac{1}{2n-1} < \alpha \leqslant \dfrac{1}{n}, \end{cases} \tag{П.21}$$

и этот максимум достигается в точке

$$p_n = p_n^*(\alpha) = \frac{1}{n} - \frac{n-1}{n}\alpha,$$

если $0 \leqslant \alpha \leqslant \dfrac{1}{2n-1}$, и в точке $p_n = \alpha$, если $\dfrac{1}{2n-1} \leqslant \alpha \leqslant \dfrac{1}{n}$.

В рассматриваемом сейчас случае больших значений α нахождение максимума

$$\max_{p_n\colon 1-\alpha\leqslant p_n\leqslant \frac{1}{n}} [H_1(p_n) - \varphi(1-\alpha, p_n)]$$

отличается от нахождения максимума в (П.21) лишь заменой параметра α на $1-\alpha$. Поэтому из (П.18) и (П.19) для $I_{\max}(\alpha)$ получаем следующую нижнюю границу:

$$I_{\max}(\alpha) \geqslant J_1'(\alpha) =$$
$$= \begin{cases} J(1-\alpha), & \text{если } 1-\dfrac{1}{2n-1} \leqslant \alpha \leqslant 1, \\ H_1(1-\alpha) - \varphi(1-\alpha, 1-\alpha), & \text{если } 1-\dfrac{1}{n} \leqslant \alpha \leqslant 1-\dfrac{1}{2n-1}, \end{cases} \tag{П.22}$$

причем первое из этих равенств для $J_1'(\alpha)$ достигается в точке

$$p_n = p_n^*(1-\alpha) = \frac{1}{n} - \frac{n-1}{n}(1-\alpha),$$

а второе – в точке $p_n = 1-\alpha$.

Теперь заметим, что справедлива другая нижняя граница:

$$I_{\max}(\alpha) \geqslant H_2(1-\alpha) \quad \text{для всех } \alpha \in \Big[1-\frac{1}{n}, 1\Big]. \tag{П.23}$$

Действительно, эта нижняя граница следует из того, что $I(X;Y) = H_2(1-\alpha)$, если компоненты матрицы $M = \|p_{ij}\|_{i,j=1}^{n}$ совместного распределения X и Y задаются равенствами

$$p_{ij} = \begin{cases} \dfrac{\alpha}{n-2}, & \text{если } i = j = 1, 2, \ldots, n-2, \\ \dfrac{1-\alpha}{2}, & \text{если } i = n-1,\ j = n \text{ и если } i = n,\ j = n-1, \\ 0 & \text{в остальных случаях,} \end{cases}$$

поскольку для этого совместного распределения $H(X) = H_2(1-\alpha)$ и $H(X\,|\,Y) = 0$. Отметим, что в отличие от случая малых значений α, когда была справедлива нижняя граница $I_{\max}(\alpha) \geqslant H_1(\alpha)$ (см. (П.7)), в данном случае больших значений α мы не можем утверждать, что $I_{\max}(\alpha) \geqslant H_1(1-\alpha)$, так как не удается построить совместное распределение X и Y такое, чтобы $H(X) = H_1(1-\alpha)$, а $H(X\,|\,Y) = 0$.

Таким образом, нам необходимо сравнить нижние границы (П.22) и (П.23) для различных значений α и выбрать из них наилучшую. Сравним вначале первую из границ в (П.22) с (П.23) на интервале $\alpha \in \big[1-\frac{1}{2n-1}, 1\big]$. Легко проверить, что разность $J(1-\alpha) - H_2(1-\alpha)$ возрастает по α на данном интервале, так как

$$\big(J(1-\alpha) - H_2(1-\alpha)\big)'_{\alpha} = \ln\frac{2\alpha(2-\alpha)}{n(n-2)(1-\alpha)^2} > 0,$$

а при $\alpha = 1$ эта разность положительна. Поэтому необходимо сравнить значения $J(1-\alpha)$ и $H_2(1-\alpha)$ при $\alpha = 1 - \frac{1}{2n-1}$. Имеем

$$J\Big(\frac{1}{2n-1}\Big) = \ln(2n-1) - \frac{2n\ln 2}{2n-1}, \tag{П.24}$$

$$H_2\Big(\frac{1}{2n-1}\Big) = \ln(2n-1) - \frac{(2n-3)\ln 2}{2n-1} - \frac{2n-2}{2n-1}\ln\frac{n-1}{n-2}, \tag{П.25}$$

так что

$$H_2\Big(\frac{1}{2n-1}\Big) = J\Big(\frac{1}{2n-1}\Big) + \frac{1}{2n-1}\Big[\ln 8 - 2(n-1)\ln\frac{n-1}{n-2}\Big]. \tag{П.26}$$

Замечая, что функция $(n-1)\ln\frac{n-1}{n-2}$ убывает по n, и учитывая, что $\ln 8 \approx 2{,}07944$, а

$$2(n-1)\ln\frac{n-1}{n-2} \approx \begin{cases} 2{,}08111, & \text{если } n = 14, \\ 2{,}07502, & \text{если } n = 15, \end{cases} \tag{П.27}$$

получаем, что нижняя граница (П.22) лучше (П.23) при всех $\alpha \in \big[1-\frac{1}{2n-1}, 1\big]$, если $3 \leqslant n \leqslant 14$, и при $\alpha \in [\widetilde{\alpha}, 1]$, если $n \geqslant 15$, где $\widetilde{\alpha}$, $1 - \frac{1}{2n-1} < \widetilde{\alpha} < 1$, – единственное решение уравнения

$$J(1-\widetilde{\alpha}) = H_2(1-\widetilde{\alpha})$$

на указанном интервале. Если же $n \geqslant 15$ и $\alpha \in \big[1 - \frac{1}{2n-1}, \widetilde{\alpha}\big]$, то граница (П.23) лучше (П.22). Таким образом, мы доказали справедливость границ (11) и (12) на интервале $\big[1 - \frac{1}{2n-1}, 1\big]$.

Сравним теперь вторую из границ в (П.22) с границей (П.23) на интервале $\alpha \in \in \big[1 - \frac{1}{n}, 1 - \frac{1}{2n-1}\big]$. Для этого заметим, что разность

$$[H_1(1-\alpha) - \varphi(1-\alpha, 1-\alpha)] - H_2(1-\alpha)$$

возрастает по α и при $\alpha = 1 - \frac{1}{n}$ она отрицательна, так как легко проверить, что

$$H_1\Big(\frac{1}{n}\Big) - \varphi\Big(\frac{1}{n}, \frac{1}{n}\Big) - H_2\Big(\frac{1}{n}\Big) = -\frac{3\ln 2}{n} + \frac{n-1}{n}\ln\Big(1 + \frac{1}{n-2}\Big) < 0,$$

поскольку $\ln 8 > 2$, а

$$(n-1)\ln\big(1 + \frac{1}{n-2}\big) \leqslant \frac{n-1}{n-2} < 2 \quad \text{при } n > 2.$$

Теперь, замечая, что при $\alpha = 1 - \frac{1}{2n-1}$ эта разность равна

$$-\frac{1}{2n-1}\left[\ln 8 - 2(n-1)\ln\frac{n-1}{n-2}\right],$$

так как нетрудно проверить, что

$$H_1\Big(\frac{1}{2n-1}\Big) - \varphi\Big(\frac{1}{2n-1}, \frac{1}{2n-1}\Big) = J\Big(\frac{1}{2n-1}\Big),$$

а для $J\big(\frac{1}{2n-1}\big)$ справедливо равенство (П.26), то снова, учитывая (П.27), убеждаемся в справедливости нижних границ (11) и (12) и на интервале $\alpha \in \big[1 - \frac{1}{n}, 1 - \frac{1}{2n-1}\big]$.

Для доказательства равенства (13) и верхней границы (14) достаточно проверить, что

$$H_1(1-\alpha) \geqslant H_2(1-\alpha) \quad \text{для всех } \alpha,\ 1 - \frac{1}{n} \leqslant \alpha \leqslant 1, \tag{П.28}$$

и что

$$J(1-\alpha) \geqslant H_1(1-\alpha) \quad \text{для } \alpha \in [\overline{\alpha}, 1], \tag{П.29}$$

$$J(1-\alpha) \leqslant H_1(1-\alpha) \quad \text{для } \alpha \in \Big[1 - \frac{1}{2n-1}, \overline{\alpha}\Big], \tag{П.30}$$

где $\overline{\alpha}$, $\overline{\alpha} \in \big[1 - \frac{1}{2n-1}, 1\big]$ – единственное решение уравнения

$$J(1-\overline{\alpha}) = H_1(1-\overline{\alpha})$$

на указанном интервале. Справедливость неравенства (П.28) следует из того, что разность $H_1(1-\alpha) - H_2(1-\alpha)$ убывает по α на отрезке $\big[1 - \frac{1}{n}, 1\big]$, а при $\alpha = 1$ эта разность положительна. Для доказательства (П.29) и (П.30) следует лишь убедиться в том, что разность $J(1-\alpha) - H_1(1-\alpha)$ возрастает по α на отрезке $\big[1 - \frac{1}{2n-1}, 1\big]$, отрицательна при $\alpha = 1 - \frac{1}{2n-1}$ и положительна при $\alpha = 1$. На этом доказательство предложения 3 заканчивается. ▲

Доказательство предложения 4. Вначале заметим, что для рассматриваемых сейчас "средних" значений параметра α, когда $\frac{1}{n} < \alpha < 1 - \frac{1}{n}$, явное выражение для $\min\limits_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y)$ при заданном распределении X получить не удается в отличие от малых и больших значений α. Поэтому для средних значений α не удается получить и аналоги нижних границ для $I_{\max}(\alpha)$, представленных в предложениях 2 и 3.

Для средних значений α естественно рассмотреть два типа явных нижних границ для $I_{\max}(\alpha)$ и сравнить их между собой: когда совместное распределение X и Y таково, что $\Pr\{Y = X\} = \alpha$, а $H(X\,|\,Y) = 0$, и когда X имеет равномерное распределение, поскольку в этом случае точное значение для $\min\limits_{Y\colon \Pr\{Y=X\}=\alpha} H(X\,|\,Y)$ известно (см. (9)). В первом случае, когда предполагается, что условная энтропия $H(X\,|\,Y)$ равна нулю и $\Pr\{Y = X\} = \alpha$, матрица совместного распределения X и Y должна содержать в каждой строке и каждом столбце ровно по одному ненулевому элементу, а сумма диагональных элементов должна быть равна α. Как нетрудно убедиться, в случае, когда

$$\frac{k}{n} < \alpha \leqslant \frac{k+1}{n}, \quad k = 1, 2, \ldots, n-3,$$

то для того чтобы максимизировать количество информации $I(X;Y) = H(X)$, необходимо выбрать наилучший из следующих двух способов расположения ненулевых элементов в этой матрице совместного распределения:

1) k чисел $\frac{\alpha}{k}$ расположить на диагонали, а остальные $n-k$ чисел $\frac{1-\alpha}{n-k}$ расположить по одному в каждой из оставшихся незанятых строк и каждом незанятом столбце, и
2) $k+1$ чисел $\frac{\alpha}{k+1}$ расположить на диагонали, а остальные $n-k-1$ чисел $\frac{1-\alpha}{n-k-1}$ расположить по одному в каждой из оставшихся незанятых строк и каждом незанятом столбце.

В первом из этих случаев $H(X) = H_k(\alpha)$, а во втором $H(X) = H_{k+1}(\alpha)$, где $H_i(x)$ определено в (2). Если же $k = n-2$, т.е. $\frac{n-2}{n} < \alpha < \frac{n-1}{n}$, то равенство $H(Y\,|\,X) = 0$ возможно только в первом из указанных выше случаев расположения ненулевых элементов в матрице совместного распределения, так что в этом случае $H(X) = H_{n-2}(\alpha)$. Теперь заметим, что

$$\max\{H_k(\alpha), H_{k+1}(\alpha)\} = \begin{cases} H_k(\alpha), & \text{если } \dfrac{k}{n} < \alpha \leqslant \widehat{\alpha}(k), \\ H_{k+1}(\alpha), & \text{если } \widehat{\alpha}(k) \leqslant \alpha \leqslant \dfrac{k+1}{n}, \end{cases} \tag{П.31}$$

где

$$\widehat{\alpha}(k) = \ln\Big(1 + \frac{1}{n-k-1}\Big) \Big/ \ln\Big(1 + \frac{n}{k(n-k-1)}\Big), \quad k = 1, 2, \ldots, n-3. \tag{П.32}$$

Действительно, равенства (П.31) и (П.32) следуют из того факта, что разность $H_k(\alpha) - H_{k+1}(\alpha)$ убывает по α,

$$H_k\Big(\frac{k}{n}\Big) - H_{k+1}\Big(\frac{k}{n}\Big) > 0, \quad H_k\Big(\frac{k+1}{n}\Big) - H_{k+1}\Big(\frac{k+1}{n}\Big) < 0,$$

так что $\widehat{\alpha}(k)$ представляет собой решение уравнения

$$H_k(\widehat{\alpha}(k)) = H_{k+1}(\widehat{\alpha}(k)).$$

Таким образом, справедливость нижней границы (18) установлена, причем ее оптимальность при

$$\alpha = \frac{k}{n}, \quad k = 0, 1, 2, \ldots, n-2, n,$$

следует из предложения 1. Нижняя граница (19) является прямым следствием равенства (9). Покажем теперь, что граница (18) лучше, чем (19).

Пусть вначале

$$\frac{k}{n} < \alpha \leqslant \frac{2k+1}{2n}, \quad k = 1, 2, \ldots, n-3.$$

В этом случае достаточно показать, что

$$H_k(\alpha) > \ln n - \varphi\Big(\frac{1}{n}, \alpha - \frac{k}{n}\Big).$$

Для этого заметим, что разность

$$H_k(\alpha) - \ln n + \varphi\Big(\frac{1}{n}, \alpha - \frac{k}{n}\Big)$$

является вогнутой функцией α, а при $\alpha = \dfrac{k}{n}$ она равна нулю. Поэтому, чтобы доказать, что в рассматриваемом случае (18) лучше (19), достаточно убедиться, что

$$H_k\Big(\frac{2k+1}{2n}\Big) > \ln n - \varphi\Big(\frac{1}{n}, \frac{1}{2n}\Big).$$

Простые выкладки показывают, что

$$H_k\Big(\frac{2k+1}{2n}\Big) = \ln n - \frac{2k+1}{2n}\ln\Big(1 + \frac{1}{2k}\Big) - \frac{2n-2k-1}{2n}\ln\Big(1 - \frac{1}{2(n-k)}\Big),$$

$$\ln n - \varphi\Big(\frac{1}{n}, \frac{1}{2n}\Big) = \ln n - \frac{1}{2n}\ln\frac{27}{4}.$$

Требуемое неравенство $H_k\Big(\dfrac{2k+1}{2n}\Big) > \ln n - \varphi\Big(\dfrac{1}{n}, \dfrac{1}{2n}\Big)$ теперь следует из того, что

$$\frac{2k+1}{2n}\ln\Big(1 + \frac{1}{2k}\Big) + \frac{2n-2k-1}{2n}\ln\Big(1 - \frac{1}{2(n-k)}\Big) \leqslant$$

$$\leqslant \frac{2k+1}{4nk} - \frac{2n-2k-1}{4n(n-k)} < \frac{1}{2n}\ln\frac{27}{4}.$$

Значит, в рассматриваемом случае, когда

$$\frac{k}{n} < \alpha \leqslant \frac{2k+1}{2n}, \quad k = 1, 2, \ldots, n-3,$$

нижняя граница (18) лучше (19).

Аналогично доказывается, что нижняя граница (18) лучше (19) и в двух оставшихся случаях, когда

$$\frac{2k+1}{2n} \leqslant \alpha \leqslant \frac{k+1}{n}, \quad k = 1, 2, \ldots, n-3,$$

и когда $k = n-2$, т.е.

$$\frac{n-2}{n} < \alpha < \frac{n-1}{n}.$$

В первом из них доказывается, что

$$H_{k+1}(\alpha) > \ln n - \varphi\Big(\frac{1}{n}, \frac{k+1}{n} - \alpha\Big),$$

а во втором – что

$$H_{n-2}(\alpha) > \ln n - \varphi\Big(\frac{1}{n}, \alpha - \frac{n-2}{n}\Big).$$

На этом доказательство предложения 4 заканчивается. ▲

В заключение автор выражает искреннюю благодарность рецензенту за указание на ряд неточностей в статье, устранение которых привело к улучшению ее качества.

СПИСОК ЛИТЕРАТУРЫ

1. *Ерохин В.* ε-Энтропия дискретного случайного объекта // Теория вероятн. и ее примен. 1958. Т. 3. № 1. С. 103–107. http://mi.mathnet.ru/tvp4919
2. *Berger T.* Rate Distortion Theory: A Mathematical Basis for Data Compression. Englewood Cliffs, NJ: Prentice Hall, 1971.
3. *Ho S.-W., Verdú S.* On the Interplay between Conditional Entropy and Error Probability // IEEE Trans. Inform. Theory. 2010. V. 56. № 12. P. 5930–5942. https://doi.org/10.1109/TIT.2010.2080891
4. *Пинскер М.С.* Об оценке информации через вариацию // Пробл. передачи информ. 2005. Т. 41. № 2. С. 3–8. http://mi.mathnet.ru/ppi91
5. *Zhang Z.* Estimating Mutual Information via Kolmogorov Distance // IEEE Trans. Inform. Theory. 2007. V. 53. № 9. P. 3280–3282. https://doi.org/10.1109/TIT.2007.903122
6. *Прелов В.В.* Обобщение одной задачи Пинскера // Пробл. передачи информ. 2011. Т. 47. № 2. С. 17–37. http://mi.mathnet.ru/ppi2043
7. *Прелов В.В.* О некоторых экстремальных задачах для взаимной информации и энтропии // Пробл. передачи информ. 2016. Т. 52. № 4. С. 3–13. http://mi.mathnet.ru/ppi2218

Прелов Вячеслав Валерьевич
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
prelov@iitp.ru

Поступила в редакцию
24.05.2022
После доработки
09.08.2022
Принята к публикации
09.08.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391.1 : 519.725

© 2022 г. **И.Ю. Могильных, Ф.И. Соловьева**

О ВЕСОВОМ СПЕКТРЕ КЛАССА КОДОВ С ПАРАМЕТРАМИ КОДОВ РИДА – МАЛЛЕРА[1]

Приведен новый метод построения дважды экспоненциального класса двоичных кодов с параметрами кодов Рида – Маллера. Исследованы весовой спектр и дистанционная инвариантность предложенных кодов. В построенном классе кодов с параметрами кода Рида – Маллера показано существование кодов с тем же весовым распределением, что и у кода Рида – Маллера, а также кодов с весовым распределением, отличным от него. Установлено, что все коды с параметрами кода Рида – Маллера, полученные конструкцией Васильева – Пулатова, отличные от расширенных совершенных кодов, либо эквивалентны оригинальным кодам Рида – Маллера, либо имеют отличное от них распределение расстояний.

Ключевые слова: код Рида – Маллера, код с параметрами кода Рида – Маллера, весовой спектр, дистанционная инвариантность, обобщенная конструкция Пулатова, свитчинговая конструкция.

DOI: 10.31857/S0555292322030032, **EDN:** EAAOCA

§ 1. Введение

Двоичный линейный *код Рида – Маллера порядка* r, $0 \leqslant r \leqslant m$, обозначаемый через $RM(r,m)$, определяется как совокупность векторов длины 2^m для любого $m \geqslant 1$, отвечающих булевым функциям от m переменных степени не более чем r. Код $RM(r,m)$ имеет мощность 2^k, $k = \sum\limits_{i=0}^{r} \binom{m}{i}$, и *кодовое расстояние* 2^{m-r}. Коды Рида – Маллера обладают рядом хороших свойств. Известно, что код $RM(m-r-1,m)$ является дуальным к коду $RM(r,m)$. Код $RM(m-2,m)$ – расширенный двоичный код Хэмминга, а $RM(1,m)$ – расширенный двоичный код Адамара длины $n = 2^m$, $m \geqslant 2$. Для любых допустимых r и m код $RM(r,m)$ обладает базисом, состоящим из кодовых слов минимального веса (см. [1]). Напомним, что код Рида – Маллера $RM(r+1,m+1)$ представим широко известной в литературе конструкцией Плоткина:

$$\{(x+y\,|\,x) :\ x \in RM(r+1,m),\ y \in RM(r,m)\} \tag{1}$$

(см., например, [1]).

Задача описания весового спектра классических двоичных кодов Рида – Маллера все еще остается открытой, несмотря на значительные усилия и полученные результаты ряда исследователей (см. [2], а также недавние работы [3, 4]).

[1] Исследование выполнено за счет гранта Российского научного фонда № 22-21-00135, https://rscf.ru/project/22-21-00135/

Коды Рида – Маллера обладают хорошими процедурами кодирования и декодирования и в течение многих десятилетий активно используются как на практике, так и в теоретических исследованиях в области теории кодирования и криптографии. Кроме того, в теории блок-схем представляют интерес 3-схемы, получаемые из совокупностей кодовых слов фиксированного веса. Каждая такая схема обладает нетривиальными комбинаторно-алгебраическими свойствами.

В 2009 г. в [5] были предложены полярные схемы, имеющие те же самые параметры, что и схемы кодов Рида – Маллера, но не изоморфные им. Этот результат опроверг широко известную гипотезу Хамады для схем, выдвинутую в 1973 г. в работе [6]. Расширение двоичного кода, натянутого на блоки полярной схемы, полученной из проективной геометрии $PG(2s, 2)$, является кодом, допускающим мажоритарное декодирование и имеет параметры кода Рида – Маллера $RM(s, 2s + 1)$, будучи не эквивалентным ему [7]. В работе [8] показано, что некоторые из этих кодов обладают исключительным свойством иметь то же самое весовое распределение, что и упомянутые коды Рида – Маллера.

Широкие классы двоичных нелинейных кодов с параметрами кодов Рида – Маллера были предложены рядом авторов (см. [9] и список литературы в работе [10], а также [11]). Среди них упомянем конструкции и исследование нетривиальных свойств $\mathbb{Z}_4$-линейных кодов Рида – Маллера (см. [10, 12–14]).

В настоящей статье приведено обобщение свитчинговой конструкции Пулатова [9] для кодов с параметрами классических двоичных кодов Рида – Маллера. Метод построения Пулатова является обобщением широко известной конструкции Васильева для совершенных кодов [15]. Следует отметить, что свитчинговый подход оказался плодотворным для решения многих проблем для совершенных q-ичных кодов, $q \geqslant 2$ (см. [16]).

В данной статье для предложенного нового класса кодов исследованы такие важные инварианты и свойства как весовой спектр и дистанционная инвариантность. Найдены условия, при которых код имеет тот же самый весовой спектр, что и код $RM(r, m)$, а также условия, при которых полученный код является дистанционно инвариантным. Доказано, что дистанционно инвариантные коды из предложенного класса кодов с тем же самым весовым спектром, что и у классического кода Рида – Маллера, но не эквивалентные ему, крайне редки. В частности показано, что таких кодов нет среди оригинальных кодов Пулатова.

§ 2. Конструкция

Основные определения и понятия см. в [1]. Всюду далее через d обозначено кодовое расстояние, а через $w(x)$ – вес вектора x. *Весовой спектр* кода C – это упорядоченный набор W_C, где $W_{C,i}$ равно числу кодовых слов кода C веса i. Двоичный код C называется *дистанционно инвариантным*, если выполняется $W_{C+x} = W_{C+y}$ для любых кодовых слов x и y из C. Вектор $y = (y_1, \ldots, y_n)$ называется *предшествующим* вектору $x = (x_1, \ldots, x_n)$, что записывается в виде $y \preceq x$, если $y_i \leqslant x_i$ для всякого $i = 1, \ldots, n$.

Далее потребуется следующий известный факт.

Утверждение 1. *Для любых векторов z, y справедливо $w(y+z \,|\, z) \geqslant w(y)$, где равенство достижимо в том и только том случае, когда $z \preceq y$.*

Напомним свитчинговую конструкцию Пулатова [9] для кодов с параметрами классических кодов Рида – Маллера.

Пусть e_i – двоичный вектор длины 2^m с 1 лишь в i-й координатной позиции. Пусть $\lambda\colon RM(r, m) \to \{0, 1\}$ – произвольная функция. Тогда множество

$$\big\{(x + y + e_1\lambda(y) \mid x + e_1\lambda(y)) :\ x \in RM(r + 1, m),\ y \in RM(r, m)\big\} \tag{2}$$

является расширенным двоичным кодом Пулатова, имеющим те же параметры, что и код Рида – Маллера $RM(r+1, m+1)$. Полагая функцию λ тождественно равной нулю, получим представление кода $RM(r+1, m+1)$ посредством конструкции Плоткина (1).

Рассмотрим следующее обобщение метода построения Пулатова. Обозначим через $\mathcal{T}$ совокупность представителей (лидеров) смежных классов кода $RM(r+1, m)$ в пространстве $\mathbb{F}^{2^m}$, взятых по одному вектору из каждого смежного класса. Пусть $\lambda\colon RM(r, m) \to \mathcal{T}$ – произвольная функция. Через $RM^{\lambda}(r+1, m+1)$ обозначим следующий код:

$$\big\{(x + y + \lambda(y) \,|\, x + \lambda(y)) : \, x \in RM(r+1, m), \, y \in RM(r, m)\big\}. \tag{3}$$

Для произвольного фиксированного кодового слова y из кода $RM(r, m)$ рассмотрим подкод

$$R_y^{\lambda} = \big\{(x + y + \lambda(y) \,|\, x + \lambda(y)) : \, x \in RM(r+1, m)\big\} \tag{4}$$

кода $RM^{\lambda}(r+1, m+1)$. Таким образом, последний представляет собой следующее объединение подкодов R_y^{λ}:

$$RM^{\lambda}(r+1, m+1) = \bigcup_{y \in RM(r,m)} R_y^{\lambda}. \tag{5}$$

Если λ – тождественно нулевая функция, то через R_y будем обозначать множество

$$\{(x + y \,|\, x) : \, x \in RM(r+1, m)\}.$$

Отсюда с учетом (1) имеем следующее представление классического кода Рида – Маллера:

$$RM(r+1, m+1) = \bigcup_{y \in RM(r,m)} R_y. \tag{6}$$

Для любого $y \in RM(r, m)$ минимальное расстояние подкода R_y^{λ} совпадает с минимальным расстоянием $d = 2^{m-r}$ кода $RM(r, m)$, так как по определению R_y^{λ} (см. (4)) его минимальное расстояние равно минимальному весу ненулевого вектора $(x \,|\, x)$, $x \in RM(r+1, m)$, т.е. $2 \times 2^{m-(r+1)} = 2^{m-r}$. Для различных y, y' из кода $RM(r, m)$ и для любых векторов x, $x' \in RM(r+1, m)$ согласно утверждению 1 справедливо

$$w\big(x + y + \lambda(y) + x' + y' + \lambda(y') \,|\, x + \lambda(y) + x' + \lambda(y')\big) \geqslant w(y + y'),$$

и в свою очередь, вес вектора $y + y'$ не меньше d. Отсюда минимальное расстояние кода $RM^{\lambda}(r+1, m+1)$ равно d. Следовательно, справедлива

Теорема 1. *Для любой функции $\lambda\colon RM(r, m) \to \mathcal{T}$ код $RM^{\lambda}(r+1, m+1)$ имеет ту же самую длину, мощность и минимальное расстояние, что и код Рида – Маллера $RM(r+1, m+1)$.*

Замечание 1. Заметим, что теорема 1 верна также в случае, когда вместо кодов $RM(r, m)$ и $RM(r+1, m)$ рассматриваются произвольные линейные коды C и D, а также для кодов над некоторыми другими метриками, в частности, над метрикой Ли. Полученные ниже весовые свойства этой конструкции кодов Рида – Маллера существенно опираются на результат Касами и Токуры [2] о несуществовании кодовых слов оригинального кода Рида – Маллера, имеющих веса между d и $3d/2$. По этой причине приводимые ниже результаты излагаются лишь для кодов Рида – Маллера.

Следствие 1. *Пусть λ и λ' – функции из $RM(r,m)$ в $\mathcal{T}$, такие что существует y из $RM(r,m)$, для которого $\lambda(y) \neq \lambda'(y)$. Тогда коды $RM^{\lambda}(r+1,m+1)$ и $RM^{\lambda'}(r+1,m+1)$ различны. В частности, имеется*

$$|RM(m-r-2,m)|^{|RM(r,m)|}$$

попарно различных кодов, полученных согласно конструкции (3).

Доказательство. Предположим, что $RM^{\lambda}(r+1,m+1) = RM^{\lambda'}(r+1,m+1)$. Для кодового слова y из $RM(r,m)$, удовлетворяющего условию следствия, и некоторого кодового слова $y' \in RM(r,m)$ рассмотрим кодовые слова

$$(x+y+\lambda(y)\,|\,x+\lambda(y)) \quad \text{и} \quad (x'+y'+\lambda'(y')\,|\,x'+\lambda'(y'))$$

кодов $RM^{\lambda}(r+1,m+1)$ и $RM^{\lambda'}(r+1,m+1)$ соответственно. Если эти векторы совпадают, то

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

Поскольку x и x' принадлежат коду $RM(r+1,m)$, из первого условия имеем

$$\lambda(y)+\lambda'(y') = x+x' \in RM(r+1,m).$$

Объединяя это со вторым равенством, получаем $y = y'$, и значит, $\lambda(y)+\lambda'(y) \in \in RM(r+1,m)$. Отсюда, так как функции λ и λ' принимают значения в $\mathcal{T}$, получаем, что $\lambda(y) = \lambda'(y)$, противоречие.

Так как код $RM(m-r-2,m)$ дуален коду $RM(r+1,m)$, то размер множества $\mathcal{T}$ равен

$$|\mathbb{F}^{2^m}/RM(r+1,m)| = |RM(m-r-2,m)|.$$

Отсюда с учетом числа различных функций λ, действующих из кода $RM(r,m)$ в множество $\mathcal{T}$, следует требуемая нижняя оценка числа кодов. ▲

§ 3. Основные результаты

В данном параграфе ограничимся случаем, когда функция λ такова, что $y \in \in RM(r,m)$ и $w(\lambda(y)) < d/4$. Всюду далее $d = 2^{m-r}$ – минимальное расстояние кодов $RM(r,m)$ и $RM^{\lambda}(r+1,m+1)$.

3.1. Нижняя граница весового спектра кода $RM^{\lambda}(r+1,m+1)$. В этом пункте коснемся весового спектра кода (5), в частности, исследуем число кодовых слов небольшого веса, меньшего $3d/2$.

Утверждение 2. *Пусть x – кодовое слово кода $RM(r+1,m)$, а y – кодовое слово кода $RM(r,m)$, такое что $w(y) > d$, где $d = 2^{m-r}$. Тогда для любого вектора $u \in \mathbb{F}^{2^m}$ выполняется*

$$w(y+u+x\,|\,u+x) \geqslant w(y) \geqslant 3d/2.$$

Доказательство. Неравенство $w(y+u+x\,|\,u+x) \geqslant w(y)$ справедливо согласно утверждению 1. Распределение небольших, т.е. близких к минимальному, весов в коде Рида – Маллера известно (см. [2]). В частности, вес, следующий за минимальным весом в коде $RM(r,m)$, равен $3d/2$. Значит, так как $w(y) > d$, то $w(y) \geqslant 3d/2$, и утверждение доказано. ▲

Рассматривая достаточно небольшие веса вектора $\lambda(y)$, из следующей леммы получим, что кодовые слова $(x+y+\lambda(y)\,|\,x+\lambda(y))$ кода $RM^{\lambda}(r+1,m+1)$ минимального веса могут возникнуть только вследствие минимальности веса вектора $(x+y\,|\,x)$, либо в случае, когда он нулевой.

Лемма 1. *Пусть* $\lambda\colon RM(r,m)\to\mathcal{T}$ *– произвольная функция, удовлетворяющая условиям* $\lambda(\mathbf{0})=\mathbf{0}$ *и* $w(\lambda(y))<d/4$ *для любого* $y\in RM(r,m)$, *где* $d=2^{m-r}$.

1. *Если* $y\in RM(r,m)$ *таков, что* $w(y)=d$ *и для некоторого* $x\in RM(r+1,m)$ *выполнено*

$$w(x+y+\lambda(y)\,|\,x+\lambda(y))=d,$$

то $w(x+y\,|\,x)=d$;

2. *Для любого* $y\in RM(r,m)$ *имеем* $W_{R_y^{\lambda},d}\leqslant W_{R_y,d}$. *Кроме того, справедливо*

$$W_{RM^{\lambda}(r+1,m+1),d}\leqslant W_{RM(r+1,m+1),d}.$$

Доказательство. 1. Предположим противное, т.е. $w(x+y\,|\,x)>d$. Заметим, что по утверждению 1 имеет место

$$w(x+y\,|\,x)\geqslant w(y)=d.$$

Тогда по теореме Касами и Токуры [2] кодовое слово $(x+y\,|\,x)$ кода $RM(r+1,m+1)$ имеет вес не менее $3d/2$. Тогда, поскольку $w(\lambda(y))<d/4$, то вес вектора $(x+y+\lambda(y)\,|\,x+\lambda(y))$ не может быть равен d, противоречие с условием.

2. Напомним, что

$$R_y^{\lambda}=\big\{(x+y+\lambda(y)\,|\,x+\lambda(y))\colon\ x\in RM(r+1,m)\big\}$$

и

$$R_y=\big\{(x+y\,|\,x)\colon\ x\in RM(r+1,m)\big\}.$$

Согласно (5) код $RM^{\lambda}(r+1,m+1)$ равен $\bigcup\limits_{y\in RM(r,m)}R_y^{\lambda}$, в то время как код Рида–Маллера $RM(r+1,m+1)$ равен $\bigcup\limits_{y\in RM(r,m)}R_y$ (см. (6)). Для каждого $y\in RM(r,m)$ сравним число векторов веса d в подкодах R_y^{λ} и R_y.

Если $y=\mathbf{0}$, то поскольку $\lambda(\mathbf{0})=\mathbf{0}$, имеем $R_{\mathbf{0}}^{\lambda}=R_{\mathbf{0}}$ и $W_{R_y^{\lambda},d}=W_{R_y,d}$.

Если $w(y)=d$, то из первого утверждения настоящей леммы следует, что существует инъективное отображение из множества векторов кода R_y^{λ}, имеющих вес d, в множество векторов из R_y веса d посредством сдвига на вектор $(\lambda(y)\,|\,\lambda(y))$. Следовательно, $W_{R_y^{\lambda},d}\leqslant W_{R_y,d}$.

Всякий вектор из R_y равен $(x+y+\lambda(y)\,|\,x+\lambda(y))$ для некоторого $x\in RM(r+1,m)$. Если $w(y)>d$, то полагая $u=\lambda(y)$ в утверждении 2, убеждаемся, что

$$w(x+y+\lambda(y)\,|\,x+\lambda(y))\geqslant 3d/2,$$

и векторов веса d в R_y^{λ} не существует. ▲

Далее рассмотрим случай, который позволяет получить коды с весовым спектром, отличным от спектра кода $RM(r+1,m+1)$.

Лемма 2. *Пусть* $\lambda\colon RM(r,m)\to\mathcal{T}$ *– произвольная функция, удовлетворяющая условиям* $\lambda(\mathbf{0})=\mathbf{0}$ *и* $w(\lambda(y))<d/4$ *для любого* $y\in RM(r,m)$, *где* $d=2^{m-r}$. *Если найдется кодовое слово* z *в* $RM(r,m)$, *такое что* $w(z)=d$ *и* $\lambda(z)\not\preceq z$, *то*

$$W_{R_z^{\lambda},d}=0\quad\text{и}\quad W_{RM^{\lambda}(r+1,m+1),d}<W_{RM(r+1,m+1),d}.$$

Доказательство. Предположим противное: пусть в $RM^{\lambda}(r+1,m+1)$ существует кодовое слово $(x+z+\lambda(z)\,|\,x+\lambda(z))$ веса d. Тогда по лемме 1, учитывая, что z – вектор веса d, вектор $(x+z\,|\,x)$ имеет вес d. Отсюда $x \preceq z$ согласно утверждению 1. Из утверждения 1, примененного к вектору $(x+z+\lambda(z)\,|\,x+\lambda(z))$, вытекает, что $x+\lambda(z) \preceq z$. Следовательно, $x \preceq z$, $x+\lambda(z) \preceq z$, а по условию $\lambda(z) \not\preceq z$. Так как одновременно все эти три свойства не выполняются, получаем противоречие. ▲

Следующее утверждение описывает "хороший случай" в том смысле, что весовые спектры кодов $RM^{\lambda}(r+1,m+1)$ и $RM(r+1,m+1)$ совпадают.

Утверждение 3. *Справедливо следующее.*

1. *Пусть* $\lambda\colon RM(r,m) \to \mathcal{T}$ *– произвольная функция, удовлетворяющая условию* $\lambda(y) \preceq y$ *для любого* $y \in RM(r,m)$. *Тогда*

$$W_{RM^{\lambda}(r+1,m+1)} = W_{RM(r+1,m+1)}.$$

2. *Для всякого* $0 \leqslant r < m$ *существует не менее*

$$\sum_{i=1}^{\frac{|RM(r,m)|}{2}-1} C^{i}_{|RM(r,m)|} \sum_{j=1}^{2^{m-r-2}-1} C^{j}_{2^{m-r}}$$

различных кодов $RM^{\lambda}(r+1,m+1)$ *с тем же весовым распределением, что и у кода* $RM(r+1,m+1)$, *и не эквивалентных ему.*

Доказательство. 1. Без ограничения общности предположим, что векторы y, x и $\lambda(y)$ имеют вид

$$y = (1,\dots,1\,|\,0,\dots,0),$$
$$x = (x_1,\dots,x_s\,|\,x_{s+1},\dots,x_n),$$
$$\lambda(y) = (1,\dots,1,0,\dots,0\,|\,0,\dots,0),$$

где $w(y)=s$ и $w(\lambda(y))=\ell<s$.

Рассмотрим векторы $(x+y+\lambda(y)\,|\,x+\lambda(y))$ и $(x+y\,|\,x)$. Очевидно, что конкатенация векторов

$$(y+x+\lambda(y)) = (x_1,\dots,x_\ell,x_{\ell+1}+1,\dots,x_s+1\,|\,x_{s+1},\dots,x_n),$$
$$(x+\lambda(y)) = (x_1+1,\dots,x_\ell+1,x_{\ell+1},\dots,x_s\,|\,x_{s+1},\dots,x_n)$$

может быть получена из конкатенации векторов

$$y+x = (x_1+1,\dots,x_s+1\,|\,x_{s+1},\dots,x_n),$$
$$x = (x_1,\dots,x_s\,|\,x_{s+1},\dots,x_n)$$

посредством некоторой подстановки. Следовательно, вектор $(x+y+\lambda(y)\,|\,x+\lambda(y))$ получается перестановкой из вектора $(x+y\,|\,x)$, и

$$W_{RM^{\lambda}(r+1,m+1)} = W_{RM(r+1,m+1)}.$$

2. Рассмотрим не тождественно нулевые функции λ, принимающие нулевые значения на хотя бы $\frac{|RM(r,m)|}{2}+1$ векторах из $RM(r,m)$, один из которых нулевой, и для всех $y \in RM(r,m)$ выполняется $\lambda(y) \preceq y$, где $w(\lambda(y)) < d/4$. Так как для всякой такой функции λ код $RM^{\lambda}(r+1,m+1)$ содержит нулевой вектор и имеет больше половины общих кодовых слов с линейным кодом $RM(r+1,m+1)$, то код $RM^{\lambda}(r+1,m+1)$ не линеен, и следовательно, не эквивалентен коду $RM(r+1,m+1)$. Согласно первому пункту данного утверждения всякий такой код имеет то же весовое распределение, что и $RM(r+1,m+1)$. Значения рассматриваемой функции λ

всегда принадлежат шару радиуса $d/4-1$ и, в свою очередь, множеству лидеров классов смежности линейного кода $RM(r+1,m)$ в $\mathbb{F}^{2^m}$. В силу утверждения 1 получаем, что все такие коды попарно различны. Из того, что каждое кодовое слово кода $RM(r,m)$ имеет вес не менее $d=2^{m-r}$, число возможностей выбрать ненулевой вектор $\lambda(y)$, $\lambda(y)\preceq y$, в шаре радиуса $d/4-1$ с центром в кодовом слове y веса не менее 2^{m-r} равно

$$\sum_{j=1}^{2^{m-r-2}-1} C_{2^{m-r}}^{j}.$$

Отсюда, предварительно выбирая множество кодовых слов из кода $RM(r,m)$, на которых значения функций λ ненулевые, получаем требуемую нижнюю оценку числа таких кодов. ▲

Замечание 2. Случай $r=m-2$ является исключительным. Для всякой функции λ код $RM^{\lambda}(m-2,m)$ является расширенным совершенным кодом, и следовательно, дистанционно инвариантным.

3.2. Дистанционная инвариантность, случай двузначных функций. В этом пункте рассмотрим двузначные функции λ, полагая ненулевое значение функции вектором, имеющим относительно небольшой вес. Покажем, что при этом ограничении на функцию λ не существует дистанционно инвариантных кодов, полученных конструкцией (3) и имеющих то же весовое распределение, что и коды Рида – Маллера, но отличных от них. Как следствие, будет построено большое количество линейных кодов с весовым распределением, отличным от весового распределения кода Рида – Маллера.

Напомним, что базис линейного кода, состоящий из кодовых слов минимального веса, называется *базисом минимального веса*.

Лемма 3. *Для любых* r,m, *таких что* $0\leqslant r\leqslant m$, $1\leqslant m$ *и* $i\in\{1,\dots,2^m\}$, *код*

$$\{u:\ u\in RM(r,m),\ u_i=0\}$$

имеет базис минимального веса.

Доказательство. Доказательство проведем по индукции. Легко убедиться, что для малых m лемма верна.

Пусть для всех $r\leqslant m-1$ и любых i, не превосходящих 2^{m-1}, коды

$$\{x:\ x\in RM(r,m-1),\ x_i=0\}$$

имеют базисы минимального веса.

Согласно конструкции Плоткина (1) выполняется

$$RM(r,m)=\big\{(x+y\,|\,x):\ x\in RM(r,m-1),\ y\in RM(r-1,m-1)\big\}.$$

Из данной конструкции, подставляя $x+y$ вместо x, без ограничения общности можно считать, что $i\geqslant 2^{m-1}+1$. Убедимся, что код

$$\big\{(x+y\,|\,x):\ x\in RM(r,m-1),\ x_i=0,\ y\in RM(r-1,m-1)\big\} \tag{7}$$

имеет базис минимального веса. Напомним, что минимальное расстояние кода $RM(r-1,m-1)$ равно 2^{m-r} и что он имеет базис минимального веса. По предположению индукции код

$$\big\{(x\,|\,x):\ x\in RM(r,m-1),\ x_i=0\big\}$$

также имеет минимальное расстояние 2^{m-r} и обладает базисом минимального веса. Следовательно, код (7) тоже обладает базисом минимального веса. ▲

Утверждение 4. *Пусть для любого $i \in \{1,\dots,2^m\}$ функция λ такова, что $\lambda(y) = y_i e_i$ для каждого $y \in RM(r,m)$. Тогда код $RM^\lambda(r+1,m+1)$ совпадает с кодом $RM(r+1,m+1)$ с точностью до перестановки координатных позиций.*

Доказательство. Зафиксируем произвольный элемент $i \in \{1,\dots,2^m\}$. Так как $\lambda(y) = y_i e_i$ для каждого кодового слова y кода $RM(r,m)$, то $\lambda(y) \preceq y$. Докажем, что транспозиция $\pi = (i, 2^m + i)$ позволяет получить

$$\pi(RM^\lambda(r+1,m+1)) = RM(r+1,m+1).$$

Пусть y – произвольное кодовое слово в $RM(r,m)$, такое что $y_i = 0$. Тогда $\lambda(y) = \mathbf{0}$, и в силу того, что

$$\pi(x+y\,|\,x) = (x+y\,|\,x) \in RM(r+1,m+1),$$

утверждение справедливо.

Если y – кодовое слово кода $RM(r,m)$, такое что $y_i = 1$, то $\lambda(y) = e_i$. Следовательно,

$$\begin{aligned}&\pi(x+y+\lambda(y)\,|\,x+\lambda(y)) = \pi(x+y+e_i\,|\,x+e_i) = \\ &= \pi(x+e_i\,|\,x+e_i) + \pi(y\,|\,\mathbf{0}) = (x+e_i\,|\,x+e_i) + (y+e_i\,|\,e_i) = \\ &= (x+y\,|\,x) \in RM(r+1,m+1). \quad ▲\end{aligned}$$

Теорема 2. *Пусть r, m таковы, что $0 \leqslant r \leqslant m$ и $1 \leqslant m$, $d = 2^{m-r}$, и пусть u – любой вектор из $\mathbb{F}^{2^m}$, такой что $w(u) < d/4$. Пусть λ – произвольная функция из кода $RM(r,m)$ в множество $\{\mathbf{0}, u\}$, где $\lambda(\mathbf{0}) = \mathbf{0}$. Тогда код $RM^\lambda(r+1,m+1)$ является дистанционно инвариантным, и*

$$W_{RM^\lambda(r+1,m+1)} = W_{RM(r+1,m+1)}$$

в том и только том случае, когда с точностью до перестановки координатных позиций он является кодом Рида – Маллера $RM(r+1,m+1)$.

Доказательство. Достаточность очевидна.

Докажем необходимость. Пусть i таково, что $u_i = 1$. Код $RM(r,m)$ равен объединению линейного подкода

$$\{y :\ y \in RM(r,m),\ y_i = 0\}$$

и его смежного класса

$$\{y :\ y \in RM(r,m),\ y_i = 1\}.$$

Обозначим эти подкоды через $RM_0(r,m)$ и $RM_1(r,m)$ соответственно. Рассмотрим значения функции λ на подкоде $RM_0(r,m)$.

Пусть код $RM^\lambda(r+1,m+1)$ дистанционно инвариантен и имеет тот же весовой спектр, что и у кода $RM(r+1,m+1)$. В первую очередь покажем, что λ – тождественно нулевая функция на $RM_0(r,m)$. Предположим противное. Поскольку $\lambda(\mathbf{0}) = \mathbf{0}$ и λ принимает значение u на некотором кодовом слове из $RM_0(r,m)$, то по лемме 3 найдется последовательность кодовых слов $y^1,\dots,y^t$ в коде $RM_0(r,m)$, удовлетворяющих условиям $y^1 = \mathbf{0}$, $\lambda(\mathbf{0}) = \mathbf{0}$, $\lambda(y^t) = u$, таких что для любого $j \in \{1,\dots,t-1\}$ выполняется $d(y^j, y^{j+1}) = d$. Следовательно, в этой последовательности найдутся два вектора из кода $RM_0(r,m)$ (обозначим их через $\widetilde{y}$ и $\overline{y}$) на расстоянии d друг от друга, удовлетворяющие условиям $\lambda(\widetilde{y}) = \mathbf{0}$, $\lambda(\overline{y}) = u$.

Так как $\lambda(\widetilde{y}) = \mathbf{0}$, то вектор $(\widetilde{y}\,|\,\mathbf{0})$ является кодовым словом кода $RM^{\lambda}(r+1, m+1)$. Покажем, что

$$W_{RM^{\lambda}(r+1,m+1)+(\widetilde{y}|\mathbf{0}),d} < W_{RM(r+1,m+1),d}.$$

Для любого $y \in RM(r,m)$ рассмотрим функцию λ', такую что

$$\lambda'(y) = \lambda(y + \widetilde{y}).$$

Согласно определению функции λ' нетрудно видеть, что

$$RM^{\lambda}(r+1, m+1) + (\widetilde{y}\,|\,\mathbf{0}) = RM^{\lambda'}(r+1, m+1).$$

По выбору векторы $\overline{y}$ и $\widetilde{y}$ в коде $RM_0(r,m)$ находятся на расстоянии d друг от друга, а вектор $\overline{y} + \widetilde{y}$ имеет вес d и принадлежит коду $RM_0(r,m)$. Более того, по определению функции λ' имеем

$$\lambda'(\overline{y} + \widetilde{y}) = \lambda(\overline{y} + \widetilde{y} + \widetilde{y}) = \lambda(\overline{y}) = u.$$

Заметим, что i-я координатная позиция кодовых слов кода $RM_0(r,m)$ нулевая, в то время как i-я координатная позиция вектора u равна 1. Отсюда $u \not\leq \overline{y} + \widetilde{y}$.

Применяя лемму 2 к функции λ' и вектору $z = \overline{y} + \widetilde{y}$ веса d, получим

$$W_{RM^{\lambda'}(r+1,m+1),d} = W_{RM^{\lambda}(r+1,m+1)+(\widetilde{y}|\mathbf{0}),d} < W_{RM(r+1,m+1),d}.$$

Следовательно, в случае, когда функция λ имеет ненулевые значения на коде $RM_0(r,m)$, код $RM^{\lambda}(r+1, m+1)$ не может быть дистанционно инвариантным и одновременно иметь весовой спектр, как у кода Рида – Маллера.

Теперь покажем, что λ принимает одинаковые значения на подкоде $RM_1(r,m)$. Предположим противное. Тогда, аналогично вышеприведенным рассуждениям, найдутся два вектора $\widetilde{y}$ и $\overline{y}$ из подкода $RM_1(r,m)$ на расстоянии d друг от друга, такие что $\lambda(\widetilde{y}) = 0$ и $\lambda(\overline{y}) = u$. Введем функцию λ', такую что $\lambda'(y) = \lambda(y + \widetilde{y})$. Заметим, что

$$RM^{\lambda}(r+1, m+1) + (\widetilde{y}\,|\,\mathbf{0}) = RM^{\lambda'}(r+1, m+1).$$

Так как вектор $z = \overline{y} + \widetilde{y}$ имеет вес d и его i-я позиция равна 0, то $\lambda(z) = u \not\leq \overline{y} + \widetilde{y}$, и по лемме 2 получаем, что

$$W_{RM^{\lambda'}(r+1,m+1),d} = W_{RM^{\lambda}(r+1,m+1)+(\widetilde{y}|\mathbf{0}),d} < W_{RM(r+1,m+1),d},$$

противоречие.

Таким образом, функция λ на $RM_1(r,m)$ либо тождественно нулевая, либо является константой, равной u. В последнем случае функция λ тождественно нулевая только на коде $\{y \in RM(r,m),\ y_i = 0\}$. Если предположить, что вес вектора u больше 1, то повторяя доказательство, приведенное выше, для любого i', $u_{i'} = 1$, $i' \neq i$, получим, что λ принимает нулевое значение только на $\{y \in RM(r,m),\ y_{i'} = 0\}$, противоречие.

Следовательно, если код $RM^{\lambda}(r+1, m+1)$ дистанционно инвариантен и имеет весовой спектр, как у кода Рида – Маллера $RM(r+1, m+1)$, то либо λ – тождественно нулевая функция, либо найдется $i \in \{1, \ldots, 2^m\}$, такое что $\lambda(y) = y_i e_i$ для произвольного $y \in RM(r,m)$. В первом случае код $RM^{\lambda}(r+1, m+1)$ совпадает с кодом $RM(r+1, m+1)$, а во втором случае согласно утверждению 4 код $RM^{\lambda}(r+1, m+1)$ эквивалентен коду $RM(r+1, m+1)$. ▲

Функцию $\lambda\colon RM(r,m) \to \mathcal{T}$ назовем *линейной*, если для любых y, $y' \in RM(r,m)$ выполнено

$$\lambda(y+y') + \lambda(y') + \lambda(y) \in RM(r+1,m).$$

Несложно видеть, что для рассматриваемых кодов имеет место следующее

Утверждение 5. *Пусть* $\lambda\colon RM(r,m) \to \mathcal{T}$ *– произвольная линейная функция. Тогда код* $RM^\lambda(r+1,m+1)$ *линеен.*

Следствие 2. *При* $r \leqslant m$ *всякий дистанционно инвариантный код с параметрами кода Рида – Маллера* $RM(r,m)$*, получаемый конструкцией Пулатова* (2) *и имеющий то же весовое распределение, что и код Рида – Маллера* $RM(r,m)$*, совпадает с ним с точностью до перестановки.*

Доказательство. Заметим, что при $r \leqslant m-3$ конструкция Пулатова является частным случаем рассматриваемой конструкции при функции λ, принимающей значения в множество, состоящее из двух фиксированных векторов веса 0 и 1. Отсюда в силу теоремы 2 получаем требуемое. При $r \geqslant m-2$ утверждение также выполнено в силу дистанционной инвариантности любого кода с параметрами расширенного кода Хэмминга или кода, состоящего из всех векторов четного веса. ▲

Широкий класс дистанционно инвариантных кодов можно получить, используя линейные функции в предложенной выше конструкции. Однако в случае, когда функция принимает только два значения достаточно малого веса, весовой спектр полученных кодов не совпадает с таковым для классического кода Рида – Маллера, либо приводит к коду Рида – Маллера с точностью до перестановки координатных позиций (см. теорему 2).

Следствие 3. *При* $r \leqslant m-3$ *существует по крайней мере*

$$(|RM(r,m)|-1)\left(-1+\sum_{i=0}^{d/4-1}\binom{2^m}{i}\right)-2^m+1$$

попарно различных линейных кодов $RM^\lambda(r+1,m+1)$ *с параметрами кодов Рида – Маллера* $RM(r+1,m+1)$*, таких что*

$$W_{RM^\lambda(r+1,m+1),d} < W_{RM(r+1,m+1),d}.$$

Доказательство. Рассмотрим произвольную не тождественно нулевую линейную функцию

$$\lambda\colon RM(r,m) \to \{\mathbf{0}, u\}, \quad w(u) < d/4.$$

Число выборов вектора u равно

$$-1+\sum_{i=0}^{d/4-1}\binom{2^m}{i}.$$

Отсюда получаем, что количество таких функций равно этому числу способов выбрать вектор u, умноженному на число способов задать значение функции λ, равному вектору u на непустом множестве базисных векторов кода $RM(r,m)$. Из полученных функций мы исключаем 2^m-1 функций вида $\lambda(y) = y_i e_i$, $i \in \{1,\ldots,2^m\}$, так как в силу утверждения 5 коды, соответствующие им, эквивалентны оригинальным кодам Рида – Маллера. В силу доказательства теоремы 2 все остальные коды $RM^\lambda(r+1,m+1)$ удовлетворяют неравенству

$$W_{RM^\lambda(r+1,m+1),d} < W_{RM(r+1,m+1),d}. \quad ▲$$

Следующий результат был получен с помощью компьютера на основе вышеизложенных ограничений на весовое распределение кодов $RM^{\lambda}(2,5)$.

Утверждение 6. *Всякий линейный код* $RM^{\lambda}(2,5)$ *либо эквивалентен коду* $RM(2,5)$, *либо удовлетворяет условию*

$$W_{RM^{\lambda}(2,5),8} < W_{RM(2,5),8}.$$

§ 4. Заключение

В статье предложена новая конструкция двоичных кодов с параметрами кодов Рида – Маллера. Из утверждения 3 вытекает существование богатого множества кодов с теми же самыми параметрами и весовым спектром, что и у кодов Рида – Маллера. Результаты § 3 позволяют сделать вывод, что достаточно трудно обнаружить дистанционно инвариантные коды с числом кодовых слов минимального веса, как у кода Рида – Маллера, но не эквивалентных ему.

Авторы выражают благодарность С.В. Августиновичу и В.Н. Потапову, обратившим внимание авторов на задачу описания весового спектра кодов с параметрами кодов Рида – Маллера.

СПИСОК ЛИТЕРАТУРЫ

1. *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки. М.: Связь, 1979.
2. *Kasami T., Tokura N.* On the Weight Structure of Reed–Muller Codes // IEEE Trans. Inform. Theory. 1970. V. 16. № 6. P. 752–759. https://doi.org/10.1109/TIT.1970.1054545
3. *Abbe E., Shpilka A., Ye M.* Reed–Muller Codes: Theory and Algorithms // IEEE Trans. Inform. Theory. 2021. V. 67. № 6. Part 1. P. 3251–3277. https://doi.org/10.1109/TIT.2020.3004749
4. *Kaufman T., Lovett S., Porat E.* Weight Distribution and List-Decoding Size of Reed–Muller Codes // IEEE Trans. Inform. Theory. 2012. V. 58. № 5. P. 2689–2696. https://doi.org/10.1109/TIT.2012.2184841
5. *Jungnickel D., Tonchev V.D.* Polarities, Quasi-symmetric Designs, and Hamada's Conjecture // Des. Codes Cryptogr. 2009. V. 51. № 2. P. 131–140. https://doi.org/10.1007/s10623-008-9249-8
6. *Hamada N.* On the p-Rank of the Incidence Matrix of a Balanced or Partially Balanced Incomplete Block Design and Its Application to Error-Correcting Codes // Hiroshima Math. J. 1973. V. 3. № 1. P. 153–226. https://doi.org/10.32917/hmj/1206137446
7. *Clark D., Tonchev V.D.* A New Class of Majority-Logic Decodable Codes Derived from Polarity Designs // Adv. Math. Commun. 2013. V. 7. № 2. P. 175–186. https://doi.org/10.3934/amc.2013.7.175
8. *Harada M., Novak E., Tonchev V.* The Weight Distribution of the Self-dual [128, 64] Polarity Design Code // Adv. Math. Commun. 2016. V. 10. № 3. P. 643–648. https://doi.org/10.3934/amc.2016032
9. *Пулатов А.К.* Нижняя оценка сложности схемной реализации для одного класса кодов // Дискретный анализ. Вып. 25. Новосибирск: Ин-т матем. СО АН СССР, 1974. С. 56–61.
10. *Соловьева Ф.И.* О пересечении кодов типа Рида – Маллера // Пробл. передачи информ. 2021. Т. 57. № 4. С. 63–73. https://doi.org/10.31857/S0555292321040057
11. *Соловьева Ф.И.* О построении кодов типа Рида – Маллера и исследовании их свойств // Тр. МФТИ. 2022. Т. 14. № 2. С. 110–123. https://www.elibrary.ru/ygvlwy
12. *Соловьева Ф.И.* О $\mathbb{Z}_4$-линейных кодах с параметрами кодов Рида – Маллера // Пробл. передачи информ. 2007. Т. 43. № 1. С. 32–38. http://mi.mathnet.ru/ppi4
13. *Pujol J., Rifà J., Solov'eva F.I.* Construction of $\mathbb{Z}_4$-Linear Reed–Muller Codes // IEEE Trans. Inform. Theory. 2009. V. 55. № 1. P. 99–104. https://doi.org/10.1109/TIT.2008.2008143

14. *Solov'eva F.I.* Minimum Weight Bases for Quaternary Reed–Muller Codes // Сиб. электрон. матем. изв. 2021. Т. 18. № 2. С. 1358–1366. https://doi.org/10.33048/semi.2021.18.103
15. *Васильев Ю.Л.* О негрупповых плотно упакованных кодах // Проблемы кибернетики. Т. 8. М.: Физматлит, 1962. С. 337–339.
16. *Solov'eva F.I.* Switchings and Perfect Codes // Numbers, Information and Complexity. Boston: Springer, 2000. P. 311–324. https://doi.org/10.1007/978-1-4757-6048-4_25

Могильных Иван Юрьевич
Институт математики им. С.Л. Соболева
СО РАН, Новосибирск
ivmog@math.nsc.ru
Соловьева Фаина Ивановна
(15.08.1952 – 09.08.2022)

Поступила в редакцию
01.04.2022
После доработки
16.06.2022
Принята к публикации
18.06.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391 : 519.72

© 2022 г. **И.В. Воробьев**[1], **В.С. Лебедев**[2]

УЛУЧШЕНИЕ ВЕРХНИХ ГРАНИЦ СКОРОСТЕЙ РАЗДЕЛЯЮЩИХ И ПОЛНОСТЬЮ РАЗДЕЛЯЮЩИХ КОДОВ

Двоичный код называется (s, ℓ)-разделяющим кодом, если для любых двух непересекающихся наборов его слов мощности не более s и ℓ соответственно существует координата, в которой все слова из одного набора имеют символ 0, а все слова из другого набора имеют символ 1. Если же вдобавок для любых наборов существует вторая координата, в которой у первого набора во всех словах стоят 1, а у второго стоят 0, то такой код называется (s, ℓ)-полностью разделяющим кодом. В статье улучшаются верхние границы скоростей разделяющих и полностью разделяющих кодов.

Ключевые слова: разделяющие коды, полностью разделяющие коды, асимптотическая скорость, граница Плоткина.

DOI: 10.31857/S0555292322030044, **EDN:** EADNOC

§ 1. Введение

Впервые задача построения двоичных разделяющих систем возникла при исследовании асинхронных конечных автоматов. Для борьбы с критическими состояниями элементов памяти автомата при его переходе из одного устойчивого внутреннего состояния в другое Ю.Л. Сагаловичем было предложено использовать двоичные $(2, 2)$-разделяющие коды [1]. В работе [2] было введено общее определение (s, ℓ)-разделяющих кодов. Отметим, что хотя изначально исследование разделяющих кодов было мотивировано задачами из теории автоматов, позднее они нашли применение при разработке способов защиты информации от нелегального копирования [3] и при построении хэш-функций [4].

Первая нижняя оценка скорости разделяющих кодов была получена с помощью случайного кодирования в работе [2]. В [5] с помощью случайного кодирования с выбрасыванием были получены нижние оценки скоростей $(2, 2)$- и $(2, 1)$-разделяющих и полностью разделяющих кодов, улучшающие оценки из [2] и совпадающие с оценкой скорости линейных $(2, 2)$-разделяющих кодов из [6]. Отметим, что доказательство с выбрасыванием точно так же работает и для случая (s, ℓ)-разделяющих и полностью разделяющих кодов.

Также отметим неожиданное улучшение этих границ для $(2, 1)$-разделяющих кодов [7].

Верхние границы для скорости $(2, 2)$-разделяющих кодов впервые были получены Сагаловичем в [8] с помощью следующей идеи. Возьмем два кодовых слова на минимальном расстоянии (Хэмминга) d и ограничим исходный код на соответствующие d

[1] Исследование выполнено за счет гранта Российского научного фонда (номер проекта 22-41-02028).

[2] Работа выполнена при финансовой поддержке совместного проекта Российского фонда фундаментальных исследований и Национального научного фонда Болгарии (номер проекта 20-51-18002).

координат, предварительно удалив из кода эти два слова. Новый код длины d будет состоять из различных двоичных слов, и следовательно, мощность исходного кода не превосходит $2^d + 2$. Это позволяет оценить сверху скорость $(2,2)$-разделяющих кодов с помощью известных верхних оценок скорости кода (см. [9, 10]).

Идея, что этот подход можно обобщать на случай (s,ℓ)-кодов, высказывалась самим Ю.Л. Сагаловичем, а также Л.А. Бассалыго и Г.А. Кабатянским на семинарах ИППИ РАН по теории кодирования. Этот подход был реализован для хэш-кодов в [11], а для (s,ℓ)-свободных от перекрытий кодов, которые тесно связаны с разделяющими кодами, – в [12, 13].

В [14] были предложены рекуррентные верхние границы скоростей (s,ℓ)-разделяющих и полностью разделяющих кодов, выражающиеся через скорости кодов с параметрами $(s-1,\ell-1)$. Отметим, что в [14] (s,ℓ)-разделяющие коды сводились к $(s-1,\ell-1)$-полностью разделяющим кодам, что позволило получить более сильные оценки, чем если бы коды сводились к разделяющим.

В [15] получены рекуррентные неравенства, связывающие верхние границы для (s,ℓ)-разделяющих и полностью разделяющих кодов со скоростями кодов с параметрами $(s-u,\ell-v)$. Эти границы, в частности, позволили показать, что скорость t-IPP-кодов [16] экспоненциально мала по t (см. [17]).

В данной статье мы доказываем новые рекуррентные неравенства, которые связывают скорости разделяющих кодов и кодов, свободных от перекрытий. Эти неравенства позволяют улучшить верхние границы скоростей разделяющих кодов для многих параметров s и ℓ. Кроме того, мы улучшаем известную верхнюю границу для $(2,1)$-полностью разделяющих кодов. Так как наилучшие верхние границы выражаются через верхние границы скоростей кодов с меньшими параметрами, улучшение границы для $(2,1)$-полностью разделяющих кодов приводит к улучшению для широкого набора параметров.

§ 2. Определения и обозначения

Пусть N, M, s, ℓ – натуральные числа, символ $\triangleq$ обозначает равенство по определению, $|A|$ – мощность множества A, $[N] \triangleq \{1, 2, \ldots, N\}$ – множество целых чисел от 1 до N. Произвольное подмножество булева куба $\{0,\ 1\}^N$ называется двоичным кодом длины N. Будем обозначать двоичную энтропию через

$$h(x) = -x\log_2(x) - (1-x)\log_2(1-x).$$

Определение 1 [2]. Двоичный код $\mathcal{C}$ называется (s,ℓ)-*разделяющим*, если для любых двух непересекающихся множеств кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существует координата i, такая что:

либо $\quad x_i = 0$ для любого $\boldsymbol{x} \in \mathcal{S}$ $\quad$ и $\quad y_i = 1$ для любого $\boldsymbol{y} \in \mathcal{L}$,
либо $\quad x_i = 1$ для любого $\boldsymbol{x} \in \mathcal{S}$ $\quad$ и $\quad y_i = 0$ для любого $\boldsymbol{y} \in \mathcal{L}$.

Определение 2. Двоичный код $\mathcal{C}$ называется (s,ℓ)-*свободным от перекрытий*, если для любых двух непересекающихся множеств кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существует координата i, для которой

$x_i = 0$ для любого $\boldsymbol{x} \in \mathcal{S}$ $\quad$ и $\quad y_i = 1$ для любого $\boldsymbol{y} \in \mathcal{L}$.

Определение 3. Двоичный код $\mathcal{C}$ называется (s,ℓ)-*полностью разделяющим*, если для любых двух непересекающихся множеств кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существуют две координаты i и j, такие что:

$x_i = 0$ для любого $\boldsymbol{x} \in \mathcal{S}$ $\quad$ и $\quad y_i = 1$ для любого $\boldsymbol{y} \in \mathcal{L}$, $\quad$ и
$x_j = 1$ для любого $\boldsymbol{x} \in \mathcal{S}$ $\quad$ и $\quad y_j = 0$ для любого $\boldsymbol{y} \in \mathcal{L}$.

Отметим, что

определение 3 $\implies$ определение 2 $\implies$ определение 1,

а также что при $s = \ell$ определения (s, ℓ)-полностью разделяющих и (s, ℓ)-свободных от перекрытий кодов совпадают. Кроме того, напомним, что $(s, 1)$-свободные от перекрытий коды известны как s-дизъюнктивные коды [18]. Полезно заметить, что если код $\mathcal{C}$ разделяющий, то и код $\mathcal{C} + \boldsymbol{a}$ тоже разделяющий для любого двоичного вектора $\boldsymbol{a}$. Если же код $\mathcal{C}$ полностью разделяющий, то и код $\mathcal{C} + \mathbf{1}$ обладает тем же свойством.

Принято думать, что (s, ℓ)-свободные от перекрытий коды были впервые определены в работе [19] в 1988 г., а полностью разделяющие системы – в работе [20] в 1973 г. На самом деле, в [20] были определены свободные от перекрытий коды (за 15 лет до работы [19]), но они были при этом названы полностью разделяющими.

Обозначим через $N_{\rm s}(M, s, \ell)$, $N_{\rm cf}(M, s, \ell)$ и $N_{\rm cs}(M, s, \ell)$ минимальную длину кодов из определений 1–3 при заданной мощности M. Определим асимптотические скорости кодов

$$R_{\rm s}(s, \ell) \triangleq \varlimsup_{M\to\infty} \frac{\log_2 M}{N_{\rm s}(M, s, \ell)}, \tag{1}$$

$$R_{\rm cf}(s, \ell) \triangleq \varlimsup_{M\to\infty} \frac{\log_2 M}{N_{\rm cf}(M, s, \ell)}, \tag{2}$$

$$R_{\rm cs}(s, \ell) \triangleq \varlimsup_{M\to\infty} \frac{\log_2 M}{N_{\rm cs}(M, s, \ell)}. \tag{3}$$

В этой статье мы улучшаем верхние границы скоростей $R_{\rm s}$ и $R_{\rm cs}$. Заметим, что в силу симметрии определений 1 и 3 справедливы равенства $R_{\rm s}(s, \ell) = R_{\rm s}(\ell, s)$ и $R_{\rm cs}(s, \ell) = R_{\rm cs}(\ell, s)$.

§ 3. Известные результаты

Верхние границы скоростей разделяющих, полностью разделяющих и свободных от перекрытий кодов получаются из различных рекуррентных неравенств, связывающих скорости кодов для разных значений параметров s и ℓ.

Для $(s, 1)$-свободных от перекрытий кодов, которые также называются s-дизъюнктивными кодами, известна [21] верхняя граница

$$R_{\rm cf}(s, 1) \leqslant \overline{R}_{\rm cf}(s, 1), \tag{4}$$

где последовательность $\overline{R}_{\rm cf}(s, 1)$ определена следующим образом: $\overline{R}_{\rm cf}(1, 1) \triangleq 1$, $\overline{R}_{\rm cf}(2, 1) \triangleq \max\limits_{0<v<1} f_2(v)$, а $\overline{R}_{\rm cf}(s, 1)$ при $s > 2$ является единственным решением уравнения

$$\overline{R}_{\rm cf}(s, 1) = f_{\rm s}\left(1 - \frac{\overline{R}_{\rm cf}(s, 1)}{\overline{R}_{\rm cf}(s-1, 1)}\right),$$

где $f_{\rm s}(v) \triangleq h(v/s) - vh(1/s)$.

В [22] было доказано, что

$$R_{\rm cf}(s, \ell) \leqslant \left(\frac{1}{R_{\rm cf}(s, \ell-1)} + \frac{1}{R_{\rm cf}(s-1, \ell)}\right)^{-1}. \tag{5}$$

В [12] (см. также [23]) было доказано рекуррентное неравенство

$$R_{\mathrm{cf}}(s,\ell) \leqslant R_{\mathrm{cf}}(s-u,\ell-v)\frac{u^u v^v}{(u+v)^{u+v}}, \quad 1 \leqslant u \leqslant s-1, \quad 1 \leqslant v \leqslant \ell-1. \tag{6}$$

В [24] (см. также [13]) были доказаны более сильные неравенства

$$R_{\mathrm{cf}}(s,\ell) \leqslant \frac{R_{\mathrm{cf}}(s-u,\ell-v)}{R_{\mathrm{cf}}(s-u,\ell-v)+\dfrac{(u+v)^{u+v}}{u^u v^v}}, \tag{7}$$

$$R_{\mathrm{cf}}(s,\ell) \leqslant h\left(1/2-\sqrt{\frac{2R_{\mathrm{cf}}(s,\ell)}{R_{\mathrm{cf}}(s-1,\ell-1)}\left(1-\frac{2R_{\mathrm{cf}}(s,\ell)}{R_{\mathrm{cf}}(s-1,\ell-1)}\right)}\right). \tag{8}$$

В [14, 25] было доказано, что скорость $(s,1)$-разделяющих кодов удовлетворяет неравенству

$$R_{\mathrm{s}}(s,1) \leqslant \frac{1}{s}. \tag{9}$$

Для разделяющих и полностью разделяющих кодов в [14] были получены следующие результаты:

$$R_{\mathrm{s}}(s,\ell) \leqslant \widehat{R}\left(\frac{R_{\mathrm{s}}(s,\ell)}{R_{\mathrm{cs}}(s-1,\ell-1)}\right), \tag{10}$$

$$R_{\mathrm{cs}}(s,\ell) \leqslant \widehat{R}\left(\frac{2R_{\mathrm{cs}}(s,\ell)}{R_{\mathrm{cs}}(s-1,\ell-1)}\right), \tag{11}$$

где $\widehat{R}(\tau)$– произвольная верхняя асимптотическая оценка скорости кода с относительным расстоянием Хэмминга $\tau = d/n$.

В [15] для разделяющих кодов были доказаны рекуррентные неравенства, аналогичные неравенствам (6). А именно,

1. Для любых $u \in [s-1]$, $v \in [\ell-1]$

$$R_{\mathrm{s}}(s,\ell) \leqslant R_{\mathrm{s}}(s-u,\ell-v)\max_{0\leqslant z\leqslant 1}\{z^u(1-z)^v+(1-z)^u z^v\}. \tag{12}$$

2. Для любого $v \in [\ell-1]$ и $u = v+s-\ell$, $1 \leqslant u \leqslant s-1$,

$$R_{\mathrm{s}}(s,\ell) \leqslant R_{\mathrm{cs}}(s-u,\ell-v)\max_{0\leqslant z\leqslant 1}\{z^u(1-z)^v+(1-z)^u z^v\}. \tag{13}$$

3. Для любого $v \in [\min(s,\ell)-1]$

$$R_{\mathrm{cs}}(s,\ell) \leqslant R_{\mathrm{cs}}(s-v,\ell-v)\frac{1}{2^{2v}}. \tag{14}$$

Кроме того, было доказано неравенство

$$R_{\mathrm{s}}(s,\ell) \leqslant \min(R_{\mathrm{cf}}(s,\ell-1), R_{\mathrm{cf}}(s-1,\ell)). \tag{15}$$

§ 4. Новые неравенства

Мы начнем с улучшения верхней границы для скорости $(2,1)$-полностью разделяющих кодов. Отметим, что известная наилучшая верхняя оценка совпадала с верхней границей скорости $(2,1)$-свободных от перекрытий кодов и равнялась 0,321929.

Теорема 1. *Справедлива оценка*

$$R_{\mathrm{cs}}(2,1) = R_{\mathrm{cs}}(1,2) \leqslant h(0{,}25) - 0{,}5 = 0{,}311278\ldots \tag{16}$$

Доказательство. Рассмотрим произвольный $(2,1)$-полностью разделяющий код $\mathcal{C}$ длины N и мощности M. Найдем вес w, такой что количество кодовых слов веса w максимально. Без ограничения общности можно считать, что $w \geqslant N/2$, так как в противном случае можно заменить код $\mathcal{C}$ на $\mathcal{C} + \mathbf{1}$.

Так как $(2,1)$-полностью разделяющий код является $(2,1)$-свободным от перекрытий, то можно применить известные оценки на количество слов фиксированного веса, доказанные в [21, 26]. Лемма 3 из [21] или теорема 1 из [26] позволяют оценить количество слов веса w как

$$4\frac{\binom{N}{\lfloor w/2\rfloor}}{\binom{2\lfloor w/2\rfloor}{\lfloor w/2\rfloor}}.$$

Тогда общее количество кодовых слов не превосходит

$$4N\frac{\binom{N}{\lfloor w/2\rfloor}}{\binom{2\lfloor w/2\rfloor}{\lfloor w/2\rfloor}},$$

что в силу хорошо известной асимптотики $\binom{N}{wN} = 2^{N(h(w)+o(1))}$ приводит к оценке

$$R_{\mathrm{cs}}(2,1) \leqslant \max_{0{,}5\leqslant w\leqslant 1}(h(w/2) - w) = h(1/4) - 1/2. \quad \blacktriangle$$

Теорема 2. *Пусть* $s, \ell \geqslant 2$. *Тогда*

$$R_{\mathrm{cs}}(s,\ell) \leqslant \frac{1}{2}\min\bigl(R_{\mathrm{cf}}(s,\ell-1), R_{\mathrm{cf}}(s-1,\ell)\bigr). \tag{17}$$

Доказательство. Так как $R_{\mathrm{s}}(s,\ell) = R_{\mathrm{s}}(\ell,s)$ и $R_{\mathrm{cf}}(s,\ell) = R_{\mathrm{cf}}(\ell,s)$, то мы докажем только неравенство

$$R_{\mathrm{cs}}(s,\ell) \leqslant \frac{1}{2}R_{\mathrm{cf}}(s,\ell-1).$$

Рассмотрим произвольный (s,ℓ)-полностью разделяющий код $\mathcal{C}$ длины N и мощности M и некоторое его слово $\boldsymbol{c}$ веса w. Без ограничения общности можно считать, что $w \leqslant N/2$, ибо в противном случае мы рассмотрим код $\mathcal{C} + \mathbf{1}$. Построим новый код $\mathcal{C}'$, удалив все координаты, в которых выбранное слово $\boldsymbol{c}$ имеет нули. Удалим также и само слово $\boldsymbol{c}$. Проекция кода $\mathcal{C} \setminus \boldsymbol{c}$ на код $\mathcal{C}'$ инъективна. Действительно, пусть два слова $a \neq b$ из кода $\mathcal{C}\setminus\boldsymbol{c}$ при проекции на координаты, где вектор $\boldsymbol{c}$ имеет 1, совпали. Тогда в исходном коде нет координаты j, такой что $a_j = c_j = 1$ и $b_j = 0$, что противоречит тому, что исходный код был $(2,1)$-полностью разделяющим. Таким образом, длина кода $\mathcal{C}'$ равна $w \leqslant N/2$, а мощность равна $M - 1$. Покажем, что код $\mathcal{C}'$ является $(s,\ell-1)$-свободным от перекрытий.

Действительно, рассмотрим произвольные непересекающиеся множества кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| = s$, $|\mathcal{L}| = \ell - 1$, $\mathcal{S}, \mathcal{L} \subset \mathcal{C}'$. Этим множествам соответствуют множества $\widehat{\mathcal{S}}$ и $\widehat{\mathcal{L}}$ исходного кода $\mathcal{C}$. Так как код $\mathcal{C}$ является (s,ℓ)-полностью разделяющим, то для множеств $\widehat{\mathcal{S}}$ и $\widehat{\mathcal{L}} \cup \boldsymbol{c}$ найдется координата i, такая что

$$x_i = 0 \text{ для любого } \boldsymbol{x} \in \widehat{\mathcal{S}} \quad \text{и} \quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \widehat{\mathcal{L}} \cup \boldsymbol{c}.$$

Так как $\boldsymbol{c}_i = 1$, то при построении кода $\mathcal{C}'$ эта координата не была удалена, и значит, в этой координате в коде $\mathcal{C}'$ выполнено

$$x_i = 0 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \mathcal{L}.$$

А это и означает, что код $\mathcal{C}'$ является $(s, \ell-1)$-свободным от перекрытий. Отсюда получаем искомое неравенство

$$R_{\mathrm{cs}}(s,\ell) \leqslant \frac{1}{2} R_{\mathrm{cf}}(s,\ell-1). \quad \blacktriangle$$

Введем дополнительное определение. Для произвольных двух непересекающихся множеств кодовых слов $U, V \subset \mathcal{C}$ определим разделяющее "расстояние" $D(U,V)$ как количество координат, в которых все слова из одного множества имеют символ 0, а все слова из другого множества – символ 1. Будем говорить, что такие координаты разделяют множества U и V. Отметим, что $(1,1)$-разделяющее расстояние – это обычное расстояние Хэмминга, но при других параметрах u и v неравенство треугольника не выполняется.

Будем называть (u,v)-разделяющим расстоянием кода $\mathcal{C}$ величину

$$d_{uv}(\mathcal{C}) = \min_{\substack{U,V\subset\mathcal{C}\\ U\cap V=\varnothing\\ |U|=u,\,|V|=v}} D(U,V),$$

и будем называть ее просто разделяющим расстоянием $d(\mathcal{C})$ кода, когда параметры u и v ясны из контекста.

Определим величину

$$\Delta(u,v) = \max_{0\leqslant z\leqslant 1} \{z^u(1-z)^v + (1-z)^u z^v\}. \tag{18}$$

Следующее утверждение является обобщением классической границы Плоткина.

Лемма 1. *Для произвольного кода $\mathcal{C}$ длины N с (u,v)-разделяющим расстоянием d справедливо*

$$|\mathcal{C}| \leqslant \frac{u+v-1}{1-\left(\dfrac{N\Delta}{d}\right)^{1/(u+v-1)}},$$

если $d/N > \Delta = \Delta(u,v)$.

Доказательство. Рассмотрим код $\mathcal{C}$ длины N и мощности M и оценим сумму S всех (u,v)-разделяющих расстояний по всем парам непересекающихся кодовых подмножеств U, V таких, что $|U| = u, |V| = v$. Очевидно, что

$$S = \sum_{\substack{U,V\subset\mathcal{C}\\ |U|=u,\,|V|=v}} D(U,V) \geqslant \binom{M}{u+v}\binom{u+v}{u} d \geqslant \frac{M(M-u-v+1)^{u+v-1}}{u!\,v!} d.$$

С другой стороны, если в i-й координате слов кода имеется A единиц (и $M-A$ нулей), то вклад этой координаты в S равен

$$\binom{A}{u}\binom{M-A}{v} + \binom{A}{v}\binom{M-A}{u},$$

и следовательно,

$$S \leqslant N \max_{0\leqslant A\leqslant M} \left\{\binom{A}{u}\binom{M-A}{v} + \binom{A}{v}\binom{M-A}{u}\right\} \leqslant N\frac{M^{u+v}}{u!\,v!}\Delta.$$

Объединяя верхнюю и нижнюю оценки величины S, получаем

$$(M-u-v+1)^{u+v-1}d \leqslant NM^{u+v-1}\Delta.$$

Отсюда следует

$$1-\frac{u+v-1}{M} \leqslant \left(\frac{N\Delta}{d}\right)^{1/(u+v-1)},$$

и так как $d/N > \Delta$ по условиям леммы, то

$$M \leqslant \frac{u+v-1}{1-\left(\frac{N\Delta}{d}\right)^{1/(u+v-1)}}. \quad \blacktriangle$$

Как уже отмечалось выше, $(1,1)$-разделяющее расстояние – это расстояние Хэмминга, $\Delta(1,1)=1/2$, и доказанная граница превращается в этом случае в классическую границу Плоткина: если $2d > N$, то $|\mathcal{C}| \leqslant \frac{2d}{2d-N}$.

Максимальную мощность кода длины N с (u,v)-разделяющим расстоянием d будем обозначать через $M_{u,v}(N,d)$. Следующая лемма является аналогом асимптотической границы Плоткина.

Лемма 2. *Для любого* τ, $0<\tau<\Delta=\Delta(u,v)$, *справедливо*

$$\log_2 M_{u,v}(N,\lfloor \tau N \rfloor) \leqslant N\left(1-\frac{\tau}{\Delta}+o(1)\right). \tag{19}$$

Доказательство. Сначала докажем неравенство

$$M_{u,v}(N,d) \leqslant 2M_{u,v}(N-1,d). \tag{20}$$

Пусть $\mathcal{C}$ – код максимальной мощности $M_{u,v}(N,d)$ длины N с (u,v)-разделяющим расстоянием d. Выберем произвольную координату и без ограничения общности будем считать, что в словах кода $\mathcal{C}$ в ней больше нулей, чем единиц. Удалим эту координату и все кодовые слова, имеющие единицу в этой координате. У полученного кода $\mathcal{C}'$ его (u,v)-разделяющее расстояние не уменьшится, длина кода уменьшится на 1, а число слов уменьшится не более чем в два раза, что и доказывает неравенство (20).

Применяя неравенство (20) i раз, получаем

$$\log_2 M_{u,v}(N,d) \leqslant i + \log_2 M_{u,v}(N-i,d). \tag{21}$$

Теперь возьмем минимальное целое i, такое что $\frac{d}{N-i} \geqslant \frac{\Delta}{1-\varepsilon}$ для некоторого $\varepsilon>0$, т.е. $i = N - \left\lfloor \frac{d(1-\varepsilon)}{\Delta} \right\rfloor$. Тогда из леммы 1 получаем

$$M_{u,v}(N-i,d) \leqslant \frac{u+v-1}{1-\left(\frac{(N-i)\Delta}{d}\right)^{1/(u+v)}} \leqslant \frac{u+v-1}{1-(1-\varepsilon)^{1/(u+v-1)}}.$$

Полагая $\varepsilon = 1/N$, получаем, что

$$M(N-i,d) \leqslant c(u,v)N + o(N),$$

где $c(u,v)$ – некоторая константа, зависящая от u и v, но не зависящая от N. Теперь оценку (21) можно переписать в виде

$$\log_2 M_{u,v}(N,d) \leqslant N - \frac{d(1-\varepsilon)}{\Delta} + \log_2((c(u,v)+o(1))N) =$$
$$= N\left(1 - \frac{d}{N\Delta}\right) + o(N). \quad \blacktriangle \tag{22}$$

Замечание 1. Из леммы, в частности, следует, что не существует разделяющих кодов с положительной асимптотической скоростью и относительным разделяющим расстоянием $\geqslant \Delta$. Несложно доказать, что для любого $\tau < \Delta$ разделяющие коды с положительной асимптотической скоростью и относительным разделяющим расстоянием τ существуют, а значит, Δ является критической точкой. Доказательство этого факта мы приводим в Приложении.

Теперь мы готовы доказать основную теорему статьи. Доказательство основывается на обобщении упомянутой во введении идеи Ю.Л. Сагаловича на случай произвольных (s,ℓ)-разделяющих кодов и на доказанной выше асимптотической границе Плоткина для разделяющих кодов.

Т е о р е м а 3. *Для любых* $u \in [s-1]$, $v \in [\ell-1]$

$$R_{\mathrm{s}}(s,\ell) \leqslant \frac{R_{\mathrm{cf}}(s-u,\ell-v)\cdot}{R_{\mathrm{cf}}(s-u,\ell-v) + \Delta^{-1}(u,v)}. \tag{23}$$

Д о к а з а т е л ь с т в о. Докажем, что у любого (s,ℓ)-разделяющего кода $\mathcal{C}$ его минимальное (u,v)-разделяющее расстояние $d = d_{u,v}(C)$ достаточно велико, а именно

$$d_{u,v}(C) \geqslant N_{\mathrm{cf}}(|C| - u - v, s-u, \ell-v). \tag{24}$$

Возьмем два непересекающихся множества кодовых слов $U, V \subset C$, $|U| = u$, $|V| = v$, на которых достигается минимальное (u,v)-разделяющее расстояние d кода $\mathcal{C}$. Обозначим через $I \subset [N]$ множество координат, разделяющих U и V. Как уже отмечалось, свойство разделимости инвариантно относительно сдвига кода на произвольный вектор, и следовательно, можно считать, что во всех координатах из множества I слова из U имеют нули, а слова из V – единицы. Рассмотрим код $\mathcal{C}'$ длины d, полученный из кода $\mathcal{C}\setminus\{U\cup V\}$ его укорочением (проекцией) на множество координат I. Покажем, что код $\mathcal{C}'$ является $(s-u,\ell-v)$-свободным от перекрытий кодом мощности $|C| - u - v$.

Действительно, рассмотрим произвольные непересекающиеся множества кодовых слов $\mathcal{S}', \mathcal{L}' \subset C'$, такие что $|\mathcal{S}'| = s-u$, $|\mathcal{L}'| = \ell - v$, и соответствующие им множества $\mathcal{S}$ и $\mathcal{L}$ исходного кода $\mathcal{C}$. Рассмотрим множества $\widehat{\mathcal{S}} = \mathcal{S} \cup U$ и $\widehat{\mathcal{L}} = \mathcal{L} \cup V$. Так как код $\mathcal{C}$ является (s,ℓ)-разделяющим, то найдется координата $i \in [N]$, которая разделяет множества $\widehat{\mathcal{S}}$ и $\widehat{\mathcal{L}}$, а значит, она разделяет и множества U и V, а следовательно, $i \in I$. Так как во всех координатах из множества I слова из U имеют нули, а слова из V – единицы, то код $\mathcal{C}'$ является $(s-u,\ell-v)$-свободным от перекрытий. Отсюда, в частности, следует, что при укорочении кода $\mathcal{C}\setminus\{U\cup V\}$ никакие два слова не могли совпасть, т.е. $|\mathcal{C}'| = |\mathcal{C}| - u - v$, и неравенство (24) доказано.

В силу неравенства (24) перепишем асимптотическую границу Плоткина (19) при $\tau < \Delta = \Delta(u,v)$ в виде

$$\log_2 M_{u,v}(N, \lfloor \tau N \rfloor) \leqslant N\left(1 - \frac{\tau}{\Delta} + o(1)\right) \leqslant$$
$$\leqslant N\left(1 - \frac{N_{\mathrm{cf}}(M_{u,v}(N,\lfloor\tau N\rfloor) - u - v, s-u, \ell-v)}{N\Delta} + o(1)\right).$$

Таблица 1

Верхние границы скоростей разделяющих кодов

$s\backslash\ell$	1	2	3	4	5	6	7	8
1	1,000000	0,500000	0,321929	0,199282	0,140457	0,105641	0,083000	0,067305
2	0,500000	0,283477	0,116879	0,066265	0,038684	0,024305	0,016336	0,011569
3	0,321929	0,116879	0,066265	0,028695	0,015326	0,008215	0,005271	0,003270
4	0,199282	0,066265	0,028695	0,015326	0,007088	0,003703	0,001912	0,001090
5	0,140457	0,038684	0,015326	0,007088	0,003703	0,001761	0,000912	0,000463
6	0,105641	0,024305	0,008215	0,003703	0,001761	0,000912	0,000439	0,000226
7	0,083000	0,016336	0,005271	0,001912	0,000912	0,000439	0,000226	0,000110
8	0,067305	0,011569	0,003270	0,001090	0,000463	0,000226	0,000110	0,000056

Таблица 2

Способ получения верхних границ скоростей кодов из табл. 1

$s\backslash\ell$	1	2	3	4	5	6	7	8
1		H9	H15	H15	H15	H15	H15	H15
2	H9	H10	H10 + T1	H10 + T2	T3 (1, 3)	T3 (1, 3)	T3 (1, 4)	T3 (1, 4)
3	H15	H10 + T1	H15	H10 + T1	H15	T3 (1, 3)	T3 (2, 5)	T3 (2, 5)
4	H15	H10 + T2	H10 + T1	H15	H10 + T1	H15	T3 (1, 3)	T3 (2, 5)
5	H15	T3 (3, 1)	H15	H10 + T1	H15	H10 + T1	H15	T3 (1, 3)
6	H15	T3 (3, 1)	T3 (3, 1)	H15	H10 + T1	H15	H10 + T1	H15
7	H15	T3 (4, 1)	T3 (5, 2)	T3 (3, 1)	H15	H10 + T1	H15	H10 + T1
8	H15	T3 (4, 1)	T3 (5, 2)	T3 (5, 2)	T3 (3, 1)	H15	H10 + T1	H15

Подставив туда очевидное соотношение $N_{\mathrm{cf}}(M, a, b) \geqslant \log_2 M(R_{\mathrm{cf}} + o(1))^{-1}$, получим, что

$$\log_2 M_{u,v}(N, \lfloor \tau N \rfloor) \leqslant N \left(1 - \frac{\log_2 M_{u,v}(N, \lfloor \tau N \rfloor)}{N\Delta(R_{\mathrm{cf}} + o(1))} + o(1)\right),$$

или, что равносильно,

$$\log_2 M_{u,v}(N, \lfloor \tau N \rfloor) \left(1 + \frac{1}{R_{\mathrm{cf}}\Delta} + o(1)\right) \leqslant N(1 + o(1)). \quad \blacktriangle$$

Подчеркнем, что эта теорема ограничивает скорость разделяющих кодов через скорость свободных от перекрытий кодов, в отличие от неравенства (12), где в правой части неравенства присутствует скорость разделяющих кодов. В неравенстве (13) скорость разделяющих кодов ограничивается через скорость полностью разделяющих кодов, однако неравенство выполняется лишь для $u = v + s - \ell$, тогда как теорема 3 выполняется для всех u и v.

§ 5. Сравнения и таблицы

В этом параграфе мы приводим численные значения верхних границ асимптотической скорости разделяющих, полностью разделяющих и свободных от перекрытий кодов. Дополнительно мы предоставляем таблицы, где указано, с помощью какой именно теоремы (T) или неравенства (H) был получен тот или иной результат.

Результаты для разделяющих кодов приведены в табл. 1, 2. В табл. 2 в скобках указаны значения параметров u и v из теоремы 3, дающие наилучшие значения. Из таблиц видно, что новые теоремы позволяют улучшить верхние оценки скорости для многих значений параметров s и ℓ.

Таблица 3

Верхние границы скоростей полностью разделяющих кодов

$s\backslash\ell$	1	2	3	4	5	6	7	8
1	1,000000	0,311278	0,160964	0,099641	0,070228	0,052820	0,041500	0,033652
2	0,311278	0,160964	0,064317	0,033133	0,021507	0,014338	0,010192	0,007299
3	0,160964	0,064317	0,033133	0,014895	0,007663	0,004861	0,003166	0,002115
4	0,099641	0,033133	0,014895	0,007663	0,003601	0,001852	0,001133	0,000719
5	0,070228	0,021507	0,007663	0,003601	0,001852	0,000887	0,000456	0,000265
6	0,052820	0,014338	0,004861	0,001852	0,000887	0,000456	0,000220	0,000113
7	0,041500	0,010192	0,003166	0,001133	0,000456	0,000220	0,000113	0,000055
8	0,033652	0,007299	0,002115	0,000719	0,000265	0,000113	0,000055	0,000028

Таблица 4

Способ получения верхних границ скоростей кодов из табл. 3

$s\backslash\ell$	1	2	3	4	5	6	7	8
1		T1	T2	T2	T2	T2	T2	T2
2	T1	CF	H11 + T1	T2	T2	T2	T2	T2
3	T2	H11 + T1	CF	H11 + T1	T2	T2	T2	T2
4	T2	T2	H11 + T1	CF	H11 + T1	T2	T2	T2
5	T2	T2	T2	H11 + T1	CF	H11 + T1	T2	T2
6	T2	T2	T2	T2	H11 + T1	CF	H11 + T1	T2
7	T2	T2	T2	T2	T2	H11 + T1	CF	H11 + T1
8	T2	T2	T2	T2	T2	T2	H11 + T1	CF

Для полностью разделяющих кодов (табл. 3, 4) мы улучшаем все значения, кроме диагональных, где границы совпадают с границами свободных от перекрытий кодов.

Для свободных от перекрытий кодов (табл. 5, 6) мы не получаем новых результатов, все оценки получены с помощью ранее известных теорем. Но, как отмечалось ранее, верхние оценки скоростей свободных от перекрытий кодов используются для получения оценок скоростей разделяющих и полностью разделяющих кодов. Насколько нам известно, ранее такие таблицы, учитывающие все известные теоремы, не публиковались.

Как и в случае разделяющих кодов, в табл. 6 значения в скобках показывают параметры u и v, использующиеся для получения оптимальных границ с помощью неравенства (7).

ПРИЛОЖЕНИЕ: НИЖНЯЯ ГРАНИЦА СКОРОСТИ КОДОВ С РАЗДЕЛЯЮЩИМ РАССТОЯНИЕМ

Теорема 4. *Максимальная мощность $M_{u,v}(N,d)$ кода с (u,v)-разделяющим расстоянием $d = \lfloor \tau N \rfloor$ удовлетворяет неравенству*

$$\log_2 M_{u,v}(N,d) \geqslant N \frac{-h(\tau) - \tau \log_2 \Delta(u,v) - (1-\tau)\log_2(1-\Delta(u,v)) + o(1)}{u+v-1} \qquad (25)$$

при $0 \leqslant \tau < \Delta(u,v)$, где величина $\Delta(u,v)$ определена в (18).

Отметим, что правая часть положительна при $\tau < \Delta(u,v)$ и стремится к нулю при $\tau \to \Delta(u,v)$. При $u = v = 1$ точка $\Delta(1,1) = 1/2$, и наша нижняя граница превращается в границу Варшамова – Гильберта.

Таблица 5

Верхние границы скоростей свободных от перекрытий кодов

$s\backslash\ell$	1	2	3	4	5	6	7	8
1	1,000000	0,321929	0,199282	0,140457	0,105641	0,083000	0,067305	0,055905
2	0,321929	0,160964	0,066265	0,043015	0,028677	0,020384	0,014598	0,011019
3	0,199282	0,066265	0,033133	0,015326	0,009722	0,006332	0,004230	0,003011
4	0,140457	0,043015	0,015326	0,007663	0,003703	0,002265	0,001438	0,000937
5	0,105641	0,028677	0,009722	0,003703	0,001852	0,000912	0,000529	0,000333
6	0,083000	0,020384	0,006332	0,002265	0,000912	0,000456	0,000226	0,000128
7	0,067305	0,014598	0,004230	0,001438	0,000529	0,000226	0,000113	0,000056
8	0,055905	0,011019	0,003011	0,000937	0,000333	0,000128	0,000056	0,000028

Таблица 6

Номер неравенства, из которого получены значения в табл. 5

$s\backslash\ell$	1	2	3	4	5	6	7	8
1		H4	H4	H4	H4	H4	H4	H4
2	H4	H5	H8	H8	H7 (1, 2)	H7 (1, 2)	H7 (1, 3)	H7 (1, 3)
3	H4	H8	H5	H8	H7 (1, 2)	H7 (1, 2)	H7 (1, 2)	H7 (1, 2)
4	H4	H8	H8	H5	H8	H7 (1, 2)	H7 (1, 2)	H7 (1, 2)
5	H4	H7 (2, 1)	H7 (2, 1)	H8	H5	H8	H7 (2, 3)	H7 (3, 5)
6	H4	H7 (2, 1)	H7 (2, 1)	H7 (2, 1)	H8	H5	H8	H7 (2, 3)
7	H4	H7 (3, 1)	H7 (2, 1)	H7 (2, 1)	H7 (3, 2)	H8	H5	H8
8	H4	H7 (3, 1)	H7 (2, 1)	H7 (2, 1)	H7 (5, 3)	H7 (3, 2)	H8	H5

Доказательство. Рассмотрим случайный код длины N и мощности M, где каждый элемент каждого кодового слова выбирается независимо и равен 1 с вероятностью p.

Вероятность того, что фиксированная координата разделяет два набора кодовых слов размера u и v равна

$$q = p^u(1-p)^v + p^v(1-p)^u.$$

Параметр p мы выберем так, чтобы максимизировать q. Отметим, что максимум вероятности разделения равен в точности величине $\Delta(u, v)$, определенной в формуле (18). Число ξ координат, разделяющих два фиксированных набора кодовых слов, имеет биномиальное распределение с параметрами N и $\Delta = \Delta(u, v)$. Оценим вероятность того, что $\xi < d$ в предположении $\Delta(u, v) > \tau$:

$$\mathcal{P} = \sum_{k=0}^{d-1} \binom{N}{k} \Delta^k (1-\Delta)^{N-k} \leqslant N 2^{N(h(\tau)+\tau\log_2\Delta+(1-\tau)\log_2(1-\Delta)+o(1))}.$$

Таким образом,

$$N^{-1}\log_2\mathcal{P} = h(\tau) + \tau\log_2\Delta + (1-\tau)\log_2(1-\Delta) + o(1).$$

Математическое ожидание количества наборов кодовых слов, разделяющее расстояние между которыми меньше d, не превосходит $\mathcal{P}\cdot M^{u+v}$. Стандартное рассуждение с выбрасыванием приводит к оценке

$$\log_2 M_{u,v}(N, d) \geqslant N\frac{-h(\tau) - \tau\log_2\Delta - (1-\tau)\log_2(1-\Delta) + o(1)}{u+v-1},$$

что и требовалось доказать. ▲

СПИСОК ЛИТЕРАТУРЫ

1. *Сагалович Ю.Л.* Метод повышения надежности конечного автомата // Пробл. передачи информ. 1965. Т. 1. № 2. С. 27–35. http://mi.mathnet.ru/ppi734
2. *Friedman A.D., Graham R.L., Ullman J.D.* Universal Single Transition Time Asynchronous State Assignments // IEEE Trans. Comput. 1969. V. 18. № 6. P. 541–547. https://doi.org/10.1109/T-C.1969.222707
3. *Barg A., Blakley G.R., Kabatiansky G.A.* Digital Fingerprinting Codes: Problem Statements, Constructions, Identification of Traitors // IEEE Trans. Inform. Theory. 2003. V. 49. № 4. P. 852–865. https://doi.org/10.1109/TIT.2003.809570
4. *Stinson D.R., Wei R., Chen K.* On Generalized Separating Hash Families // J. Combin. Theory Ser. A. 2008. V. 115. № 1. P. 105–120. https://doi.org/10.1016/j.jcta.2007.04.005
5. *Сагалович Ю.Л.* Полностью разделяющие системы // Пробл. передачи информ. 1982. Т. 18. № 2. С. 74–82. http://mi.mathnet.ru/ppi1227
6. *Пинскер М.С., Сагалович Ю.Л.* Нижняя граница мощности кода состояний автомата // Пробл. передачи информ. 1972. Т. 8. № 3. С. 58–66. http://mi.mathnet.ru/ppi854
7. *Randriambololona H.* $(2,1)$-Separating Systems beyond the Probabilistic Bound // Israel J. Math. 2013. V. 195. № 1. P. 171–186. https://doi.org/10.1007/s11856-012-0126-9
8. *Сагалович Ю.Л.* Верхняя граница мощности кода состояний автомата // Пробл. передачи информ. 1973. Т. 9. № 1. С. 73–83. http://mi.mathnet.ru/ppi884
9. *Körner J., Simonyi G.* Separating Partition Systems and Locally Different Sequences // SIAM J. Discrete Math. 1988. V. 1. № 3. P. 355–359. https://doi.org/10.1137/0401035
10. *Сагалович Ю.Л.* Новые верхние границы мощности разделяющих систем // Пробл. передачи информ. 1993. Т. 29. № 2. С. 109–111. http://mi.mathnet.ru/ppi182
11. *Bassalygo L.A., Burmester M., Dyachkov A., Kabatianskii G.* Hash Codes // Proc. 1997 IEEE Int. Sympos. on Information Theory (ISIT'97). Ulm, Germany. June 29 – July 4, 1997. P. 174. https://doi.org/10.1109/ISIT.1997.613089
12. *D'yachkov A.G., Vilenkin P.A., Yekhanin S.M.* Upper Bounds on the Rate of Superimposed (s,ℓ)-Codes Based on Engel's Inequality // Proc. 8th Int. Workshop on Algebraic and Combinatorial Coding Theory (ACCT-8). Tsarskoe Selo, Russia. Sept. 8–14, 2002. P. 95–99.
13. *Лебедев В.С.* Асимптотическая верхняя граница для скорости кодов, свободных от (w,r)-перекрытий // Пробл. передачи информ. 2003. Т. 39. № 4. С. 3–9. http://mi.mathnet.ru/ppi311
14. *Cohen G.D., Schaathun H.G.* Asymptotic Overview on Separating Codes // Tech. Rep. № 248. Dept. of Informatics, Univ. of Bergen. Bergen, Norway, 2003. Available at http://www.ii.uib.no/~georg/sci/inf/coding/hyperpdf/cs03rep.pdf.
15. *Воробьев И.В.* Границы скоростей разделяющих кодов // Пробл. передачи информ. 2017. Т. 53. № 1. С. 34–46. http://mi.mathnet.ru/ppi2225
16. *Hollmann H.D.L., van Lint J.H., Linnartz J.-P., Tolhuizen L.M.G.M.* On Codes with the Identifiable Parent Property // J. Combin. Theory Ser. A. 1998. V. 82. № 2. P. 121–133. https://doi.org/10.1006/jcta.1997.2851
17. *Кабатянский Г.А.* Идентифицирующие коды и их обобщения // Пробл. передачи информ. 2019. Т. 55. № 3. С. 93–105. https://doi.org/10.1134/S0555292319030070
18. *Kautz W., Singleton R.* Nonrandom Binary Superimposed Codes // IEEE Trans. Inform. Theory. 1964. V. 10. № 4. P. 363–377. https://doi.org/10.1109/TIT.1964.1053689
19. *Mitchell C.J., Piper F.C.* Key Storage in Secure Networks // Discrete Appl. Math. 1988. V. 21. № 3. P. 215–228. https://doi.org/10.1016/0166-218X(88)90068-6
20. *Magó G.* Monotone Functions in Sequential Circuits // IEEE Trans. Comput. 1973. V. 22. № 10. P. 928–933. https://doi.org/10.1109/T-C.1973.223620
21. *Дьячков А.Г., Рыков В.В.* Границы длины дизъюнктивных кодов // Пробл. передачи информ. 1982. Т. 18. № 3. С. 7–13. http://mi.mathnet.ru/ppi1232

22. *Stinson D.R., Wei R., Zhu L.* Some New Bounds for Cover-Free Families // J. Combin. Theory Ser. A. 2000. V. 90. № 1. P. 224–234. https://doi.org/10.1006/jcta.1999.3036
23. *D'yachkov A., Vilenkin P., Macula A., Torney D.* Families of Finite Sets in Which No Intersection of ℓ Sets Is Covered by the Union of s Others // J. Combin. Theory Ser. A. 2002. V. 99. № 2. P. 195–218. https://doi.org/10.1006/jcta.2002.3257
24. *Lebedev V.S.* Some Tables for (w, r)-Superimposed Codes // Proc. 8th Int. Workshop on Algebraic and Combinatorial Coding Theory (ACCT-8). Tsarskoe Selo, Russia. Sept. 8–14, 2002. P. 185–189.
25. *Blackburn S.R.* Frameproof Codes // SIAM J. Discrete Math. 2003. V. 16. № 3. P. 499–510. https://doi.org/10.1137/S0895480101384633
26. *Erdős P., Frankl P., Füredi Z.* Families of Finite Sets in Which No Set Is Covered by the Union of Two Others // J. Combin. Theory Ser. A. 1982. V. 33. № 2. P. 158–166. https://doi.org/10.1016/0097-3165(82)90004-8

Воробьев Илья Викторович
Сколковский институт науки и технологий (Сколтех), Москва
vorobyev.i.v@yandex.ru
Лебедев Владимир Сергеевич
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
lebedev37@mail.ru

Поступила в редакцию
14.04.2022
После доработки
28.07.2022
Принята к публикации
30.07.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 2022 Вып. 3

УДК 621.391.1 : 519.725

© 2022 г. **Ф.И. Соловьева**

РАЗБИЕНИЯ НА СОВЕРШЕННЫЕ КОДЫ В МЕТРИКАХ ХЭММИНГА И ЛИ[1]

Предложены новые комбинаторные конструкции разбиений на совершенные коды в метриках Хэмминга и Ли. Кроме того, приведен новый комбинаторный метод построения диаметральных совершенных кодов в метрике Ли, который развит для построения разбиений на такие коды. Для метрики Ли улучшены известные нижние оценки числа совершенных и диаметральных совершенных кодов Ли, предложенные Этционом в 2011 г.

Ключевые слова: совершенный код, совершенный код в метрике Хэмминга, совершенный код в метрике Ли, диаметральный совершенный код в метрике Ли, разбиения, разбиения на совершенные коды.

DOI: 10.31857/S0555292322030056, **EDN:** EAJWAD

§ 1. Введение

Целью данной статьи является развитие новых комбинаторных методов построения разбиений на совершенные коды в метриках Хэмминга и Ли, а также конструкций диаметральных совершенных кодов Ли и разбиений на такие коды. Оказалось, что идеи некоторых подходов для построения совершенных q-ичных кодов, $q \geqslant 2$, в метрике Хэмминга могут быть развиты для построения кодов и разбиений на совершенные коды Ли и диаметральные совершенные коды Ли. Приведенные конструкции позволяют существенно улучшить известные нижние оценки числа совершенных и диаметральных совершенных кодов Ли, предложенные Этционом в 2011 г. (см. [1]).

В отличие от метрики Хэмминга, для которой получено большое число различных свитчинговых и каскадных методов построения совершенных кодов и разбиений, для построения кодов и разбиений в метрике Ли предложено лишь незначительное число конструкций. Мотивация исследования совершенных и диаметральных совершенных кодов Ли и основательный обзор полученных по данной тематике результатов могут быть найдены в работе [1], где представлены две конструкции совершенных и диаметральных совершенных кодов Ли, и поэтому они не приводятся в настоящей статье. Алгебраические методы построения совершенных и диаметральных совершенных кодов Ли см. также в работах [2, 3]. Гипотеза о несуществовании совершенных кодов, исправляющих две ошибки, была выдвинута в 1970 г. в работе [4]. О совершенных кодах Ли, исправляющих две ошибки, см. также в [5]; о существовании и некоторых необходимых условиях для квазисовершенных кодов в метрике Ли см. [6].

Упор в данной статье сделан на построение разбиений, поскольку этот вопрос оставался недостаточно глубоко изученным как для метрики Ли, так и для метрики Хэмминга. Кроме того, это дает возможность строить коды, используя эти

[1] Исследование выполнено за счет гранта Российского научного фонда № 22-21-00135, https://rscf.ru/project/22-21-00135/

разбиения, что позволило получить новые нижние оценки числа совершенных и диаметральных совершенных кодов Ли. Отметим также, что все конструкции кодов и разбиений в данной статье являются комбинаторными, что может представлять интерес с практической точки зрения.

Статья имеет следующую структуру: в § 2 приводятся необходимые определения и понятия, § 3 посвящен построению разбиений на совершенные коды над полем Галуа $\mathbb{F}_q$, $q > 2$. В § 4 приведена конструкция построения разбиений на совершенные коды Ли, идейно восходящая к подходу, использованному для построения разбиений в § 3. В § 5 предложены новые диаметральные совершенные коды Ли и разбиения на такие коды. В каждом из последних трех параграфов приводятся нижние оценки числа кодов и разбиений и сравнение их с полученными ранее нижними оценками.

§ 2. Совершенные и диаметральные совершенные коды в метрике Ли

Векторное пространство размерности n над полем Галуа $\mathbb{F}_q$ по отношению к метрике Хэмминга обозначается через $\mathbb{F}_q^n$. Основные определения, касающиеся кодов в метрике Хэмминга, см. в [7].

В данном параграфе приведем необходимые определения и понятия для метрики Ли. Сначала рассмотрим определения для совершенных кодов Ли, затем для диаметральных совершенных кодов Ли.

Через $\mathbb{Z}_s^n$ обозначим множество слов длины n над кольцом вычетов $\mathbb{Z}_s$ по модулю s. *Вес Ли* $w_L(x)$ слова x из множества слов $\mathbb{Z}_s^n$ определяется как сумма весов Ли его координатных позиций. *Расстояние Ли*, обозначаемое через $d_L(x, y)$, для произвольных слов $x, y \in \mathbb{Z}_s^n$ определяется как

$$d_L(x, y) = \sum_{i=1}^{n} \min(|x_i - y_i|, s - |x_i - y_i|).$$

В настоящей статье рассматриваются совершенные коды Ли с минимальным расстоянием Ли $d_L = 3$. Такой код длины n над кольцом вычетов $\mathbb{Z}_s$ имеет мощность $s^n/(2n+1)$, где $2n+1$ – размер сферы Ли радиуса 1, а $s \geqslant 2n+1$. Код в $\mathbb{Z}_s^n$ *линеен*, если он образует подгруппу кольца $\mathbb{Z}_s^n$. Линейный совершенный код Ли длины n с минимальным расстоянием 3 над кольцом вычетов $\mathbb{Z}_s$ существует согласно [2] тогда и только тогда, когда $\tau \,|\, s$, где τ равно произведению всех простых делителей числа $2n + 1$. При этом наименьшее s, для которого существует совершенный код Ли с минимальным расстоянием 3 над $\mathbb{Z}_s$, равно τ. В случаях $s = 2$ и $s = 3$ метрика Ли совпадает с метрикой Хэмминга для двоичных и троичных кодов, что неверно при $s > 3$.

Как и совершенные коды в метрике Хэмминга, совершенные коды в метрике Ли с минимальным расстоянием 3 разбивают множество слов $\mathbb{Z}_s^n$ на смежные классы посредством сдвигов на слова веса Ли, равного единице.

В случае диаметральных совершенных кодов Ли важным является понятие антикода, как и для диаметральных совершенных кодов в метриках Хэмминга, Джонсона и Грассмана (см. [8,9]). Антикод с максимальным расстоянием $d-1$ (подмножество слов A в $\mathbb{Z}_s^n$, таких что $d_L(x, y) < d$, где $x, y \in A$), удовлетворяет неравенству Дельсарта [8]

$$|C||A| \leqslant s^n,$$

здесь C – код с минимальным расстоянием d в $\mathbb{Z}_s^n$. Если оценка в неравенстве Дельсарта точна, то C называется *диаметральным совершенным кодом Ли*. В статье рассматриваются диаметральные совершенные коды Ли только с $d_L = 4$, объем

такого кода равен $|C| = s^n/4n$, где $4n$ – размер антикода. Пусть

$$n = 2^i p_1^{i_1} \dots p_k^{i_k}$$

– разложение числа n в произведение степеней простых сомножителей, где $p_r > 2$, $r = 1, \dots, k$, и может быть, $i = 0$. Тогда согласно [3, теорема 13] линейный диаметральный совершенный код Ли длины n с минимальным расстоянием 4 существует над кольцом $\mathbb{Z}_s$ тогда и только тогда, когда

$$s = 2^{i'} p_1^{i'_1} \dots p_k^{i'_k}, \quad \text{где } 2 \leqslant i' \leqslant i+2 \text{ и } 1 \leqslant i'_r \leqslant i_r.$$

Наименьшее s, для которого существует диаметральный совершенный код Ли длины n с минимальным расстоянием 4 над $\mathbb{Z}_s$, равно 4τ, где $\tau = p_1 \dots p_k$, т.е. s четно. Длина кода не обязательно равна степени числа 2.

Аналогично расширенным совершенным кодам в метрике Хэмминга диаметральные совершенные коды Ли разбивают множество слов четного и нечетного весов в $\mathbb{Z}_s^n$ на смежные классы. Существенная разница между двоичными совершенными кодами в метрике Хэмминга и диаметральными совершенными кодами Ли состоит в следующем. Произвольный двоичный совершенный код в метрике Хэмминга можно расширить посредством общей проверки на четность, в результате чего получится диаметральный совершенный код (и наоборот при выкалывании последнего). При этом минимальное расстояние кода увеличится на единицу. Подобная процедура в случае диаметральных совершенных кодов Ли невозможна.

Для q-ичных совершенных кодов, $q > 2$, задача о расширении кодов с минимальным расстоянием 3 до кодов с расстоянием 4 все еще остается нерешенной. Беспалов в [10] доказал несуществование q-ичных расширенных совершенных кодов в метрике Хэмминга в случаях, когда $q = 3$, 4, а $n > q+2$, или оба числа n и q нечетны, а также доказал несуществование нелинейных МДР-кодов с параметрами $(q+2, q^{q-1}, 4)_q$ в случае, когда q нечетно.

§ 3. Разбиения $\mathbb{F}_q^n$ на совершенные коды

В этом параграфе рассмотрим коды в метрике Хэмминга и применение конструкции Думера [11] для построения разбиений на q-ичные совершенные коды в $\mathbb{F}_q^n$, $q > 2$.

Пусть $P_1 = \{C_1, C_2, \dots, C_{q^r}\}$ – произвольное разбиение на q-ичные совершенные коды в $\mathbb{F}_q^n$, $q > 2$, с параметрами $(n, q^{n-r}, 3)_q$, где $n = \dfrac{q^r - 1}{q - 1}$. Пусть D – произвольный q^r-ичный совершенный код в $\mathbb{F}_{q^r}^\ell$, $q > 2$, где

$$\ell = \frac{q^{rs} - 1}{q^r - 1}, \quad |D| = q^{r(\ell - s)},$$

т.е. код, имеющий параметры $(\ell, q^{r(\ell-s)}, 3)_{q^r}$. Множество векторов

$$C = \bigcup_{(x_1, x_2, \dots, x_\ell) \in D} C_{x_1} \times C_{x_2} \times \dots \times C_{x_\ell} \tag{1}$$

является q-ичным совершенным кодом, имеющим параметры $(n\ell, |D||C_1|^\ell, 3)_q$, где длина кода равна $n\ell = \dfrac{q^{rs} - 1}{q - 1}$. Эта каскадная конструкция является спецификацией обобщенной каскадной конструкции Зиновьева [12] для q-ичных кодов, а также непосредственным обобщением каскадной конструкции для двоичных совершенных кодов из [13,14]. Заметим, что автор работы [1] (см. теорему 10 в ней) ошибочно полагает, что эта конструкция является новой. В отличие от оригинальной конструкции

Думера [11], где рассматривается разбиение на классы смежности совершенного кода C_1, в (1) используются произвольные разбиения на q-ичные совершенные коды в $\mathbb{F}_q^n$, $q > 2$.

Кроме того, к кодам C_{x_i}, $i \in \{1, 2, \ldots, \ell\}$, где $x = (x_1, x_2, \ldots, x_\ell) \in D$, можно применить произвольные ℓ подстановок π_t, $t \in \{1, 2, \ldots, \ell\}$, на алфавите кода D, т.е. на $\mathbb{F}_{q^r}$. Пусть $C_{\pi(x_i)}$ – результат действия подстановки π на коде C_{x_i}. Без ограничения общности первую подстановку можно положить тождественной, т.е. $C_{\pi_1(x_1)} = C_{x_1}$. В этом случае конструкция (1) преобразуется в следующую:

$$C = \bigcup_{x=(x_1,x_2,\ldots,x_\ell)\in D} C_{x_1} \times C_{\pi_2(x_2)} \times \ldots \times C_{\pi_\ell(x_\ell)}. \tag{2}$$

Перейдем к построению разбиений посредством конструкции (2). Для этой цели рассмотрим помимо произвольного разбиения P_1, введенного выше, произвольное разбиение $P_2 = \{D_1, D_2, \ldots, D_{q^{rs}}\}$ на q^r-ичные совершенные коды D_i длины ℓ, $i \in \{1, 2, \ldots, q^{rs}\}$. Для построения кодов C_i^* применим конструкцию (2) к кодам D_i и разбиению P_1, а именно

$$C_i^* = \bigcup_{x=(x_1,x_2,\ldots,x_\ell)\in D_i} C_{x_1} \times C_{\pi_2(x_2)} \times \ldots \times C_{\pi_\ell(x_\ell)}, \quad i = 1, 2, \ldots, q^{rs}. \tag{3}$$

Т е о р е м а 1. *Совокупность кодов* (3) *образует разбиение пространства* $\mathbb{F}_q^{n\ell}$, $q > 2$, *на* q*-ичные совершенные коды длины* $n\ell$, *где* $n\ell = \dfrac{q^{rs} - 1}{q - 1}$.

Д о к а з а т е л ь с т в о. Число совершенных кодов длины $n\ell$ в разбиении пространства $\mathbb{F}_q^{n\ell}$ на совершенные коды должно быть равно q^{rs}, что совпадает с числом кодов, построенных в (3). Убедимся, что $C_i^* \cap C_j^* = \varnothing$, где

$$C_j^* = \bigcup_{(x_1,x_2,\ldots,x_\ell)\in D_j} C_{x_1} \times C_{\pi_2(x_2)} \times \ldots \times C_{\pi_\ell(x_\ell)},$$

$i \neq j$, $i, j \in \{1, 2, \ldots, q^{rs}\}$. Поскольку D_i и D_j являются элементами разбиения P_2, то $D_i \cap D_j = \varnothing$. Отсюда $x \neq x'$ для любых x, x', таких что $x \in D_i$, $x' \in D_j$. Следовательно, найдется $k \in \{1, 2, \ldots, \ell\}$, такое что $x_k \neq x'_k$, откуда

$$C_{x_k} \cap C_{x'_k} = \varnothing \quad \text{и} \quad C_{\pi_k(x_k)} \cap C_{\pi_k(x'_k)} = \varnothing.$$

А значит, выполняется

$$C_{x_1} \times C_{\pi_2(x_2)} \times \ldots \times C_{\pi_\ell(x_\ell)} \cap C_{x'_1} \times C_{\pi_2(x'_2)} \times \ldots \times C_{\pi_\ell(x'_\ell)} = \varnothing,$$

и как следствие, имеем $C_i^* \cap C_j^* = \varnothing$ для любых i, j из множества $\{1, 2, \ldots, q^{rs}\}$, где $i \neq j$. ▲

Пусть $M_{n,q}^H$ и $\widetilde{M}_{\ell,q}^H$ обозначают числа различных разбиений на q- и q^r-ичные совершенные коды длины n и ℓ соответственно, где верхний индекс H указывает, что разбиения рассматриваются в метрике Хэмминга. Число подстановок π_t в силу произвольности их выбора равно $(\ell - 1)!$. Из конструкции разбиений, приведенной в теореме 1, легко найти нижнюю оценку числа таких разбиений.

С л е д с т в и е 1. *Число различных разбиений пространства* $\mathbb{F}_q^{n\ell}$ *на совершенные коды, полученных конструкцией* (3), *не меньше*

$$M_{n\ell,q}^H \geqslant M_{n,q}^H \widetilde{M}_{\ell,q}^H (\ell - 1)!. \tag{4}$$

В работе [15] была представлена наилучшая дважды экспоненциальная нижняя оценка числа различных разбиений на q-ичные, $q > 2$, совершенные коды длины

$$N = (q^m - 1)/(q^m - 1), \quad q = p^m, \quad m > 1.$$

Конструкция основана на свитчинговом методе так называемых простых i-компонент кода Хэмминга, с использованием латинских квадратов. Было показано, что число различных разбиений пространства $\mathbb{F}_q^N$ на q-ичные совершенные коды при $p \to \infty$ не меньше, чем

$$p^{p^{n(2m-1)+m+1}(1-o(1))}. \tag{5}$$

Множитель в нижней оценке числа различных разбиений, который дают подстановки π_t в конструкции (2), всего лишь экспоненциальный, поскольку

$$(\ell-1)^{\ell-1+\frac{1}{2}}e^{-\ell+1} \leqslant (\ell-1)! \leqslant (\ell-1)^{\ell-1+\frac{1}{2}}e^{-\ell+2}.$$

Отсюда и из того факта, что с ростом r величина q^r растет существенно быстрее, чем q, величина $M^H_{\ell,q}$ больше, чем $M^H_{n,q}(\ell-1)!$. Таким образом, имеем

$$\widetilde{M}^H_{n\ell,q} > M^H_{n,q}(\ell-1)!.$$

Очевидно, что оценка (4) с использованием (5) для $\widetilde{M}^H_{\ell,q}$ является дважды экспоненциальной по числу ℓmr, где ℓ – длина q^r-ичных кодов в разбиении P_2, $q = p^m$. Однако, как нетрудно убедиться, она уступает оценке из [15], которая по-прежнему является лучшей на сегодняшний день.

§ 4. Разбиения на совершенные коды Ли

Построение разбиений на совершенные коды Ли основано на конструкции Этциона для совершенных кодов Ли из [1]. Для полноты изложения приведем этот метод построения кодов. Он в свою очередь основан на конструкции (2), благодаря чему легко проследить связь результатов настоящего параграфа с результатами, приведенными в § 3.

Пусть $P_1 = \{C_1, C_2, \ldots, C_{q^r}\}$ – произвольное разбиение множества слов $\mathbb{Z}^n_{\tau(2n+1)}$ на совершенные коды Ли длины $n = \dfrac{q^r-1}{2}$ над алфавитом из $\tau(2n+1)$ символов, где q нечетно и равно p^m, $m > 2$. Напомним, что здесь $q^r = 2n+1$ – размер шара Ли радиуса 1 в $\mathbb{Z}^n_{\tau(2n+1)}$, где τ равно произведению всех простых делителей числа $2n+1$, т.е. $\tau = p$. Минимальное расстояние кода C_i равно 3, а его мощность

$$|C_i| = (\tau(2n+1))^n/(2n+1).$$

Пусть D – произвольный q^r-ичный совершенный код в $\mathbb{F}^\ell_{q^r}$, $q > 2$, где

$$\ell = \frac{q^{rs}-1}{q^r-1}, \quad |D| = q^{r(\ell-s)},$$

т.е. D имеет параметры $(\ell, q^{r(\ell-s)}, 3)_{q^r}$. Подчеркнем, что здесь код D рассматривается в метрике Хэмминга.

Пусть $x = (x_1, x_2, \ldots, x_\ell)$ – произвольное кодовое слово кода D. К кодам C_{x_i}, $i \in \{1, 2, \ldots, \ell\}$, применим произвольные ℓ подстановок π_t на множестве $\{1, 2, \ldots, q^r\}$, где $t \in \{1, 2, \ldots, \ell\}$. Пусть π_t являются подстановками на элементах алфавита кода D. Таким образом, имеем ℓ подстановок, первую из которых без ограничения

общности полагаем тождественной, т.е. $C_{\pi_1(x_1)} = C_{x_1}$. Множество слов

$$C = \bigcup_{x=(x_1,x_2,\ldots,x_\ell)\in D} C_{x_1} \times C_{\pi_2(x_2)} \times \ldots \times C_{\pi_\ell(x_\ell)} \tag{6}$$

является совершенным кодом Ли в алфавите из $\tau(2n+1)$ символов, имеющим мощность

$$|C| = (\tau(2n+1))^{n\ell}/(2n\ell+1) \tag{7}$$

и длину $n\ell$. В отличии от конструкции Этциона [1], где P_1 является разбиением лишь на сдвиги совершенного кода Ли C_1, для построения кода C в (6) были взяты более широкие классы разбиений $\mathbb{Z}^n_{\tau(2n+1)}$ на совершенные коды Ли. Ниже убедимся, что такие нетривиальные разбиения существуют.

Разбиения на совершенные коды Ли построим, в свою очередь, с помощью конструкции (6). Рассмотрим помимо разбиения P_1, упомянутого выше, любое разбиение

$$P_2 = \{D_1, D_2, \ldots, D_{q^{rs}}\}$$

пространства $\mathbb{F}^\ell_{q^r}$ на q^r-ичные совершенные коды D_i длины ℓ, $i \in \{1, 2, \ldots, q^{rs}\}$. Для построения кодов C^*_i, $i \in \{1, 2, \ldots, q^{rs}\}$, применим конструкцию (6) к кодам D_i и разбиению P_1:

$$C^*_i = \bigcup_{x=(x_1,x_2,\ldots,x_\ell)\in D_i} C_{x_1} \times C_{\pi_2(x_2)} \times \ldots \times C_{\pi_\ell(x_\ell)}, \quad i = 1, 2, \ldots, q^{rs}. \tag{8}$$

Т е о р е м а 2. *Совокупность кодов* (8) *образует разбиение множества* $\mathbb{Z}^{n\ell}_{\tau(2n+1)}$ *на совершенные коды Ли длины* $n\ell = (q^{rs}-1)/2$ *над алфавитом из* $\tau(2n+1)$ *символов.*

Д о к а з а т е л ь с т в о. Согласно [1, теорема 14] все коды C^*_i, $i \in \{1, 2, \ldots, q^{rs}\}$, являются совершенными кодами Ли длины $n\ell = (q^{rs}-1)/2$. По построению число таких кодов в каждом разбиении равно q^{rs}. Следовательно, с учетом (7) и того, что $2n\ell + 1 = q^{rs}$, имеем

$$|C^*_i|q^{rs} = \frac{\tau^{n\ell}(2n+1)^{n\ell}}{2n\ell+1} q^{rs} = \tau^{n\ell}(2n+1)^{n\ell} = \left|\mathbb{Z}^{n\ell}_{\tau(2n+1)}\right|,$$

что соответствует необходимому числу совершенных кодов Ли в разбиении множества $\mathbb{Z}^{n\ell}_{\tau(2n+1)}$.

Доказательство того факта, что любые два кода C^*_i и C^*_j, где $i \neq j$, не пересекаются, идентично факту непересечения аналогичных кодов в метрике Хэмминга (см. теорему 1), и поэтому опущено. ▲

Пусть $M^L_{n,q}$ обозначает число различных разбиений длины n на совершенные коды Ли. Подчеркнем, что верхний индекс L указывает на то, что разбиения рассматриваются в метрике Ли. Из конструкции разбиений (8) получаем нижнюю оценку числа разбиений на совершенные коды Ли, аналогичную полученной для q-ичных совершенных кодов в $\mathbb{F}^{n\ell}_q$ (см. следствие 1).

С л е д с т в и е 2. *Для числа различных разбиений множества* $\mathbb{Z}^{n\ell}_{\tau(2n+1)}$ *на совершенные коды Ли длины* $n\ell$, *полученных конструкцией* (8), *справедлива оценка снизу*

$$M^L_{n\ell,q} \geqslant M^L_{n,q} \widetilde{M}^H_{\ell,q} ((q^r)!)^{\ell-1}. \tag{9}$$

Поскольку второй сомножитель $\widetilde{M}^H_{\ell,q}$, $q = p^m$, в (9) отражает число разбиений на q^r-ичные совершенные коды длины ℓ в метрике Хэмминга, которое является дважды экспоненциальным от числа ℓmr (см. конец § 3 настоящей статьи), то и результирующая оценка для разбиений на совершенные коды в метрике Ли является дважды экспоненциальной от числа ℓmr. Подставляя ее в конструкцию (6) для совершенных кодов Ли, получаем число различных совершенных кодов Ли, большее чем в [1], где при оценке числа различных совершенных кодов Ли рассматриваются только тривиальные разбиения на совершенные коды Ли. Поскольку оценка Этциона явно представлена не была (см. [1, раздел V]), уточним ее. Она имеет вид

$$N^L_{n\ell,q} \geqslant N^H_{\ell,q}((q^r)!)^{\ell-1}, \tag{10}$$

где $N^L_{n\ell,q}$ и $N^H_{\ell,q}$ обозначают число различных совершенных кодов Ли и совершенных кодов в $\mathbb{F}^\ell_{q^r}$ в метрике Хэмминга длин $n\ell$ и ℓ соответственно.

Используя для построения совершенных кодов Ли конструкцию (6), получаем оценку снизу

$$N^L_{n\ell,q} \geqslant M^L_{n,q} N^H_{\ell,q}((q^r)!)^{\ell-1},$$

что существенно больше оценки Этциона (10) в силу дважды экспоненциальной оценки от числа nm в (9), вычисленной для $M^L_{n,q}$ – числа различных разбиений длины n на совершенные коды Ли.

§ 5. Построение диаметральных совершенных кодов Ли и разбиений

Настоящий параграф посвящен построению диаметральных совершенных кодов Ли и разбиений на такие коды. Сначала рассмотрим построение кодов, а затем на их основе разовьем конструкцию разбиений. Конструкция диаметральных совершенных кодов Ли основана на идее каскадной конструкции Фелпса 1984 г. для двоичных расширенных совершенных кодов (см. [16]). Отметим, что конструкция из [16] является частным случаем обобщенной каскадной конструкции [12]. При построении разбиений используется предложенная конструкция диаметральных совершенных кодов Ли и подход, развитый в работе [17].

5.1. Построение диаметральных совершенных кодов Ли. Пусть $C^0_1, C^0_2, \ldots, C^0_{2r}$ и $C^1_1, C^1_2, \ldots, C^1_{2r}$ – произвольные разбиения множеств четно- и нечетновесовых слов в множестве слов $\mathbb{Z}^r_s$ на диаметральные совершенные коды Ли длины r с минимальным расстоянием $d_L = 4$ над кольцом $\mathbb{Z}_s$, где s четно. Пусть C^m – произвольный расширенный двоичный совершенный код длины $m = 2^p$ в $\mathbb{F}^m_2$ с минимальным расстоянием $d_H = 4$. Далее код C^m будем называть *базовым кодом*. Для каждого вектора μ из C^m рассмотрим $2r$-ичный МДР-код C_μ с минимальным расстоянием Хэмминга 2 длины m и мощности

$$|C_\mu| = (2r)^{m-1}$$

над алфавитом из $2r$ символов. Отметим, что как и в конструкции из § 3, в данной конструкции для построения диаметральных совершенных кодов используются коды над различными алфавитами и метриками.

Теорема 3. *Множество*

$$\mathcal{C} = \bigcup_{\mu\in C^m} \bigcup_{j\in C_\mu} C^{\mu_1}_{j_1} \times C^{\mu_2}_{j_2} \times \ldots \times C^{\mu_m}_{j_m} \tag{11}$$

является диаметральным совершенным кодом Ли длины $n = mr$ над кольцом $\mathbb{Z}_s$.

Доказательство. Очевидно, что длина кода $\mathcal{C}$ равна $n = mr$. Также, с учетом того, что мощности входящих в построение кодов равны

$$|C_{j_i}^{\mu_i}| = \frac{s^r}{4r}, \quad |C_\mu| = (2r)^{m-1}, \quad |C^m| = 2^{m-\log m-1},$$

несложно вычислить мощность кода:

$$|\mathcal{C}| = |(C_{j_i}^{\mu_i})|^m |C_\mu||C^m| = \Big(\frac{s^r}{4r}\Big)^m (2r)^{m-1} 2^{m-\log m-1} = \frac{s^n}{4n}.$$

Убедимся, что минимальное расстояние Ли в коде $\mathcal{C}$ равно 4.

По построению для двух различных кодовых слов u и u' кода $\mathcal{C}$, таких что $\mu = \mu'$ и $j = j'$, выполняется неравенство $d_L(u, u') \geqslant 4$.

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

$$d_L\big(C_{j_1}^{\mu_1} \times \ldots \times C_{j_m}^{\mu_m}, C_{j'_1}^{\mu'_1} \times \ldots \times C_{j'_m}^{\mu'_m}\big) \geqslant 4. \tag{12}$$

Возможны следующие случаи.

1. Пусть $\mu = \mu'$, $j \neq j'$.

Тогда $d_H(j, j') \geqslant 2$, и найдутся координаты a, b, такие что $j_a \neq j'_a$, $j_b \neq j'_b$. Отсюда, учитывая, что $C_{j_a}^{\mu_a}$ и $C_{j'_a}^{\mu_a}$ одновременно являются четно- или нечетновесовыми диаметральными совершенными кодами Ли (аналогичное верно для кодов $C_{j_b}^{\mu_b}$ и $C_{j'_b}^{\mu_b}$), имеем $d_L(C_{j_a}^{\mu_a}, C_{j'_a}^{\mu_a}) \geqslant 2$ и $d_L(C_{j_b}^{\mu_t}, C_{j'_b}^{\mu_b}) \geqslant 2$. Следовательно,

$$d_L\big(C_{j_1}^{\mu_1} \times \ldots \times C_{j_a}^{\mu_a} \times C_{j_b}^{\mu_b} \times \ldots \times C_{j_m}^{\mu_m}, C_{j'_1}^{\mu_1} \times \ldots \times C_{j'_a}^{\mu_a} \times C_{j'_b}^{\mu_b} \times \ldots \times C_{j'_m}^{\mu_m}\big) \geqslant 4.$$

2. Пусть $\mu \neq \mu'$.

Векторы μ и μ' принадлежат базовому коду C^m, т.е. $d_H(\mu, \mu') \geqslant 4$, и найдутся четыре координаты a, b, a' и b', в которых различаются μ и μ'. Следовательно, имеются четыре пары диаметральных совершенных кодов Ли $C_{j_i}^{\mu_i}$ и $C_{j'_i}^{\mu'_i}$, $i \in \{a, b, a', b'\}$, удовлетворяющие

$$d_L\big(C_{j_i}^{\mu_i}, C_{j'_i}^{\mu'_i}\big) \geqslant 1.$$

Отсюда справедливо (12), и как следствие, минимальное расстояние Ли диаметрального совершенного кода $\mathcal{C}$ длины n равно 4. ▲

Обозначим через $N_{n,s}^L$, $\widetilde{D}_{r,s}^L$, N_m^H и R_m^H, соответственно, число различных диаметральных совершенных кодов Ли длины n над $\mathbb{Z}_s$, разбиений на диаметральные совершенные коды Ли длины r над $\mathbb{Z}_s$, число различных двоичных совершенных кодов и число различных МДР-кодов с кодовым расстоянием 2 длины m; здесь верхние индексы L и H указывают на метрики Ли и Хэмминга. Из конструкции диаметральных кодов, приведенной выше, получаем следующую нижнюю оценку числа диаметральных совершенных кодов Ли.

Следствие 3. *Для числа различных диаметральных совершенных кодов Ли длины n, полученных конструкцией* (11), *справедлива оценка снизу*

$$N_{n,s}^L \geqslant \widetilde{D}_{r,s}^L N_m^H R_m^H. \tag{13}$$

Как второй N_m^H, так и третий R_m^H сомножители в (13) являются дважды экспоненциальными от чисел $m/2$ и $r\log_2(m-2)$ в силу [18] и [19] соответственно. Следовательно, эта нижняя оценка существенно лучше нижней оценки числа различных

диаметральных кодов Этциона длины $2^{t-1}n$ в [1], которая имеет вид

$$\prod_{i=1}^{t}(2^i\frac{n}{2}-1)!^{2^{t-i}} \tag{14}$$

и, как несложно убедиться, не превосходит

$$2^t\cdot 2^{n2^{t-1}\log_2\frac{n}{2}}.$$

Оценку (13) можно дополнительно усилить, используя широкие классы разбиений $\widetilde{D}^L_{2r,s}$, построение которых рассмотрим ниже.

5.2. Построение разбиений на диаметральные совершенные коды Ли. Пусть $C^0_{i,1}$, $C^0_{i,2},\dots,C^0_{i,2r}$ и $C^1_{i,1},C^1_{i,2},\dots,C^1_{i,2r}$, $i=1,\dots,m$, – произвольные разбиения множеств четно- и нечетновесовых слов в множестве $\mathbb{Z}_s^r$ на диаметральные совершенные коды Ли длины r с минимальным расстоянием $d_L=4$ над алфавитом $\mathbb{Z}_s$, где s четно. Как и выше, в этой конструкции будут рассмотрены коды в разных алфавитах.

Рассмотрим два разбиения множеств четно- и нечетновесовых слов в пространстве $\mathbb{F}^m$, $m=2^p$, на смежные классы базового двоичного расширенного совершенного кода C^m длины m с минимальным расстоянием $d_H=4$:

$$C^0_i=C^m+e_i+e_m,\quad C^1_i=C^m+e_i,\quad i=1,2,\dots,m;$$

здесь все кодовые слова кода C^0_i имеют четный вес, а кодовые слова C^1_i – нечетный. Для каждого вектора μ из C^m возьмем $2r$-ичный МДР-код C_μ с кодовым расстоянием 2 длины m и мощности $|C_\mu|=(2r)^{m-1}$ над алфавитом из $2r$ символов.

Кроме того, для каждого вектора μ из базового кода C^m рассмотрим латинский квадрат L_μ порядка $2r$, k-я строка которого равна

$$\big(\ell^\mu_{k,1},\ell^\mu_{k,2},\dots,\ell^\mu_{k,2r}\big),\quad k=1,2,\dots,2r.$$

Латинские квадраты и МДР-коды могут быть взяты как различными, так и совпадающими. Определим коды над $\mathbb{Z}_s$ длины $n=mr$ следующим образом:

$$\mathcal{C}_{i,k}=\bigcup_{\mu\in C^0_i}\bigcup_{j\in C_\mu}C^{\mu_1}_{i,j_1}\times\ldots\times C^{\mu_{m-1}}_{i,j_{m-1}}\times C^{\mu_m}_{i,\ell^\mu_{k,j_m}},\quad i=1,2,\dots,m,\quad k=1,2,\dots,2r, \tag{15}$$

$$\mathcal{D}_{i,k}=\bigcup_{\mu\in C^1_i}\bigcup_{j\in C_\mu}C^{\mu_1}_{i,j_1}\times\ldots\times C^{\mu_{m-1}}_{i,j_{m-1}}\times C^{\mu_m}_{i,\ell^\mu_{k,j_m}},\quad i=1,2,\dots,m,\quad k=1,2,\dots,2r. \tag{16}$$

Идея конструкции, описанной в (15) и (16), перекликается с одним из методов построения разбиений на двоичные расширенные совершенные коды из [17].

Теорема 4. *Совокупность кодов $\mathcal{C}_{i,k}$ и $\mathcal{D}_{i,k}$, где $i=1,2,\dots,m$, $k=1,2,\dots,2r$, из (15), (16) образует разбиение множества $\mathbb{Z}_s^n$, $n=mr$, на диаметральные совершенные коды Ли длины n над алфавитом $\mathbb{Z}_s$.*

Доказательство. Согласно теореме 3 все коды $\mathcal{C}_{i,k}$ и $\mathcal{D}_{i,k}$, $i=1,2,\dots,m$, $k=1,2,\dots,2r$, являются диаметральными совершенными кодами Ли над алфавитом $\mathbb{Z}_s$, имеющими длину $n=mr$. По построению число таких кодов в разбиении

равно $4mr$. Следовательно,

$$|\mathcal{C}_{i,k}| \cdot 4mr = \frac{s^n}{4mr} \cdot 4mr = s^n = |\mathbb{Z}_s^n|,$$

что соответствует необходимому числу диаметральных совершенных кодов Ли в разбиении множества $\mathbb{Z}_s^n$.

Очевидно, что четновесовые коды $\mathcal{C}_{i,k}$ и нечетновесовые $\mathcal{D}_{i',k'}$ не пересекаются. Докажем, что любые два кода $\mathcal{C}_{i,k}$ и $\mathcal{C}_{i',k'}$, где $(i,k) \neq (i',k')$, не пересекаются.

Возможны следующие случаи.

1. Пусть $i \neq i'$. Помимо кода $\mathcal{C}_{i,k}$ рассмотрим код $\mathcal{C}_{i',k}$, где

$$\mathcal{C}_{i',k} = \bigcup_{\mu' \in C_{i'}^0} \bigcup_{j' \in C_{\mu'}} C_{i'j'_1}^{\mu'_1} \times \ldots \times C_{i',j'_{m-1}}^{\mu'_{m-1}} \times C_{i',\ell_{k,j'_m}^{\mu'}}^{\mu'_m}.$$

Так как $i \neq i'$, то $C_i^0 \neq C_{i'}^0$. Отсюда в силу того, что оба кода C_i^0 и $C_{i'}^0$ четновесовые, получаем, что существуют $\mu \in C_i^0$ и $\mu' \in C_{i'}^0$, такие что

$$d_H(\mu,\mu') = d_H\big((\mu_1,\ldots,\mu_m),(\mu'_1,\ldots,\mu'_m)\big) \equiv 0 \pmod 2,$$

т.е. $d_H(\mu,\mu') \geqslant 2$. Значит, найдутся по крайней две координаты a и b, в которых различаются слова μ и μ', т.е. $\mu_a \neq \mu'_a$, $\mu_b \neq \mu'_b$. Следовательно, имеются две пары диаметральных совершенных кодов Ли $(C_{i,j_a}^{\mu_a}, C_{i',j'_a}^{\mu'_a})$ и $(C_{i,j_b}^{\mu_b}, C_{i',j'_b}^{\mu'_b})$, такие что

$$d_L\big(C_{i,j_a}^{\mu_a}, C_{i',j'_a}^{\mu'_a}\big) \geqslant 1 \quad \text{и} \quad d_L\big(C_{i,j_b}^{\mu_b}, C_{i',j'_b}^{\mu'_b}\big) \geqslant 1.$$

Как следствие, имеем

$$\mathcal{C}_{i,k} \cap \mathcal{C}_{i',k} = \varnothing.$$

2. Пусть $i = i'$ и $k \neq k'$. Пусть

$$\mathcal{C}_{i,k'} = \bigcup_{\mu' \in C_i^0} \bigcup_{j' \in C_{\mu'}} C_{ij'_1}^{\mu'_1} \times \ldots \times C_{i,j'_{m-1}}^{\mu'_{m-1}} \times C_{i,\ell_{k',j'_m}^{\mu'}}^{\mu'_m}.$$

При $\mu \neq \mu'$ рассуждения идентичны проведенным в случае 1.

Пусть $\mu = \mu'$ и пересечение кодов $\mathcal{C}_{i,k}$ и $\mathcal{C}_{i,k'}$ непусто. В этом случае найдутся два вектора j и j' в МДР-коде C_μ, такие что они совпадают в первых $m-1$ координатных позициях и выполняется

$$\ell_{k,j_m}^{\mu} = \ell_{k',j'_m}^{\mu}. \tag{17}$$

Так как код C_μ имеет минимальное расстояние 2, то это возможно, только если $j_m = j'_m$. Отсюда и из (17) с учетом того, что L_μ – латинский квадрат, имеем $k = k'$, т.е. $\mathcal{C}_{i,k} = \mathcal{C}_{i,k'}$, противоречие.

Для нечетновесовых кодов $\mathcal{D}_{i,k}$, $i = 1,2,\ldots,m$, $k = 1,2,\ldots,2r$, доказательство аналогично, поэтому оно опущено. Как следствие, получаем разбиение множества слов $\mathbb{Z}_s^n$, $n = mr$, на диаметральные совершенные коды Ли длины n. ▲

Оценим снизу число $\widetilde{D}_{r,s}^L$ разбиений множества слов $\mathbb{Z}_s^r$, $r = m'r'$, на диаметральные совершенные коды Ли длины r. Пусть $\widetilde{D}_{r',s}^L$, $N_{m'}^H$, $R_{m'}^H$ и $N_{2r,\mathrm{Lat}}$ – соответственно, число различных разбиений $\mathbb{Z}_s^{r'}$ на диаметральные совершенные коды Ли длины r' над $\mathbb{Z}_s$, число различных двоичных расширенных совершенных кодов длины m', различных МДР-кодов с кодовым расстоянием 2 длины m' и различных

латинских квадратов порядка $2r$. Здесь, как и ранее, верхние индексы L и H указывают на принадлежность к метрикам Ли и Хэмминга. Из конструкции теоремы 4 имеем следующую нижнюю оценку числа разбиений множества слов $\mathbb{Z}_s^r$, $r = m'r'$, на диаметральные совершенные коды Ли длины r.

Следствие 4. *Для числа разбиений множества слов $\mathbb{Z}_s^r$, $r = m'r'$, на диаметральные совершенные коды Ли длины r справедлива оценка снизу*

$$\widetilde{D}_{r,s}^{L} \geqslant \left(\widetilde{D}_{r',s}^{L}\right)^{2m'} N_{m'}^{H} \left(R_{m'}^{H} N_{2r,\mathrm{Lat}}\right)^{2|C^{m'}|}. \tag{18}$$

Все сомножители в оценке (18) можно оценить снизу, как и в предыдущих параграфах, что, очевидно, при подстановке $\widetilde{D}_{r,s}^{L}$ в формулу (13) еще больше усиливает нижнюю оценку числа $N_{n,s}^{L}$ диаметральных совершенных кодов Ли длины $n = mr$ по сравнению с оценкой, предложенной Этционом в [1].

СПИСОК ЛИТЕРАТУРЫ

1. *Etzion T.* Product Construction for Perfect Lee Codes // IEEE Trans. Inform. Theory. 1994. V. 57. № 11. P. 7473–7481. https://doi.org/10.1109/TIT.2011.2161133
2. *AlBdaiwi B., Horak P., Milazzo L.* Enumerating and Decoding Perfect Linear Lee Codes // Des. Codes Cryptogr. 2009. V. 52. № 2. P. 155–162. https://doi.org/10.1007/s10623-009-9273-3
3. *Horak P., AlBdaiwi B.* Diameter Perfect Lee Codes // IEEE Trans. Inform. Theory. 2012. V. 58. № 8. P. 5490–5499. https://doi.org/10.1109/TIT.2012.2196257
4. *Golomb S.W., Welch L.R.* Perfect Codes in the Lee Metric and the Packing of the Polyominoes // SIAM J. App. Math. 1970. V. 18. № 2. P. 302–317. https://doi.org/10.1137/0118025
5. *Kim D.* Nonexistence of Perfect 2-Error-Correcting Lee Codes in Certain Dimensions // European J. Combin. 2017. V. 63. P. 1–5. https://doi.org/10.1016/j.ejc.2017.01.007
6. *Mesnager S., Tang C., Qi Y.* 2-Correcting Lee Codes: (Quasi)-Perfect Spectral Conditions and Some Constructions // IEEE Trans. Inform. Theory. 2018. V. 64. № 4. P. 3031–3041. https://doi.org/10.1109/TIT.2018.2789921
7. *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки. М.: Связь, 1979.
8. *Delsarte P.* An Algebraic Approach to the Association Schemes of Coding Theory // Philips Res. Rep. Suppl. 1973. № 10 (97 pp.).
9. *Ahlswede R., Aydinian H.K., Khachatrian L.H.* On Perfect Codes and Related Concepts // Des. Codes Cryptogr. 2001. V. 22. № 3. P. 221–237. https://doi.org/10.1023/A:1008394205999
10. *Bespalov E.* On the Non-existence of Extended 1-Perfect Codes and MDS Codes // J. Combin. Theory Ser. A. 2022. V. 189. Article ID 105607 (11 pp.). https://doi.org/10.1016/j.jcta.2022.105607
11. *Cohen G., Honkala I., Listyn S., Lobstein A.* Covering Codes. Amsterdam: Elsevier, 1997.
12. *Zinoviev V.A.* On Generalized Concatenated Codes // Topics in Information Theory (Proc. 2nd Colloq. on Information Theory. Keszthely, Hungary. August 25–29, 1975). Colloq. Math. Soc. János Bolyai. V. 16. Amsterdam: North-Holland, 1977. P. 587–592.
13. *Зиновьев В.А.* Комбинаторные методы построения и анализа нелинейных корректирующих кодов: Дисс. ... докт. ф.-м.н.: 01.01.09. М., 1987.
14. *Соловьева Ф.И.* Класс двоичных плотно упакованных кодов, порождаемых q-ичными кодами // Методы дискретного анализа в изучении булевых функций и графов. Вып. 48. Новосибирск: Ин-т матем. СО АН СССР, 1989. С. 70–72.
15. *Соловьева Ф.И., Лось А.В.* О построении разбиений F_q^N на совершенные q-значные коды // Дискретн. анализ и исслед. опер. 2009. V. 16. № 3. P. 63–73. http://mi.mathnet.ru/da574

16. *Phelps K.T.* A General Product Construction for Error Correcting Codes // SIAM J. Algebr. Discrete Methods. 1984. V. 5. № 2. P. 224–228. https://doi.org/10.1137/0605023
17. *Avgustinovich S.V., Lobstein A.C., Soloveva F.I.* Intersection Matrices for Partitions by Binary Perfect Codes // IEEE Trans. Inform. Theory. 2001. V. 47. № 4. P. 1621–1624. https://doi.org/10.1109/18.923749
18. *Krotov D.S., Avgustinovich S.V.* On the Number of 1-Perfect Binary Codes: A Lower Bound // IEEE Trans. Inform. Theory. 2008. V. 54. № 4. P. 1760–1765. https://doi.org/10.1109/TIT.2008.917692
19. *Потапов В.Н., Кротов Д.С.* О числе n-арных квазигрупп конечного порядка // Дискретная математика. 2012. Т. 24. № 1. С. 60–69. https://doi.org/10.4213/dm1172

Соловъева Фаина Ивановна
(15.08.1952 – 09.08.2022)

Поступила в редакцию
06.06.2022
После доработки
06.06.2022
Принята к публикации
24.08.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391 : 519.24

© 2022 г. **М.В. Бурнашев**

О МИНИМАКСНОМ ОБНАРУЖЕНИИ ГАУССОВСКИХ СТОХАСТИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С НЕТОЧНО ИЗВЕСТНЫМИ СРЕДНИМИ И КОВАРИАЦИОННЫМИ МАТРИЦАМИ[1]

Рассматривается задача обнаружения (проверки) гауссовских стохастических последовательностей (сигналов) с неточно известными средними и ковариационными матрицами. Альтернативой являются независимые одинаково распределенные гауссовские случайные величины с нулевыми средними и единичными дисперсиями. При заданной вероятности "*ложной тревоги*" (вероятности ошибки 1-го рода) качество минимаксного обнаружения определяется наилучшей экспонентой "*вероятности пропуска*" (вероятности ошибки 2-го рода) при растущем интервале наблюдений. Исследуется максимальное множество средних и ковариационных матриц (сложная гипотеза), такое что его минимаксную проверку можно заменить проверкой одной конкретной пары среднего и ковариационной матрицы (простая гипотеза) без ухудшения экспоненты обнаружения. В статье полностью описывается это максимальное множество.

Ключевые слова: минимаксная проверка гипотез, вероятность ошибки 1-го рода, вероятность ошибки 2-го рода, экспонента вероятности ошибки, лемма Стейна.

DOI: 10.31857/S0555292322030068, **EDN:** EANVWL

§ 1. Введение и основные результаты

1.1. Постановка задачи. Одна из традиционных задач проверки простых гипотез $\mathcal{H}_0$ и $\mathcal{H}_1$ относительно гауссовского сигнального вектора $\boldsymbol{\eta}_n$ на фоне гауссовского шума $\boldsymbol{\xi}_n$ (т.е. задачи обнаружения сигнала на фоне шума) по наблюдениям $\boldsymbol{y}_n^T = \boldsymbol{y}_n' = (y_1, \ldots, y_n) \in \mathbb{R}^n$ имеет вид

$$\begin{aligned} &\mathcal{H}_0 \colon \boldsymbol{y}_n = \boldsymbol{\xi}_n, \quad \boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n), \\ &\mathcal{H}_1 \colon \boldsymbol{y}_n = \boldsymbol{\eta}_n, \quad \boldsymbol{\eta}_n \sim \mathcal{N}(\boldsymbol{a}_n, \boldsymbol{M}_n), \end{aligned} \tag{1}$$

где выборка $\boldsymbol{\xi}_n^T = (\xi_1, \ldots, \xi_n)$ представляет собой "шум" и состоит из независимых одинаково распределенных гауссовских случайных величин с нулевыми средними и дисперсиями 1, а $\boldsymbol{I}_n$ – единичная ковариационная матрица. Стохастический "сигнал" $\boldsymbol{\eta}_n$ является гауссовским случайным вектором с известным средним $\boldsymbol{a}_n$ и известной ковариационной матрицей $\boldsymbol{M}_n$.

Однако на практике среднее $\boldsymbol{a}_n$ и матрица $\boldsymbol{M}_n$ обычно не известны в точности, и поэтому в реальности модель (1) принимает вид

$$\begin{aligned} &\mathcal{H}_0 \colon \boldsymbol{y}_n = \boldsymbol{\xi}_n, \quad \boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n), \\ &\mathcal{H}_1 \colon \boldsymbol{y}_n = \boldsymbol{\eta}_n, \quad \boldsymbol{\eta}_n \sim \mathcal{N}(\boldsymbol{a}_n, \boldsymbol{M}_n), \quad \boldsymbol{a}_n \in \mathcal{A}_n, \quad \boldsymbol{M}_n \in \mathcal{M}_n, \end{aligned} \tag{2}$$

[1] Исследование выполнено при частичной финансовой поддержке Российского фонда фундаментальных исследований (номер проекта 19-01-00364).

где $\mathcal{A}_n$ – заданное множество возможных средних $\boldsymbol{a}_n$, а $\mathcal{M}_n$ – заданное множество возможных ковариационных матриц $\boldsymbol{M}_n$ (возможно, зависящее от $\boldsymbol{a}_n$). Будем также обозначать для удобства

$$\boldsymbol{B}_n = (\boldsymbol{b}_n, \boldsymbol{V}_n), \quad \boldsymbol{b}_n \in \mathcal{A}_n, \quad \boldsymbol{V}_n \in \mathcal{M}_n, \quad \mathcal{F}_n = \{\boldsymbol{B}_n\} = (\mathcal{A}_n, \mathcal{M}_n).$$

Далее для модели (2) рассматривается задача минимаксной проверки [1–3] простой гипотезы $\mathcal{H}_0$ против сложной альтернативы $\mathcal{H}_1$ по наблюдениям $\boldsymbol{y}_n^T = \boldsymbol{y}_n' = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Если для принятия решения в пользу $\mathcal{H}_0$ выбирается область $\mathcal{D} \in \mathbb{R}^n$, такая что

$$\boldsymbol{y}_n \in \mathcal{D} \;\Rightarrow\; \mathcal{H}_0, \quad \boldsymbol{y}_n \notin \mathcal{D} \;\Rightarrow\; \mathcal{H}_1, \tag{3}$$

то тогда вероятности ошибки 1-го рода ("*ложной тревоги*") $\alpha(\mathcal{D})$ и 2-го рода ("*вероятность пропуска*") $\beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n)$ определяются, соответственно, формулами

$$\alpha(\mathcal{D}) = \mathbf{P}(\boldsymbol{y}_n \notin \mathcal{D} \,|\, \mathcal{H}_0) \tag{4}$$

и

$$\beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n) = \sup_{\boldsymbol{a}_n \in \mathcal{A}_n} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \mathbf{P}(\boldsymbol{y}_n \in \mathcal{D} \,|\, \boldsymbol{M}_n, \boldsymbol{a}_n). \tag{5}$$

Нас интересует минимально возможная вероятность $\beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n)$ ошибки 2-го рода (см. (4) и (5)) при заданной вероятности ошибки 1-го рода α, $0 < \alpha < 1$:

$$\beta(\alpha, \mathcal{A}_n, \mathcal{M}_n) = \inf_{\mathcal{D}:\, \alpha(\mathcal{D}) \leqslant \alpha} \beta(\mathcal{D}, \mathcal{A}_n, \mathcal{M}_n), \tag{6}$$

и соответствующая оптимальная область принятия решения $\mathcal{D}(\alpha)$ из (3).

В статье рассматривается случай, когда величина α фиксирована (или медленно убывает при $n \to \infty$). Этот случай иногда называют задачей Неймана – Пирсона минимаксного различения гипотез. В этом случае ошибки 1-го и 2-го рода влекут очень разные потери для статистика, и он в основном хочет минимизировать вероятность ошибки 2-го рода $\beta = \mathbf{P}\{\mathcal{H}_0 \,|\, \mathcal{H}_1\}$. Этот случай очень популярен в разных приложениях (см., например, работу [4] и библиографию в ней).

Для заданных среднего $\boldsymbol{a}_n$, матрицы $\boldsymbol{M}_n$ и величины α через $\beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ обозначим минимально возможную вероятность ошибки 2-го рода (см. (6)). Соответствующая оптимальная область принятия решения $\mathcal{D}(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ описывается леммой Неймана – Пирсона [1, 2]. Ясно, что

$$\sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) \leqslant \beta(\alpha, \boldsymbol{a}_n, \mathcal{M}_n). \tag{7}$$

Также для фиксированного α и заданных множеств $\mathcal{A}_n, \mathcal{M}_n$ обозначим через $\beta(\alpha, \mathcal{A}_n, \mathcal{M}_n)$ минимально возможную вероятность ошибки 2-го рода (см. (6)). Тогда аналогично (7) имеем

$$\sup_{\boldsymbol{a}_n \in \mathcal{A}_n} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) \leqslant \beta(\alpha, \mathcal{A}_n, \mathcal{M}_n). \tag{8}$$

Во многих практических случаях величина $\beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ убывает экспоненциально по $n \to \infty$. Поэтому естественно (во всяком случае, проще и продуктивнее) исследовать соответствующие экспоненты $n^{-1} \ln \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ и $n^{-1} \ln \beta(\alpha, \mathcal{A}_n, \mathcal{M}_n)$ при $n \to \infty$ (некоторые результаты относительно равенства в (7) имеются в [5]).

В данной статье исследуются множества $\mathcal{F}_n = (\mathcal{A}_n, \mathcal{M}_n)$, для которых в соотношении (8) выполняется следующее асимптотическое равенство:

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\alpha, \mathcal{A}_n, \mathcal{M}_n) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n). \tag{9}$$

Мотивация исследования минимаксной проверки гипотез (обнаружения сигналов) подробно описана в [1–4]. Если для заданных множеств средних $\mathcal{A}_n$ и матриц $\mathcal{M}_n$ выполняется соотношение (9), то тогда можно заменить (без асимптотических потерь) все множество $\mathcal{F}_n$ одной конкретной парой $(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Напомним, что оптимальный тест для конкретной пары $(\boldsymbol{a}_n, \boldsymbol{M}_n)$ описывается леммой Неймана – Пирсона и сводится к простому тесту отношения правдоподобия (LR-тесту). В противном случае (без соотношения (9)) оптимальным минимаксным тестом является значительно более сложный байесовский тест относительно *наименее благоприятного* априорного распределения на множестве $\mathcal{F}_n$. Поэтому естественно исследовать, когда заданное множество $\mathcal{F}_n$ можно заменить конкретной парой $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$. Однако технически удобнее рассмотреть эквивалентную задачу: для заданной пары $\boldsymbol{F}_n$ найти максимальное множество пар $\mathcal{F}_n(\boldsymbol{F}_n)$, которое может быть заменено парой $\boldsymbol{F}_n$. Эта задача в основном и рассматривается в настоящей статье.

Замечание 1. Модели (1) и (2) можно свести к эквивалентным моделям с диагональной матрицей $\boldsymbol{M}_n$. Действительно, так как $\boldsymbol{M}_n$ – ковариационная матрица (т.е. симметричная и положительно определенная), то существуют ортогональная матрица $\boldsymbol{T}_n$ и диагональная матрица $\boldsymbol{\Lambda}_n$, такие что $\boldsymbol{M}_n = \boldsymbol{T}_n \boldsymbol{\Lambda}_n \boldsymbol{T}_n'$ (см. [6, §§ 4.7–4.9; 7, теорема 4.1.5]). При этом диагональная матрица $\boldsymbol{\Lambda}_n = \boldsymbol{T}_n' \boldsymbol{M}_n \boldsymbol{T}_n$ состоит из собственных значений $\{\lambda_i\}$ матрицы $\boldsymbol{M}_n$. Отметим также, что для любой ортогональной матрицы $\boldsymbol{T}_n$ вектор $\boldsymbol{T}_n' \boldsymbol{\xi}_n$ имеет то же самое распределение, что и сам вектор $\boldsymbol{\xi}_n$ (для простой гипотезы $\mathcal{H}_0$ в (2)). Поэтому, умножая обе части (2) на $\boldsymbol{T}_n'$, можно свести модель (2) к эквивалентной модели с диагональной матрицей $\boldsymbol{M}_n$.

О п р е д е л е н и е 1. Для фиксированного α и заданной последовательности пар $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$ обозначим через $\mathcal{F}_0(\boldsymbol{F}_n)$ последовательность наибольших множеств пар, таких что равенство (9) принимает вид

$$\lim_{n\to\infty} \frac{1}{n} \ln \beta(\mathcal{F}_0(\boldsymbol{F}_n)) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{F}_n). \tag{10}$$

Ясно, что $\boldsymbol{F}_n \in \mathcal{F}_0(\boldsymbol{F}_n)$.

Другими словами, для заданной вероятности ошибки 1-го рода α последовательность $\mathcal{F}_0(\boldsymbol{F}_n)$ представляет собой наибольшее множество пар, которое может быть заменено (без асимптотических потерь для $\beta(\mathcal{F}_0(\boldsymbol{F}_n))$) одной парой $\boldsymbol{F}_n$. Далее (теорема 1) описывается наибольшее множество $\mathcal{F}_0(\boldsymbol{F}_n)$, удовлетворяющее (10). Это обобщает аналогичный результат из [8], где рассматривался случай $\boldsymbol{a}_n = \boldsymbol{0}_n$. Это также усиливает аналогичный результат из [4], где для множества $\mathcal{F}_0(\boldsymbol{0}_n, \boldsymbol{M}_n)$ были получены некоторые оценки снизу.

Удобно сначала исследовать максимальные множества $\mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{F}_n)$, аналогичные $\mathcal{F}_0(\boldsymbol{F}_n)$, но которые возникают, если используется LR-детектор (см. определение 2). Будет показано, что $\mathcal{F}_0(\boldsymbol{F}_n) = \mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{F}_n)$, т.е. LR-детектор является асимптотически оптимальным.

В моделях (1) и (2) обозначим через $\mathbf{P}_{\boldsymbol{I}_n}$ распределение величины $\boldsymbol{y}_n = \boldsymbol{\xi}_n$, где $\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$. Аналогично обозначим через $\mathbf{Q}_{\boldsymbol{F}_n}$, $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$, распределение величины $\boldsymbol{y}_n = \boldsymbol{\eta}_n$, где $\boldsymbol{\eta}_n \sim \mathcal{N}(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Обозначим также через $p_{\boldsymbol{I}_n}(\boldsymbol{y}_n)$ и $p_{\boldsymbol{F}_n}(\boldsymbol{y}_n)$, $\boldsymbol{y}_n \in \mathbb{R}^n$, соответствующие плотности распределения вероятностей. Для $(n \times n)$-мат-

рицы $\boldsymbol{M}_n$ будем обозначать $|\boldsymbol{M}_n| = \det \boldsymbol{M}_n$. Заметим, что если $|\boldsymbol{M}_n| \neq 0$, то

$$\begin{aligned} &\ln p_{\boldsymbol{I}_n}(\boldsymbol{y}_n) = -\frac{1}{2}[n\ln(2\pi) + (\boldsymbol{y}_n, \boldsymbol{y}_n)], \quad \boldsymbol{y}_n \in \mathbb{R}^n, \\ &\ln p_{\boldsymbol{F}_n}(\boldsymbol{y}_n) = -\frac{1}{2}\big[n\ln(2\pi) + \ln|\boldsymbol{M}_n| + (\boldsymbol{y}_n - \boldsymbol{a}_n, \boldsymbol{M}_n^{-1}(\boldsymbol{y}_n - \boldsymbol{a}_n))\big]. \end{aligned} \tag{11}$$

При $|\boldsymbol{M}_n| \neq 0$ введем также логарифм отношения правдоподобия (см. (11))

$$\begin{aligned} &r_{\boldsymbol{F}_n}(\boldsymbol{y}_n) = \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{F}_n}}(\boldsymbol{y}) = \\ &= \frac{1}{2}\Big[\ln|\boldsymbol{M}_n| + (\boldsymbol{y}, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{y}) - 2(\boldsymbol{y}, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\Big]. \end{aligned} \tag{12}$$

Рассмотрим сначала LR-детекторы. Введем соответствующие области принятия решения $\mathcal{D}_{\mathrm{LR}}(\boldsymbol{F}_n, \alpha)$ в пользу гипотезы $\mathcal{H}_0$ (т.е. в пользу матрицы $\boldsymbol{I}_n$), когда проверяются простые гипотезы $\boldsymbol{I}_n$ и $\boldsymbol{F}_n$:

$$\mathcal{D}_{\mathrm{LR}}(\boldsymbol{F}_n, \alpha) = \{\boldsymbol{y}_n \in \mathbb{R}^n : \, r_{\boldsymbol{F}_n}(\boldsymbol{y}_n) \geqslant \gamma\}, \tag{13}$$

где γ такое, что (см. (12))

$$\begin{aligned} &\alpha = \mathbf{P}_{\boldsymbol{I}_n}\{\mathcal{D}^c_{\mathrm{LR}}(\boldsymbol{F}_n, \alpha)\} = \mathbf{P}_{\boldsymbol{I}_n}\{r_{\boldsymbol{F}_n}(\boldsymbol{\xi}_n) \leqslant \gamma\} = \mathbf{P}_{\boldsymbol{I}_n}\Big\{[\ln|\boldsymbol{M}_n| + \\ &+ (\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{\xi}_n) - 2(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)] \leqslant 2\gamma\Big\}. \end{aligned} \tag{14}$$

О п р е д е л е н и е 2. Для фиксированного α и заданной последовательности пар $\boldsymbol{F}_n = (\boldsymbol{a}_n, \boldsymbol{M}_n)$ обозначим через $\mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{F}_n)$ последовательность максимальных множеств пар $(\boldsymbol{b}_n, \boldsymbol{V}_n)$, таких что

$$\lim_{n\to\infty} \frac{1}{n} \ln \sup_{(\boldsymbol{b}_n, \boldsymbol{V}_n) \in \mathcal{F}_n^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)} \beta(\boldsymbol{b}_n, \boldsymbol{V}_n) = \lim_{n\to\infty} \frac{1}{n} \ln \beta(\boldsymbol{a}_n, \boldsymbol{M}_n), \tag{15}$$

при условии, что используются решающие области $\mathcal{D}_{\mathrm{LR}}(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$.

Ниже в теореме 2 описывается множество $\mathcal{F}_0^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ для модели (2).

Нам понадобится также следующее определение [9].

О п р е д е л е н и е 3. Для вероятностных мер $\mathbf{P}$ и $\mathbf{Q}$ на измеримом пространстве $(\mathcal{X}, \mathcal{B})$ введем функцию (расстояние (или дивергенция) Кульбака – Лейблера для мер $\mathbf{P}$ и $\mathbf{Q}$)

$$D(\mathbf{P} \,\|\, \mathbf{Q}) = \mathbf{E}_{\mathbf{P}} \ln \frac{d\,\mathbf{P}}{d\,\mathbf{Q}}(\boldsymbol{x}) \geqslant 0, \tag{16}$$

где математическое ожидание берется по мере $\mathbf{P}$.

С помощью формул (11) и (16) имеем

$$\begin{aligned} &D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n, \boldsymbol{M}_n}) = \mathbf{E}_{\boldsymbol{\xi}_n} \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{\xi}_n) = \\ &= \frac{1}{2}\Big[\ln|\boldsymbol{M}_n| + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + \mathbf{E}_{\boldsymbol{\xi}_n}(\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I}_n)\boldsymbol{\xi}_n)\Big] = \\ &= \frac{1}{2}\left[\sum_{i=1}^{n}(\ln\lambda_i + \frac{1}{\lambda_i} - 1) + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\right], \end{aligned} \tag{17}$$

где $\{\lambda_1, \ldots, \lambda_n\}$ – собственные значения (все положительные) ковариационной матрицы $\boldsymbol{M}_n$, а $\boldsymbol{a}_n = (a_1, \ldots, a_n)$.

1.2. Предположения. В модели (2) обозначим через $\lambda_1(\boldsymbol{M}_n), \ldots, \lambda_n(\boldsymbol{M}_n)$ собственные значения (все положительные) ковариационной матрицы $\boldsymbol{M}_n$. Предположим, что выполняются следующие предположения:

I. Для всех ковариационных матриц $\boldsymbol{M}_n \in \mathcal{M}_n(\boldsymbol{M}_n)$ существует предел (см. (17))

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \left(\ln \lambda_i(\boldsymbol{M}_n) + \frac{1}{\lambda_i(\boldsymbol{M}_n)} - 1 \right) \tag{18}$$

(заметим, что $\ln z + 1/z - 1 \geqslant 0$, $z > 0$).

II. Для какого-нибудь $\delta > 0$ имеем

$$\lim_{n\to\infty} \frac{1}{n} \sup_{\boldsymbol{M}_n \in \mathcal{M}_n} \sum_{i=1}^{n} \left| \frac{1}{\lambda_i(\boldsymbol{M}_n)} - 1 \right|^{1+\delta} < \infty. \tag{19}$$

1.3. Основные результаты. Сделаем сначала важное пояснение.

Замечание 2. При описании максимальных множеств $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ имеется следующая техническая трудность. Соотношение (9) имеет асимптотический (при $n \to \infty$) характер. Поэтому и максимальные множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ можно описать только асимптотически (при $n \to \infty$). Для этого удобнее всего описать наиболее простую последовательность множеств, которые в пределе дают максимальные множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

В статье для $(n \times n)$-матрицы $\boldsymbol{A}_n$ будем обозначать $|\boldsymbol{A}_n| = \det \boldsymbol{A}_n$. Через $(\boldsymbol{x}, \boldsymbol{y})$ будем обозначать скалярное произведение векторов $\boldsymbol{x}, \boldsymbol{y}$. Будем писать $\boldsymbol{A}_n > \boldsymbol{0}$, если матрица $\boldsymbol{A}_n$ положительно определена.

Пусть $\mathcal{C}_n$ – множество всех ковариационных (т.е. симметричных и положительно определенных) $(n \times n)$-матриц в $\mathbb{R}^n$. Для любых $\boldsymbol{M}_n, \boldsymbol{V}_n \in \mathcal{C}_n$, и $\boldsymbol{a}_n, \boldsymbol{b}_n \in \mathbb{R}^n$ определим функцию

$$f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) = \frac{|\boldsymbol{M}_n| e^{-K}}{|\boldsymbol{V}_n||\boldsymbol{B}_n|}, \tag{20}$$

где

$$\begin{aligned} &\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}, \quad \boldsymbol{d} = \boldsymbol{B}_n^{-1}(\boldsymbol{V}_n^{-1}\boldsymbol{b}_n - \boldsymbol{M}_n^{-1}\boldsymbol{a}_n), \\ &K = (\boldsymbol{b}_n, \boldsymbol{V}_n^{-1}\boldsymbol{b}_n) - (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) - (\boldsymbol{d}, \boldsymbol{B}_n\boldsymbol{d}). \end{aligned} \tag{21}$$

Для последовательности пар $(\boldsymbol{a}_n, \boldsymbol{M}_n)$ введем следующую последовательность множеств пар $(\boldsymbol{b}_n, \boldsymbol{V}_n)$:

$$\begin{aligned} \mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n) &= \left\{ (\boldsymbol{b}_n, \boldsymbol{V}_n) : \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)} \right\} = \\ &= \left\{ (\boldsymbol{b}_n, \boldsymbol{V}_n) : \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1} > \boldsymbol{0},\ f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant e^{o(n)} \right\}, \end{aligned} \tag{22}$$

где функция $f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ определена в (20).

Следующая теорема является главным результатом статьи и описывает множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ и $\mathcal{F}^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ из (10) и (15) соответственно.

Теорема 1. *Если выполняются предположения* (18), (19), *то при* $n \to \infty$

$$\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n) = \mathcal{F}^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n) = \mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n), \tag{23}$$

где равенства понимаются в смысле замечания 2.

Замечание 3. Ясно, что $(\boldsymbol{a}_n, \boldsymbol{M}_n) \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Также множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ и $\mathcal{F}^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ являются выпуклыми по $\boldsymbol{b}_n, \boldsymbol{V}_n$. Действительно, известно [6, § 8.5, теорема 4; 7, теорема 7.6.7], что функция $f(\boldsymbol{A}_n) = \ln|\boldsymbol{A}_n|$ является строго вогнутой на выпуклом множестве всех положительно определенных симметричных матриц в $\mathbb{R}^n$. Поэтому множество $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ является выпуклым, т.е. любые матрицы $\boldsymbol{V}_n^{(1)} \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ и $\boldsymbol{V}_n^{(2)} \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ удовлетворяют условию

$$a\boldsymbol{V}_n^{(1)} + (1-a)\boldsymbol{V}_n^{(2)} \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n) \quad \text{для любого } 0 \leqslant a \leqslant 1.$$

В определенном смысле $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ – это множество $\mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n)$, расширенное на "тонкий слой" с шириной порядка $o(n)$. Другими словами, $\mathcal{F}_0(\boldsymbol{a}_n, \boldsymbol{M}_n)$ можно рассматривать как "ядро" множества $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

Приведем также следующее упрощающее следствие из теоремы 1. Без ограничения общности можно считать, что матрица $\boldsymbol{M}_n$ является диагональной (см. замечание 1) с собственными значениями $\{\lambda_i\}$ (все положительные). Также ограничимся в (23) только диагональными матрицами $\boldsymbol{V}_n$ с положительными собственными значениями $\{\nu_i\}$. Матрица $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ является диагональной с собственными значениями $\{\mu_i\}$:

$$\mu_i = 1 + \frac{1}{\nu_i} - \frac{1}{\lambda_i}, \quad i = 1, \ldots, n. \tag{24}$$

Тогда для $\boldsymbol{a}_n = (a_{1,n}, \ldots, a_{n,n})$, $\boldsymbol{b}_n = (b_{1,n}, \ldots, b_{n,n})$ из (21) имеем

$$K = \sum_{i=1}^{n} \Bigl[\frac{b_{i,n}^2}{\nu_i} - \frac{a_{i,n}^2}{\lambda_i} - \frac{1}{\mu_i}\Bigl(\frac{b_{i,n}}{\nu_i} - \frac{a_{i,n}}{\lambda_i}\Bigr)^2 \Bigr]. \tag{25}$$

Введем выпуклое множество $\mathcal{C}_{\mathrm{diag},n}$ диагональных положительно определенных матриц $\boldsymbol{V}_n$:

$$\mathcal{C}_{\mathrm{diag},n} = \{\boldsymbol{V}_n \in \mathcal{C}_n : \ \boldsymbol{V}_n > \boldsymbol{0} \ \text{ и } \ \boldsymbol{V}_n \text{ – диагональная матрица}\}.$$

Если $\boldsymbol{M}_n, \boldsymbol{V}_n \in \mathcal{C}_{\mathrm{diag},n}$, то функция $f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ из (20) принимает вид

$$f^{(0)}_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) = e^{-K} \prod_{i=1}^{n} \Bigl(\frac{\lambda_i}{\nu_i \mu_i}\Bigr), \tag{26}$$

где $\{\mu_i\}$ определены в (24), а K определено в (25). Предполагается также, что $\mu_i > 0$, $i = 1, \ldots, n$.

Для последовательности пар $(\boldsymbol{a}_n, \boldsymbol{M}_n)$, $\boldsymbol{M}_n \in \mathcal{C}_{\mathrm{diag},n}$, введем следующее множество пар $(\boldsymbol{b}_n, \boldsymbol{V}_n)$, $\boldsymbol{V}_n \in \mathcal{C}_{\mathrm{diag},n}$:

$$\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n) = \\ = \Bigl\{(\boldsymbol{b}_n, \boldsymbol{V}_n) : \ 1 + 1/\nu_i - 1/\lambda_i > 0, \ i = 1, \ldots, n, \ \ln f^{(0)}_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant o(n)\Bigr\}, \tag{27}$$

где функция $f^{(0)}_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ определена в (26).

Тогда справедлива следующая "внутренняя граница" для $\mathcal{M}(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

Теорема 2. *Если выполняются предположения* (18), (19), *то множество* $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ *содержит множество* $\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n)$:

$$\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n) \subseteq \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n), \quad \boldsymbol{M}_n \in \mathcal{C}_{\mathrm{diag},n}, \tag{28}$$

где множество $\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ *определено в* (27).

Множество $\mathcal{V}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ выпукло по $\boldsymbol{V}_n$ (см. замечание 3).

Далее в § 2 приводится вспомогательная теорема 3. В § 3 доказывается теорема 1, а в § 4 в качестве примеров рассматриваются некоторые частные случаи задачи.

§ 2. Вспомогательная теорема

В моделях (1), (2) рассмотрим сначала проверку простых гипотез: пара $(\mathbf{0}_n, \boldsymbol{I}_n)$ против пары $(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Обозначим

$$D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) = D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n, \boldsymbol{M}_n}).$$

Следующая теорема является основным вспомогательным результатом, используемым в настоящей статье. Ее доказательство следует доказательству теоремы 3 в [8]. Более общий результат содержится в [10].

Теорема 3. *Для минимально возможного $\beta(\alpha)$, $0 < \alpha < 1$, справедливы границы*

$$\ln \beta(\alpha) \geqslant -\frac{D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) + h(\alpha)}{1-\alpha}, \quad h(\alpha) = -\alpha \ln \alpha - (1-\alpha)\ln(1-\alpha), \tag{29}$$

и

$$\ln \beta(\alpha) \leqslant -D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) + \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n), \tag{30}$$

где $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ определяется соотношением

$$\mathbf{P}_{\boldsymbol{I}_n}\left\{ \ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) - \mu_0 \right\} = \alpha. \tag{31}$$

Отметим, что обе границы (29) и (30) являются чисто аналитическими соотношениями, без каких-либо предельных операций. Нижняя граница (29) и верхняя граница (30) близки друг к другу, если величина $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ значительно меньше, чем $D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n)$ (которая обычно имеет порядок n).

Следующий результат дает для величины $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ из (31) оценку сверху порядка $n^{1/p}$, $p > 1$. Ее доказательство (см. Приложение) следует доказательству леммы 1 в [8].

Лемма 1. *Для $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ из (30) справедлива оценка сверху (см. (19))*

$$\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) \leqslant \left(\frac{24}{\alpha} \sum_{i=1}^{n} \left| \frac{1}{\lambda_i(\boldsymbol{M}_n)} - 1 \right|^p \right)^{1/p} + 3 \big\| \boldsymbol{M}_n^{-1} \boldsymbol{a}_n \big\| \sqrt{\ln(1/\alpha)}. \tag{32}$$

§ 3. Доказательство теоремы 1

Так как $\mathcal{F}_n^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n) \subseteq \mathcal{F}_n(\boldsymbol{a}_n, \boldsymbol{M}_n)$, то для доказательства теоремы 1 достаточно получить "внутреннюю границу" для $\mathcal{F}_n^{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n)$, а затем получить аналогичную "внешнюю границу" для $\mathcal{F}_n(\boldsymbol{a}_n, \boldsymbol{M}_n)$.

3.1. "Внутренняя граница" для $\boldsymbol{\mathcal{F}_n^{\mathrm{LR}}(a_n, M_n)}$. Оценим сначала сверху величину $\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n)$. Для этого в модели (2) рассмотрим проверку простой гипотезы $(\mathbf{0}_n, \boldsymbol{I}_n)$ против простой альтернативы $(\boldsymbol{a}_n, \boldsymbol{M}_n)$, когда $\boldsymbol{a}_n$ известна. Используем оптимальный LR-тест с решающей областью $\mathcal{D}_{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n, \alpha) = \mathcal{A}_{\mu_0}$ в пользу $(\mathbf{0}_n, \boldsymbol{I}_n)$ (см. (13), (14)), где $\mu_0 = \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) > 0$ определено в (31). Рассмотрим какую-либо другую пару $(\boldsymbol{b}_n, \boldsymbol{V}_n)$ и оценим вероятность ошибки 2-го рода $\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n)$ при

условии, что используется решающая область $\mathcal{A}_{\mu_0}$. Тогда

$$\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n) = \int\limits_{\mathcal{A}_{\mu_0}} p_{\boldsymbol{b}_n, \boldsymbol{V}_n}(\boldsymbol{x}) d\boldsymbol{x} = \int\limits_{\mathcal{A}_{\mu_0}} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) \frac{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}{p_{\boldsymbol{I}_n}}(\boldsymbol{x}) p_{\boldsymbol{I}_n}(\boldsymbol{x}) d\boldsymbol{x} =$$

$$= e^{-D(\boldsymbol{I}_n \| \boldsymbol{a}_n, \boldsymbol{M}_n) + \mu_1} \int\limits_{\mathcal{A}_{\mu_0}} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) p_{\boldsymbol{I}_n}(\boldsymbol{x}) d\boldsymbol{x} \leqslant$$

$$\leqslant \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) e^{\mu_0} \, \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}), \tag{33}$$

где $0 \leqslant \mu_1 \leqslant \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$. В силу предположения (19) и оценки (32) имеем

$$\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) = O(n^{1/(1+\delta)}) = o(n), \quad n \to \infty. \tag{34}$$

Поэтому если

$$\sup_{(\boldsymbol{b}_n, \boldsymbol{V}_n) \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)} \mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \quad n \to \infty, \tag{35}$$

то в силу (33)–(35) при $n \to \infty$

$$\sup_{(\boldsymbol{b}_n, \boldsymbol{V}_n) \in \mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)} \ln \beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant \ln \beta(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n) + o(n). \tag{36}$$

3.2. "Внешняя граница" для $\mathcal{F}_n(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Получим теперь аналогичную оценку снизу для $\beta(\alpha, \boldsymbol{b}_n, \boldsymbol{V}_n)$. Рассмотрим сначала проверку простой гипотезы $(\mathbf{0}_n, \boldsymbol{I}_n)$ против простой альтернативы $(\boldsymbol{a}_n, \boldsymbol{M}_n)$. Используем для этого оптимальный LR-тест с решающей областью $\mathcal{D}_{\mathrm{LR}}(\boldsymbol{a}_n, \boldsymbol{M}_n, \alpha) = \mathcal{A}_{\mu_0}$ в пользу $(\mathbf{0}_n, \boldsymbol{I}_n)$ (см. (13), (14)). Тогда, обозначая $p = p_{\boldsymbol{I}_n}$ и $q = p_{\boldsymbol{a}_n, \boldsymbol{M}_n}$, для вероятностей ошибок имеем

$$\alpha = \mathbf{P}_{\boldsymbol{I}_n}(\mathcal{A}_{\mu_0}), \quad \beta_{\boldsymbol{a}_n, \boldsymbol{M}_n} = \int\limits_{\mathcal{A}_{\mu_0}} q(\boldsymbol{x}) \, d\boldsymbol{x} = \beta(\alpha). \tag{37}$$

Рассмотрим какую-либо другую пару $(\boldsymbol{b}_n, \boldsymbol{V}_n)$. Пусть $\mathcal{D} \in \mathbb{R}^n$ – какая-то решающая область в пользу $(\mathbf{0}_n, \boldsymbol{I}_n)$, а $\beta_{\boldsymbol{b}_n, \boldsymbol{V}_n}(\mathcal{D})$ и $\alpha = \alpha(\mathcal{D})$ – соответствующие вероятности ошибок. Тогда, обозначая $q_1 = p_{\boldsymbol{b}_n, \boldsymbol{V}_n}$, для вероятности ошибки 2-го рода $\beta_{\boldsymbol{b}_n, \boldsymbol{V}_n}(\mathcal{D})$ (см. (37)) должно выполняться

$$\beta_{\boldsymbol{b}_n, \boldsymbol{V}_n}(\mathcal{D}) = \int\limits_{\mathcal{D}} q_1(\boldsymbol{x}) \, d\boldsymbol{x} \leqslant \beta(\alpha) e^{o(n)}, \quad \alpha = \mathbf{P}_{\boldsymbol{I}_n}(\mathcal{D}^c). \tag{38}$$

Для некоторого δ, $0 \leqslant \delta \leqslant 1$, рассмотрим также плотность вероятности

$$q_\delta(\boldsymbol{x}) = (1 - \delta) q(\boldsymbol{x}) + \delta q_1(\boldsymbol{x}) \tag{39}$$

и соответствующую величину β_δ для нее:

$$\beta_\delta = \int\limits_{\mathcal{D}} q_\delta(\boldsymbol{x}) \, d\boldsymbol{x} = (1 - \delta)\beta(\alpha) + \delta \beta_{\boldsymbol{b}_n, \boldsymbol{V}_n}. \tag{40}$$

В силу (38) и (40) имеем

$$\beta_\delta \leqslant \beta(\alpha)\big(1 - \delta + \delta e^{o(n)}\big). \tag{41}$$

Отметим, что плотность вероятности $q_\delta(\boldsymbol{x})$ соответствует байесовской постановке задачи, когда альтернативная гипотеза $\mathcal{H}_1$ с вероятностью $1-\delta$ совпадает с $(\boldsymbol{a}_n, \boldsymbol{M}_n)$, а с вероятностью δ – с $(\boldsymbol{b}_n, \boldsymbol{V}_n)$. Величина β_δ является соответствующей вероятностью ошибки 2-го рода.

Оценим снизу величину β_δ. Сначала имеем

$$\ln\frac{\beta_\delta}{1-\alpha} = \ln\left[\frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}} p(\boldsymbol{x})\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x}\right] \geqslant \frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} =$$

$$= -\frac{D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x}))}{1-\alpha} - \frac{1}{(1-\alpha)}\int\limits_{\mathcal{D}^c} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x}. \tag{42}$$

Для последнего члена в правой части (42) имеем

$$\int\limits_{\mathcal{D}^c} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} \leqslant \alpha\ln\left[\frac{1}{\alpha}\int\limits_{\mathcal{D}^c} q_\delta(\boldsymbol{x})\,d\boldsymbol{x}\right] = \alpha\ln\frac{1-\beta_\delta}{\alpha} \leqslant \alpha\ln\frac{1}{\alpha}.$$

Поэтому получаем

$$\ln\beta_\delta \geqslant -\frac{D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x})) + h(\alpha)}{1-\alpha}. \tag{43}$$

Рассмотрим величину $D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x}))$ в правой части (43). Обозначая

$$r(\boldsymbol{x}) = \frac{q_1(\boldsymbol{x})}{q(\boldsymbol{x})}, \tag{44}$$

в силу (39) и (44) имеем

$$\frac{q_\delta(\boldsymbol{x})}{q(\boldsymbol{x})} = 1-\delta+\delta\frac{q_1(\boldsymbol{x})}{q(\boldsymbol{x})} = 1-\delta+\delta r(\boldsymbol{x}).$$

Поэтому

$$D(p(\boldsymbol{x})\,\|\,q_\delta(\boldsymbol{x})) = -\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})\ln\frac{q_\delta}{p}(\boldsymbol{x})\,d\boldsymbol{x} = D(p(\boldsymbol{x})\,\|\,q(\boldsymbol{x})) + g(\delta), \tag{45}$$

где

$$g(\delta) = -\int\limits_{\mathbb{R}^n} p(\mathbf{x})\ln[1-\delta+\delta r(\boldsymbol{x})]\,d\boldsymbol{x}. \tag{46}$$

Поэтому в силу (41), (45) и (46) необходимо иметь

$$g(\delta) \geqslant -\ln(1-\delta+\delta e^{o(n)}) \quad \text{для всех } 0<\delta\leqslant 1. \tag{47}$$

Заметим, что так как $\ln\mathbf{E}\,\xi \geqslant \mathbf{E}\ln\xi$, то из (46) имеем

$$g(\delta) \leqslant \ln\int\limits_{\mathbb{R}^n}\frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \quad \text{для всех } 0<\delta\leqslant 1.$$

Поэтому для того чтобы выполнялось (47), необходимо иметь

$$\int\limits_{\mathbb{R}^n}\frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \geqslant \frac{1}{1-\delta+\delta e^{o(n)}}, \quad 0<\delta\leqslant 1. \tag{48}$$

Так как $\int p(\boldsymbol{x})\,d\boldsymbol{x} = 1$, то соотношение (48) эквивалентно условию

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})(r(\boldsymbol{x})-1)}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \leqslant \frac{e^{o(n)}-1}{1-\delta+\delta e^{o(n)}}, \quad 0<\delta\leqslant 1. \tag{49}$$

Заметим, что

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})}{1-\delta+\delta r(\boldsymbol{x})}d\boldsymbol{x} \leqslant \frac{1}{1-\delta}, \quad 0<\delta\leqslant 1.$$

Поэтому для того чтобы выполнялось (49), необходимо по меньшей мере иметь

$$\int\limits_{\mathbb{R}^n} \frac{p(\boldsymbol{x})r(\boldsymbol{x})}{1+\delta r(\boldsymbol{x})}d\boldsymbol{x} \leqslant \frac{e^{o(n)}}{(1-\delta)(1-\delta+\delta e^{o(n)})}. \tag{50}$$

Полагая $\delta\downarrow 0$, получаем из (50) необходимое условие

$$\int\limits_{\mathbb{R}^n} p(\boldsymbol{x})r(\boldsymbol{x})\,d\boldsymbol{x} = \mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) \leqslant e^{o(n)}, \tag{51}$$

что дает "внешнюю границу" для $\mathcal{F}_n(\boldsymbol{a}_n,\boldsymbol{M}_n)$ (см. (23)).

Заметим, что "внутренняя граница" (35), (36) для $\mathcal{F}_n(\boldsymbol{a}_n,\boldsymbol{M}_n)$ совпадает с (51). Поэтому для завершения доказательства теоремы 1 остается выразить условие (51) через матрицы $\boldsymbol{M}_n,\boldsymbol{V}_n$ и средние $\boldsymbol{a}_n,\boldsymbol{b}_n$. Для этого используем следующий результат.

Лемма 2. *Если* $\boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}>\boldsymbol{0}$, *то справедлива формула* (*см.* (20)–(22))

$$\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2}e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}|\boldsymbol{B}_n|^{1/2}} = f^{1/2}_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n,\boldsymbol{V}_n), \tag{52}$$

где функция $f_{\boldsymbol{a}_n,\boldsymbol{M}_n}(\boldsymbol{b}_n,\boldsymbol{V}_n)$ *определена в* (20).

Если матрица $\boldsymbol{I}_n+\boldsymbol{V}_n^{-1}-\boldsymbol{M}_n^{-1}$ *не является положительно определенной, то*

$$\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) = \infty. \tag{53}$$

Доказательство. Обозначая

$$\zeta = \big(\boldsymbol{\xi}_n-\boldsymbol{b}_n,\boldsymbol{V}_n^{-1}(\boldsymbol{\xi}_n-\boldsymbol{b}_n)\big) - \big(\boldsymbol{\xi}_n-\boldsymbol{a}_n,\boldsymbol{M}_n^{-1}(\boldsymbol{\xi}_n-\boldsymbol{a}_n)\big),$$

с помощью (11) получаем

$$\begin{aligned}\mathbf{E}_{\boldsymbol{I}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}) &= \mathbf{E}_{\boldsymbol{\xi}_n}\frac{p_{\boldsymbol{b}_n,\boldsymbol{V}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{\xi}_n) = \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{V}_n|^{1/2}}\,\mathbf{E}_{\boldsymbol{\xi}_n}\,e^{-\zeta/2} = \\ &= \frac{|\boldsymbol{M}_n|^{1/2}}{|\boldsymbol{V}_n|^{1/2}(2\pi)^{n/2}}\int\limits_{\mathbb{R}^n} e^{-\frac{1}{2}\left[(\boldsymbol{x},\boldsymbol{x})+(\boldsymbol{x}-\boldsymbol{b}_n,\boldsymbol{V}_n^{-1}(\boldsymbol{x}-\boldsymbol{b}_n))-(\boldsymbol{x}-\boldsymbol{a}_n,\boldsymbol{M}_n^{-1}(\boldsymbol{x}-\boldsymbol{a}_n))\right]}\,d\boldsymbol{x}.\end{aligned} \tag{54}$$

Заметим, что (см. (54))

$$(\boldsymbol{x},\boldsymbol{x})+(\boldsymbol{x}-\boldsymbol{b},\boldsymbol{V}^{-1}(\boldsymbol{x}-\boldsymbol{b}))-(\boldsymbol{x}-\boldsymbol{a},\boldsymbol{M}^{-1}(\boldsymbol{x}-\boldsymbol{a})) = (\boldsymbol{x}-\boldsymbol{d},\boldsymbol{B}(\boldsymbol{x}-\boldsymbol{d}))+K,$$

где (см. также (21))

$$\boldsymbol{B} = \boldsymbol{I} + \boldsymbol{V}^{-1} - \boldsymbol{M}^{-1}, \quad \boldsymbol{d} = \boldsymbol{B}^{-1}(\boldsymbol{V}^{-1}\boldsymbol{b} - \boldsymbol{M}^{-1}\boldsymbol{a}),$$
$$K = (\boldsymbol{b}, \boldsymbol{V}^{-1}\boldsymbol{b}) - (\boldsymbol{a}, \boldsymbol{M}^{-1}\boldsymbol{a}) - (\boldsymbol{d}, \boldsymbol{B}\boldsymbol{d}).$$

Поэтому (54) можно продолжить следующим образом:

$$\mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2} e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{-((\boldsymbol{x}-\boldsymbol{d}), \boldsymbol{B}_n(\boldsymbol{x}-\boldsymbol{d}))/2}\, d\boldsymbol{x} =$$
$$= \frac{|\boldsymbol{M}_n|^{1/2} e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{-(\boldsymbol{x}, \boldsymbol{B}_n\boldsymbol{x})/2}\, d\boldsymbol{x}. \tag{55}$$

Рассмотрим интеграл в правой части (55). Если $\boldsymbol{B}_n > \boldsymbol{0}$, то [6, § 6.9, теорема 3]

$$\int\limits_{\mathbb{R}^n} e^{-(\boldsymbol{x}, \boldsymbol{B}_n\boldsymbol{x})/2}\, d\boldsymbol{x} = \frac{(2\pi)^{n/2}}{|\boldsymbol{B}_n|^{1/2}}. \tag{56}$$

В противном случае

$$\int\limits_{\mathbb{R}^n} e^{-(\boldsymbol{x}, \boldsymbol{B}_n\boldsymbol{x})/2}\, d\boldsymbol{x} = \infty. \tag{57}$$

Предположим сначала, что $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1} > \boldsymbol{0}$, т.е. матрица $\boldsymbol{B}_n$ положительно определена. Тогда из (55), (56) получаем

$$\mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) = \frac{|\boldsymbol{M}_n|^{1/2} e^{-K/2}}{|\boldsymbol{V}_n|^{1/2}|\boldsymbol{B}_n|^{1/2}}. \tag{58}$$

Если же матрица $\boldsymbol{B}_n = \boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}$ не является положительно определенной, то в силу (57)

$$\mathbf{E}_{\boldsymbol{I}_n} \frac{p_{\boldsymbol{b}_n, \boldsymbol{V}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{x}) = \infty, \tag{59}$$

и поэтому условие (51) не может быть выполнено. Из (58), (59) следует лемма 2. ▲

Продолжим доказательство теоремы. Определим $\mathcal{F}(\boldsymbol{a}_n, \boldsymbol{M}_n)$ как максимальное множество, удовлетворяющее условию

$$f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n) \leqslant e^{o(n)}, \quad n \to \infty. \tag{60}$$

Это множество совпадает с определением (22). Поэтому из (35), (51), (52) и (60) следует теорема 1.

§ 4. Примеры. Частные случаи

4.1. Известные среднее $\boldsymbol{a_n}$ и ковариационная матрица $\boldsymbol{M_n}$. Рассмотрим сначала простейший случай известных среднего $\boldsymbol{a}_n = (a_1, \ldots, a_n)$ и матрицы $\boldsymbol{M}_n$ и применим теорему 3. Это позволит оценить скорость сходимости в теореме 1. Без ограничения общности в модели (2) можно считать ковариационную матрицу $\boldsymbol{M}_n$ диагональной с положительными собственными значениями $\lambda_1, \ldots, \lambda_n$ (см. замечание 1).

Тогда (см. (17))

$$D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) = D(\boldsymbol{\xi}_n \,\|\, \boldsymbol{a}_n + \boldsymbol{\eta}_n) = \frac{1}{2}\left[\sum_{i=1}^{n}\Big(\ln\lambda_i + \frac{1}{\lambda_i} - 1 + \frac{a_i^2}{\lambda_i}\Big)\right]. \tag{61}$$

В силу (29), (30) для $D = D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n})$ получаем

$$-\frac{D+1}{1-\alpha} \leqslant \ln\beta(\alpha) \leqslant -D + \mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n), \tag{62}$$

где $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ оценено в (32).

Для того чтобы оценить $\mu_0(\alpha, \boldsymbol{a}_n, \boldsymbol{M}_n)$ проще, чем (32), предположим дополнительно, что выполняется следующее условие:

III. Существует $C > 0$, такое что

$$\mathbf{E}_{\mathbf{P}}\left[D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\right]^2 \leqslant C^2 D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}). \tag{63}$$

Тогда, используя неравенство Чебышева, имеем

$$\alpha_\mu = \mathbf{P}_{\boldsymbol{I}_n}\left\{D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}_n) \geqslant \mu\right\} \leqslant$$

$$\leqslant \mu^{-2}\,\mathbf{E}_{\mathbf{P}}\left[D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\right]^2 \leqslant C^2\mu^{-2} D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}). \tag{64}$$

Для того чтобы правая часть (64) не превышала α, достаточно положить

$$\mu = C\sqrt{\frac{D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n})}{\alpha}},$$

и тогда (62) принимает вид

$$-\frac{D+1}{1-\alpha} \leqslant \ln\beta(\alpha) \leqslant -D + C\sqrt{\frac{D}{\alpha}},$$

который оценивает скорость сходимости в (62).

Заметим также, что аналогично (74), (75) нетрудно получить

$$\mathbf{E}_{\mathbf{P}}\left[D(\mathbf{P}_{\boldsymbol{I}_n} \,\|\, \mathbf{Q}_{\boldsymbol{a}_n,\boldsymbol{M}_n}) - \ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x})\right]^2 = \frac{1}{2}\sum_{i=1}^{n}\left[\Big(1 - \frac{1}{\lambda_i}\Big)^2 + 2\frac{a_i^2}{\lambda_i^2}\right]. \tag{65}$$

Поэтому условие **III** эквивалентно неравенству (см. (61) и (65))

$$\sum_{i=1}^{n}\left[\Big(1 - \frac{1}{\lambda_i}\Big)^2 + 2\frac{a_i^2}{\lambda_i^2}\right] \leqslant C^2\left[\sum_{i=1}^{n}\Big(\ln\lambda_i + \frac{1}{\lambda_i} - 1 + \frac{a_i^2}{\lambda_i}\Big)\right].$$

Замечание 4. Предположение (63) выполняется, например, в естественном "регулярном" случае, когда элементы $\boldsymbol{a}_{n+1}$, $\boldsymbol{M}_{n+1}$ являются "продолжениями" элементов $\boldsymbol{a}_n$, $\boldsymbol{M}_n$.

4.2. Неизвестное среднее $\boldsymbol{a}_n$ и известная ковариационная матрица $\boldsymbol{M}_n$. Рассмотрим случай модели (2), в котором известна ковариационная матрица $\boldsymbol{M}_n$, но не известно среднее $\boldsymbol{a}_n$. Без ограничения общности можно считать ковариационную

матрицу $\boldsymbol{M}_n$ диагональной с положительными собственными значениями $\lambda_1, \dots, \lambda_n$ (см. замечание 1). Тогда функция $f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{M}_n)$ из (20) принимает вид

$$f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{M}_n) = e^{-K},$$

где при $\boldsymbol{a}_n = (a_{1,n}, \dots, a_{n,n})$ и $\boldsymbol{b}_n = (b_{1,n}, \dots, b_{n,n})$ имеем

$$K = (\boldsymbol{b}_n, \boldsymbol{M}_n^{-1}\boldsymbol{b}_n) - (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) - \big(\boldsymbol{M}_n^{-1}(\boldsymbol{b}_n - \boldsymbol{a}_n), \boldsymbol{M}_n^{-1}(\boldsymbol{b}_n - \boldsymbol{a}_n)\big) =$$
$$= \sum_{i=1}^{n} \left[\frac{b_{i,n}^2 - a_{i,n}^2}{\lambda_i} - \frac{(b_{i,n} - a_{i,n})^2}{\lambda_i^2} \right]. \tag{66}$$

Соответствующее максимальное множество $\mathcal{F}_1(\boldsymbol{a}_n, \boldsymbol{M}_n) = \{\boldsymbol{b}_n\}$ в этом случае принимает вид (см. (22))

$$\mathcal{F}_1(\boldsymbol{a}_n, \boldsymbol{M}_n) = \{\boldsymbol{b}_n : \, K \geqslant o(n)\}, \tag{67}$$

где функция $K = K(\boldsymbol{a}_n, \boldsymbol{M}_n, \boldsymbol{b}_n)$ определена в (66).

Отметим, что в случае $\boldsymbol{M}_n = \boldsymbol{I}_n$ (т.е. когда гипотезы различаются только сдвигом $\boldsymbol{a}_n$) формулы (66), (67) принимают особенно простой вид:

$$K = 2(\boldsymbol{a}_n, \boldsymbol{b}_n - \boldsymbol{a}_n), \quad \mathcal{F}_1(\boldsymbol{a}_n, \boldsymbol{I}_n) = \{\boldsymbol{b}_n : \, (\boldsymbol{a}_n, \boldsymbol{b}_n - \boldsymbol{a}_n) \geqslant o(n)\}. \tag{68}$$

Эти результаты следуют также из работ [11, 12] (где эта задача рассматривалась в гильбертовом и банаховом пространствах).

4.3. Известное среднее $\boldsymbol{a}_n$ и неизвестная ковариационная матрица $\boldsymbol{M}_n$. Ограничимся случаем $\boldsymbol{a}_n = \boldsymbol{0}_n$. Тогда функция $f_{\boldsymbol{a}_n, \boldsymbol{M}_n}(\boldsymbol{b}_n, \boldsymbol{V}_n)$ из (20) при $\boldsymbol{a}_n = \boldsymbol{b}_n = \boldsymbol{0}_n$ принимает вид

$$f_{\boldsymbol{0}_n, \boldsymbol{M}_n}(\boldsymbol{0}_n, \boldsymbol{V}_n) = \frac{|\boldsymbol{M}_n|}{|\boldsymbol{V}_n| \cdot \big|\boldsymbol{I}_n + \boldsymbol{V}_n^{-1} - \boldsymbol{M}_n^{-1}\big|}. \tag{69}$$

Соответствующее максимальное множество $\mathcal{F}_1(\boldsymbol{0}_n, \boldsymbol{M}_n) = \{\boldsymbol{V}_n\}$ в этом случае принимает вид (см. (22))

$$\mathcal{F}_1(\boldsymbol{0}_n, \boldsymbol{M}_n) = \{\boldsymbol{V}_n : \, f_{\boldsymbol{0}_n, \boldsymbol{M}_n}(\boldsymbol{0}_n, \boldsymbol{V}_n) \leqslant e^{o(n)}\}. \tag{70}$$

Формулы (69), (70) совпадают с соответствующими результатами в [8, теорема 1].

ПРИЛОЖЕНИЕ: ДОКАЗАТЕЛЬСТВО ЛЕММЫ 1

Пусть $\boldsymbol{\xi}_n$ – гауссовский случайный вектор с распределением $\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, а $\boldsymbol{A}_n$ – симметричная $(n \times n)$-матрица с собственными значениями $\{a_i\}$. Рассмотрим квадратичную форму $(\boldsymbol{\xi}_n, \boldsymbol{A}_n\boldsymbol{\xi}_n)$. Существует ортогональная матрица $\boldsymbol{T}_n$, такая что $\boldsymbol{T}_n'\boldsymbol{A}_n\boldsymbol{T}_n = \boldsymbol{B}_n$, где $\boldsymbol{B}_n$ – диагональная матрица с диагональными элементами $\{a_i\}$ [6, § 4.7]. Так как $\boldsymbol{T}_n\boldsymbol{\xi}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, то квадратичные формы $(\boldsymbol{\xi}_n, \boldsymbol{A}_n\boldsymbol{\xi}_n)$ и $(\boldsymbol{\xi}_n, \boldsymbol{B}_n\boldsymbol{\xi}_n)$ имеют одинаковые распределения. Поэтому из формулы (12) имеем

$$\ln \frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n, \boldsymbol{M}_n}}(\boldsymbol{y}_n) \stackrel{d}{=} \frac{1}{2}\big[\ln |\boldsymbol{M}_n| + (\boldsymbol{a}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) + \eta_n\big], \tag{71}$$

где

$$\eta_n = (\boldsymbol{y}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I})\boldsymbol{y}_n) - 2(\boldsymbol{y}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n). \tag{72}$$

Введем величину (см. (31))

$$\alpha_\mu = \mathbf{P}_{\boldsymbol{I}_n}\left\{\ln\frac{p_{\boldsymbol{I}_n}}{p_{\boldsymbol{a}_n,\boldsymbol{M}_n}}(\boldsymbol{x}_n) \leqslant D(\boldsymbol{I}_n \,\|\, \boldsymbol{a}_n, \boldsymbol{M}_n) - \mu\right\}. \tag{73}$$

Тогда с помощью (71), (72) и (17) для α_μ из (73) имеем

$$\begin{aligned}\alpha_\mu &\leqslant \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|(\boldsymbol{\xi}_n, (\boldsymbol{M}_n^{-1} - \boldsymbol{I})\boldsymbol{\xi}_n) - 2(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) - \sum_{i=1}^n\Big(\frac{1}{\lambda_i} - 1\Big)\right| > 2\mu\right\} = \\ &= \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|\sum_{i=1}^n\Big(\frac{1}{\lambda_i} - 1\Big)(\xi_i^2 - 1) - 2(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\right| > 2\mu\right\} \leqslant P_1 + P_2,\end{aligned} \tag{74}$$

где

$$\begin{aligned}P_1 &= \mathbf{P}_{\boldsymbol{\xi}_n}\left\{\left|\sum_{i=1}^n\Big(\frac{1}{\lambda_i} - 1\Big)(\xi_i^2 - 1)\right| > \mu\right\}, \\ P_2 &= \mathbf{P}_{\boldsymbol{\xi}_n}\big\{\big|(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)\big| > \mu/2\big\}.\end{aligned} \tag{75}$$

Для того чтобы оценить величину P_1 из (75), используем следующий результат [13, п. III.5.15]: пусть $\zeta_1, \ldots, \zeta_n$ – независимые случайные величины с $\mathbf{E}\,\zeta_i = 0$, $i = 1, \ldots, n$. Тогда для любого $1 \leqslant p \leqslant 2$

$$\mathbf{E}\left|\sum_{i=1}^n \zeta_i\right|^p \leqslant 2\sum_{i=1}^n \mathbf{E}\,|\zeta_i|^p. \tag{76}$$

Поэтому, используя для P_1 неравенство Чебышева и (76), получаем

$$\begin{aligned}P_1 &\leqslant \mu^{-p}\,\mathbf{E}\left|\sum_{i=1}^n\Big(\frac{1}{\lambda_i} - 1\Big)(\xi_i^2 - 1)\right|^p \leqslant 2\mu^{-p}\sum_{i=1}^n\Big|\frac{1}{\lambda_i} - 1\Big|^p\,\mathbf{E}\,|\xi_i^2 - 1|^p \leqslant \\ &\leqslant 2\mu^{-p}\sum_{i=1}^n\Big|\frac{1}{\lambda_i} - 1\Big|^p\big(\mathbf{E}\,|\xi^2 - 1|^2\big)^{p/2} \leqslant 2\mu^{-p}6^{p/2}\sum_{i=1}^n\Big|\frac{1}{\lambda_i} - 1\Big|^p \leqslant \\ &\leqslant 12\mu^{-p}\sum_{i=1}^n\Big|\frac{1}{\lambda_i} - 1\Big|^p.\end{aligned} \tag{77}$$

Для того чтобы оценить величину P_2 из (74), (75), заметим, что

$$(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) \sim \mathcal{N}(0, \|\boldsymbol{M}_n^{-1}\boldsymbol{a}_n\|),$$

и тогда

$$(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n) \stackrel{d}{=} \|\boldsymbol{M}_n^{-1}\boldsymbol{a}_n\|\xi.$$

Поэтому, используя стандартную оценку

$$\mathbf{P}(|\xi| \geqslant z) \leqslant e^{-z^2/2}, \quad z \geqslant 0,$$

получаем ($\xi_i \sim \mathcal{N}(0,1)$)

$$P_2 = \mathbf{P}_{\boldsymbol{\xi}_n}\big\{|(\boldsymbol{\xi}_n, \boldsymbol{M}_n^{-1}\boldsymbol{a}_n)| > \mu/2\big\} \leqslant e^{-\mu^2/(8\|\boldsymbol{M}_n^{-1}\boldsymbol{a}_n\|^2)}. \tag{78}$$

Для выполнения условия $\alpha_\mu \leqslant \alpha$ выберем μ так, что $\max\{P_1, P_2\} \leqslant \alpha/2$. В силу (77) и (78) для этого достаточно положить μ удовлетворяющим оценке (32).

СПИСОК ЛИТЕРАТУРЫ

1. *Вальд А.* Статистические решающие функции // Позиционные игры. М.: Наука, 1967. С. 300–522.
2. *Леман Э.* Проверка статистических гипотез. М.: Наука, 1979.
3. *Poor H.V.* An Introduction to Signal Detection and Estimation. New York: Springer, 1994.
4. *Zhang W., Poor H.V.* On Minimax Robust Detection of Stationary Gaussian Signals in White Gaussian Noise // IEEE Trans. Inform. Theory. 2011. V. 57. № 6. P. 3915–3924. https://doi.org/10.1109/TIT.2011.2136210
5. *Бурнашев М.В.* Об обнаружении гауссовских стохастических последовательностей // Пробл. передачи информ. 2017. Т. 53. № 4. С. 49–68. http://mi.mathnet.ru/ppi2252
6. *Беллман Р.* Введение в теорию матриц. М.: Наука, 1976.
7. *Хорн Р., Джонсон Ч.* Матричный анализ. М.: Мир, 1989.
8. *Бурнашев М.В.* О минимаксном обнаружении гауссовских стохастических последовательностей и гауссовских стационарных сигналов // Пробл. передачи информ. 2021. Т. 57. № 3. С. 55–72. https://doi.org/10.31857/S0555292321030049
9. *Кульбак С.* Теория информации и статистика. М.: Наука, 1967.
10. *Burnashev M.V.* On Stein's Lemma in Hypotheses Testing in General Non-Asymptotic Case // Stat. Inference Stoch. Process. 2022. Online First article. https://doi.org/10.1007/s11203-022-09278-4
11. *Бурнашев М.В.* О минимаксном обнаружении неточно известного сигнала на фоне белого гауссовского шума // Теория вероятн. и ее примен. 1979. Т. 24. № 1. С. 106–118. http://mi.mathnet.ru/tvp957
12. *Бурнашев М.В.* О различении гипотез для гауссовских мер и одна геометрическая характеризация гауссовского распределения // Матем. заметки. 1982. Т. 32. № 4. С. 549–556. http://mi.mathnet.ru/mz6021
13. *Петров В.В.* Суммы независимых случайных величин. М.: Наука, 1972.

Бурнашев Марат Валиевич
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
burn@iitp.ru

Поступила в редакцию
28.03.2022
После доработки
18.08.2022
Принята к публикации
19.08.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 621.391 : 517.938 : 519.766

© 2022 г. **М.Л. Бланк**

ВОССТАНАВЛИВАЕМЫЙ ФОРМАЛЬНЫЙ ЯЗЫК

Изучается задача восстановления искаженных произвольно длинных сообщений, записанных на некотором динамически заданном формальном языке. Получены необходимые и достаточные условия на задание языка для наличия допустимого сообщения в окрестности искаженного при условии, что локальные искажения происходят редко.

Ключевые слова: кодирование, формальный язык, динамическая система, отслеживание псевдотраекторий.

DOI: 10.31857/S055529232203007X, **EDN:** EAQKEE

§ 1. Введение

Один из основных вопросов теории кодирования состоит в том, как записать сообщение в такой форме, чтобы искажение некоторой части записи не помешало восстановлению сообщения в исходной форме (см., например, [1–4]). Мы попробуем взглянуть на эту задачу с несколько неожиданной точки зрения: какими свойствами должен обладать формальный язык, на котором пишется сообщение[1] для того, чтобы после "разумных" искажений (уже не являющееся допустимым) исходное сообщение могло быть хотя бы приблизительно восстановлено. Чуть более формально это означает, что для любого "разумно" искаженного сообщения произвольно большой длины (уже не принадлежащего нашему формальному языку) нашлось бы близкое к нему допустимое сообщение. Довольно близкая задача, но с разрешением пропусков кусков сообщений и их перестановками, изучалась, например, в [5].

Под формальным языком мы будем понимать следующее. Рассмотрим ориентированный граф G с конечным множеством вершин $\mathcal{A} := \{1, 2, \ldots, r\}$ (алфавитом языка), $r > 1$, и дискретной метрикой

$$\rho(a, b) := \begin{cases} 0 & \text{при } a = b, \\ 1 & \text{в противном случае,} \end{cases}$$

где $a, b \in \mathcal{A}$. Используемую нами терминологию, связанную с теорией графов, можно найти, например, в [6].

Определение 1. Любой бесконечный в обе стороны путь в ориентированном графе G будем называть *допустимым сообщением*, а объединение всех допустимых сообщений назовем *формальным языком*.

Не теряя общности, мы полагаем, что граф G не содержит "стоков" (вершин, из которых нельзя выйти) и "источников" (вершин, в которые нельзя попасть). В противном случае этого можно добиться модификацией (уменьшением) списка вершин.

[1] Сообщения, принадлежащие нашему формальному языку, назовем допустимыми.

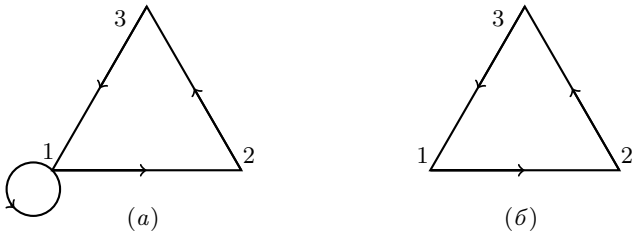

Рис. 1. Примеры графов, порождающих (*а*) восстанавливаемые и (*б*) невосстанавливаемые формальные языки

Заметим, что более сложные грамматические правила, связанные с запретом переходов между выделенными "словами" из нескольких букв, нетрудно реализовать при увеличении размера алфавита.

О п р е д е л е н и е 2. *Возмущением* допустимого сообщения $\vec{v} := \{v_t\}_{t\in\mathbb{Z}}$ в момент времени t_0 назовем двустороннюю последовательность вершин $\vec{w} := \{w_t\}_{t\in\mathbb{Z}}$, совпадающую с $\vec{v}$ до момента времени t_0, а начиная с t_0 стартующую из произвольной вершины графа G, такую что при любом $t > t_0$ ребро $(w_{t-1}, w_t) \in G$.

Иными словами, возмущенное сообщение состоит из части допустимого сообщения $\vec{v}$ до момента времени $t_0 - 1$ и части некоторого другого допустимого сообщения с момента времени t_0. В терминах блуждания по графу G это означает, что по времени от $-\infty$ до $t_0 - 1$ мы следуем по какому-то пути на графе G, в момент t_0 происходит перескок в произвольную вершину графа, а после этого мы вновь следуем пути на графе G.

Напомним, что плотностью последовательности целых чисел $\{t_i\}_{i\in\mathbb{Z}}$ называется

$$\limsup_{n\to\infty} \frac{\#\{i \in \mathbb{Z} : \ |t_i| \leqslant n\}}{2n+1}.$$

О п р е д е л е н и е 3. *ε-искаженным* сообщением в моменты времени $\ldots < t_{-1} < t_0 < t_1 < \ldots$ назовем результат применения последовательности возмущений при условии, что плотность последовательности моментов возмущения не превосходит значения $\varepsilon > 0$.

Таким образом, низкая плотность последовательности моментов возмущения связана с тем, что они происходят редко по времени. Здесь важно отметить, что в силу отсутствия "стоков" и "источников" как сами сообщения, так и результат их возмущения описываются двусторонне бесконечными последовательностями.

О п р е д е л е н и е 4. Будем говорить, что произвольная последовательность вершин $\vec{w}$ графа G *отслеживается в среднем* допустимым сообщением $\vec{v}$ с точностью δ, если

$$\limsup_{n\to\infty} \frac{1}{2n+1} \sum_{k=-n}^{n} \rho(w_k, v_k) \leqslant \delta. \tag{1}$$

О п р е д е л е н и е 5. Будем говорить, что формальный язык, построенный по ориентированному графу G, является *восстанавливаемым*, если $\forall\delta > 0 \ \exists\varepsilon > 0$, такое что любое ε-искаженное сообщение отслеживается с точностью δ.

Наш основной результат дает необходимые и достаточные условия восстанавливаемости формального языка, построенного по ориентированному графу G.

Обозначим через π матрицу переходов в графе G, т.е. $\pi_{ij} = 1$, если ребро (i, j) принадлежит G, и $\pi_{ij} = 0$ в противном случае.

Теорема 1. *Формальный язык, построенный по ориентированному графу G, является восстанавливаемым тогда и только тогда, когда найдется такое натуральное N, что $\pi^N > 0$ (т.е. все элементы матрицы π^N строго положительны).*

§ 2. Доказательство теоремы 1

С точки зрения теории динамических систем последовательное построение допустимого пути на графе G можно описать с помощью многозначного отображения $T\colon \mathcal{A} \to 2^{\mathcal{A}}$, ставящего в соответствие элементу $a \in \mathcal{A}$ все вершины графа G, в которые можно из него попасть за один шаг. Поэтому задача проверки восстанавливаемости языка сводится к известной в теории динамических систем задаче отслеживания псевдотраекторий (траекторий возмущенного отображения) – см., например, [7; 8; 9, раздел 6; 10, глава 6]. Однако, во-первых, на сегодня нет решения задачи отслеживания для многозначных отображений, а во-вторых, поскольку фазовое пространство $\mathcal{A}$ дискретно, возмущения отображения ни в каком смысле не являются малыми.

Поэтому мы пойдем другим путем. Рассмотрим отображение левого сдвига σ, действующее в пространстве двусторонних последовательностей $\vec{X} := \mathcal{A}^{\mathbb{Z}}$ с элементами из алфавита $\mathcal{A}$, т.е. $(\sigma\vec{x})_i := x_{i+1} \ \forall \vec{x} \in \vec{X}$, $i \in \mathbb{Z}$.

Определение 6. *Топологической марковской цепью* (subshift of finite type) с матрицей переходов π называется ограничение левого сдвига σ на инвариантное множество *допустимых* последовательностей $\vec{X}_\pi$, определяемых следующим условием: $\vec{x} \in \vec{X}_\pi$ тогда и только тогда, когда $\pi_{x_i x_j} > 0 \ \forall i \in \mathbb{Z}$.

В соответствии с терминологией, описанной в предыдущем параграфе, множество допустимых последовательностей совпадает с множеством допустимых сообщений. Как и ранее, мы предполагаем отсутствие “стоков” и “источников”. Поэтому допустимые последовательности являются двусторонне бесконечными.

Определение 7. Для упорядоченной пары допустимых последовательностей $\vec{x}, \vec{y}$ под *возмущением* в момент времени t будем понимать такую новую последовательность $\vec{z} \in \vec{X}$ (вообще говоря, не лежащую в $\vec{X}_\pi$), что

$$z_i = x_i \ \forall i < t, \quad z_i = y_i \ \forall i \geqslant t.$$

Определение 8. *Псевдотраекторией* $\widetilde{y}$ левого сдвига σ назовем траекторию возмущенной системы с произвольным набором моментов возмущения $\ldots < t_{-1} < < t_0 < t_1 < \ldots$ и выбором самих возмущений.

Определим расстояние между последовательностями:

$$\vec{\rho}(\vec{w}, \vec{v}) := \limsup_{n\to\infty} \sum_{k=-n}^{n} 2^{-|k|}\rho(w_k, v_k). \tag{2}$$

Определение 9. Будем говорить, что псевдотраектория $\widetilde{y} := \{\vec{y}_n\}_{n\in\mathbb{Z}}$ *отслеживается в среднем* с точностью δ траекторией $\{\sigma^n\vec{x}\}_{n\in\mathbb{Z}}$ точки $\vec{x} \in \vec{X}$, если

$$\limsup_{n\to\infty} \frac{1}{2n+1} \sum_{k=-n}^{n} \vec{\rho}(\vec{y}_k, \sigma^k\vec{x}) \leqslant \delta. \tag{3}$$

Замечание 1. Заметим, что при $\vec{w} := \vec{y}$, $\vec{v} := \vec{x}$ левые части неравенств (1) и (3) согласованы в том смысле, что первая из них не превышает второй.

Аналогично предыдущему параграфу низкая плотность моментов возмущения соответствует редким возмущениям.

Теперь мы уже находимся в рамках теории однозначных динамических систем, и для анализа возможности отслеживания можно воспользоваться следующим недавно полученным результатом [11].

Определение 10. Отображение σ метрического пространства $(\vec{X}, \vec{\rho})$ в себя удовлетворяет условию *склеивания со скоростью* $\varphi\colon \mathbb{Z} \to \mathbb{R}$, если для любой пары односторонних последовательностей элементов $\vec{X}$ вида $\{\sigma^k \vec{x}\}_{k<0}$ и $\{\sigma^k \vec{y}\}_{k\geqslant 0}$ найдется такая последовательность $\vec{z} \in \vec{X}_\pi$, что

$$\vec{\rho}(\sigma^k \vec{x}, \sigma^k \vec{z}) \leqslant \varphi(k) \cdot \vec{\rho}(\vec{x}, \vec{y}) \ \ \forall k < 0, \quad \vec{\rho}(\sigma^k \vec{y}, \sigma^k \vec{z}) \leqslant \varphi(k) \cdot \vec{\rho}(\vec{x}, \vec{y}) \ \ \forall k \geqslant 0.$$

Другими словами, траектория точки $\vec{z}$ одновременно отслеживает траекторию назад точки $\vec{x}$ и траекторию вперед точки $\vec{y}$ со скоростью, контролируемой функцией φ.

Теорема 2 [11]. *Если возмущение ε-редко, то при выполнении условия склеивания с суммируемой функцией φ любая псевдотраектория отслеживается в среднем с точностью $C\varepsilon$, где $C < \infty$ не зависит от выбора возмущений.*

Доказательство. Для применения этого результата при проверке достаточности условий теоремы 1 нужно доказать выполнение условия склеивания с суммируемой точностью аппроксимации φ для отображения сдвига. Согласно условию π^N – положительная матрица. Поэтому за время N мы можем перейти из любого элемента алфавита $\mathcal{A}$ в любой другой. Таким образом, полагая функцию φ равной индикаторной функции целочисленного отрезка $[-N, N]$, получаем требуемый результат.

Остается проверить необходимость. Предположим, что условие теоремы не выполнено и $\nexists N : \ \pi^N := \left(\pi_{ij}^{(n)}\right) > 0$. Из этого следует, что

$$\forall n \in \mathbb{Z}_+ \ \ \exists i_n, j_n : \ \pi_{i_n, j_n}^{(n)} = 0.$$

Действительно, если бы это было не так, то нашлось бы такое $k \in \mathbb{Z}_+$, что $\pi^k > 0$. Но в этом случае $\pi^n > 0 \ \ \forall n > k$, что противоречит предположению.

Как мы уже отмечали, матрица переходов π индуцирует многозначное отображение алфавита в себя по формуле $\pi a := \{b \in \mathcal{A} : \ \pi_{ab} > 0\}$. В этих терминах из отсутствия строгой положительности π и конечности $\mathcal{A}$ следует, что при некотором $M \in \mathbb{Z}_+$ имеется разбиение $\mathcal{A} := \bigsqcup_i \mathcal{A}_i$ на непустые π^M-инвариантные подмножества, т.е. $(\pi^M)^{-1}\mathcal{A}_i = \mathcal{A}_i \ \ \forall i$. Из этого следует, что при x_0, y_0, принадлежащих различным элементам этого разбиения, соответствующие допустимые последовательности $\vec{x}, \vec{y}$ не пересекаются, т.е. $x_i \neq y_i \ \ \forall i \in \mathbb{Z}$. Поэтому не существует их "склейки" с суммируемой функцией φ.

С другой стороны, отслеживание в среднем в рассматриваемой постановке эквивалентно тому, что отслеживающая траектория может отличаться от отслеживаемой лишь на конечном временном интервале. В противном случае в силу дискретности метрики ρ усредненное расстояние (левая часть формулы (3)) не может быть малой величиной. Поэтому здесь отслеживание в среднем эквивалентно выполнению условия склеивания с суммируемой функцией φ. ▲

Автор благодарен рецензенту за конструктивные и полезные замечания.

СПИСОК ЛИТЕРАТУРЫ

1. *Марков А.А.* Введение в теорию кодирования. М.: Наука, 1982.
2. *Хэмминг Р.У.* Теория кодирования и теория информации. М.: Радио и связь, 1983.

3. *Мак-Вильямс Ф.Дж., Слоэн Н.Дж.А.* Теория кодов, исправляющих ошибки. М.: Связь, 1979.
4. *Ромащенко А.Е., Румянцев А.Ю., Шень А.* Заметки по теории кодирования. М.: МЦНМО, 2017.
5. *Elishco O., Barg A.* Recoverable Systems, arXiv:2010.00589v2 [cs.IT], 2022.
6. *Bollobás B.* Modern Graph Theory. New York: Springer, 1998.
7. *Аносов Д.В.* Об одном классе инвариантных множеств гладких динамических систем // Тр. V Междунар. конф. по нелинейным колебаниям (25 августа – 4 сентября 1969 г.). Т. 2: Качественные методы. Киев: Ин-т матем. АН Украины, 1970. С. 39–45.
8. *Blank M.* Metric Properties of ε-Trajectories of Dynamical Systems with Stochastic Behaviour // Ergodic Theory Dynam. Systems. 1988. V. 8. № 3. P. 365–378. https://doi.org/10.1017/S014338570000451X
9. *Blank M.* Discreteness and Continuity in Problems of Chaotic Dynamics. Providence, R.I.: Amer. Math. Soc., 1997.
10. *Katok A., Hasselblatt B.* Introduction to the Modern Theory of Dynamical Systems. Cambridge: Cambridge Univ. Press, 1995.
11. *Blank M.* Average Shadowing and Gluing Property, arXiv:2202.13407v1 [math.DS], 2022.

Бланк Михаил Львович
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
Национальный исследовательский университет
"Высшая школа экономики", Москва
blank@iitp.ru

Поступила в редакцию
23.03.2022
После доработки
11.06.2022
Принята к публикации
11.06.2022

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 **2022** **Вып. 3**

УДК 519.651 : 517.589

© 2022 г. **Е.А. Карацуба**

БЫСТРЫЕ АЛГОРИТМЫ ВЫЧИСЛЕНИЯ ЭЛЕМЕНТАРНЫХ АЛГЕБРАИЧЕСКИХ И ОБРАТНЫХ ФУНКЦИЙ С ПРИМЕНЕНИЕМ БВЕ

Построены новые быстрые алгоритмы вычисления элементарных алгебраических и обратных функций, основанные на применении двух методов – метода А.А. Карацубы 1960 г. и авторского метода БВЕ 1990 г. Сложность вычисления близка к оптимальной. Алгоритмы допускают частичное распараллеливание.

Ключевые слова: быстрые алгоритмы, сложность вычисления, метод А.А. Карацубы, метод БВЕ, метод Ньютона, элементарные алгебраические функции, обратные функции, рациональная функция, логарифмическая функция.

DOI: 10.31857/S0555292322030081, **EDN:** EAXCNG

Памяти Юрия Ивановича Журавлёва
(14.01.1935–14.01.2022)

"Не огорчайся! Метод, который не украли, – это плохой метод, потому что он никому не нужен!" (Ю.И. Журавлёв, 2000 г.)

§ 1. Введение. Основные определения

Под алгебраическими функциями понимают функции, которые в окрестности каждой точки области определения могут быть заданы алгебраическими уравнениями. Под элементарными алгебраическими функциями будем иметь в виду степенную функцию и рациональную функцию. Под обратными функциями будем иметь в виду функции, обратные к простейшим трансцендентным (экспоненциальной и тригонометрическим) функциям, т.е. логарифм, арктангенс, арксинус и т.п.

Далее считаем, что числа записаны в двоичной системе счисления, знаки которой 0 и 1 называются битами. Таким образом, числа записываются в виде

$$z = x_{-j}2^j + x_{-j+1}2^{j-1} + \ldots + x_0 + x_1 2^{-1} + x_2 2^{-2} + \ldots, \quad x_j = 0 \text{ или } 1.$$

Определение 1. Запись знаков 0, 1, плюс, минус, скобка, а также сложение, вычитание и умножение двух битов назовем одной битовой операцией (далее – просто операцией).

Определение 2. Пусть элементарная алгебраическая или простейшая трансцендентная функция $y = f(z)$ задана в некоторой ограниченной области $D \in E$, $y = f(z)$ не имеет в $\bar{D}$ особенностей и ограничена вместе со своей производной. Переменная и значение функции записываются последовательностями

$$z = (\widetilde{x}_{-j}, \ldots, \widetilde{x}_0, \widetilde{x}_1, \ldots), \quad y = (\widetilde{y}_{-j}, \ldots, \widetilde{y}_0, \widetilde{y}_1, \ldots), \tag{1}$$

где $\widetilde{x}_i, \widetilde{y}_j$ равны 0 или 1. Вычислить функцию $y = f(z)$ в точке $z = z_0 \in D$ с точностью до n знаков означает найти такое число S_n, что

$$|f(z_0) - S_n| \leqslant c2^{-n}, \tag{2}$$

где постоянная c не зависит от n.

Вычисление функции $f(z)$ в точке $z = z_0$ можно заменить ее вычислением в точке $z = z_n$, где $|z_0 - z_n| < 2^{-n}$, поскольку $f'(z)$ ограничена в $\bar{D}$. Таким образом, все участвующие в вычислениях числа можно записывать в виде конечных последовательностей, причем из (1), (2) легко видеть, что z и y можно представить в виде целой части и $C_0 n$ двоичных знаков после запятой, $C_0 = \text{const}$. Поскольку целые части $[z]$, $[y]$ из (1) являются фиксированными величинами, а основным растущим параметром является точность вычисления n, $n \to +\infty$, то действия фактически производятся над числами, имеющими порядка n знаков. Такое n-значное число можно записать в виде

$$z = 2^{n-1} + \varepsilon_{-n+2}2^{n-2} + \ldots + \varepsilon_2 2^2 + \varepsilon_1 2 + \varepsilon_0,$$

где $\varepsilon_{-n+2}, \varepsilon_{-n+3}, \ldots, \varepsilon_2, \varepsilon_1, \varepsilon_0$ равны 0 или 1.

Определение 3. Сложностью умножения двух n-значных чисел $M(n)$ называется число операций, достаточное для вычисления произведения двух n-значных чисел.

О первом быстром методе – методе умножения – см. [1–3]. Далее предполагаем, что для сложности умножения двух n-значных чисел справедлива оценка

$$M(n) = O\big(n \log^C n\big),$$

где C – константа, т.е. сложность используемого алгоритма умножения не хуже, чем сложность Шёнхаге – Штрассена (см. алгоритмы из [4, 5]).

Определение 4. Количество операций, достаточное для вычисления функции $f(z)$ в точке z_0 с точностью до n знаков посредством данного алгоритма, называется сложностью вычисления $f(z)$ в точке z_0 и обозначается $s_f(n) = s_{f,z_0}(n)$.

Определение 5. Будем называть *быстрыми* такие методы и алгоритмы вычисления функции f, что для них

$$s_f(n) = O\big(n \log^K n\big), \quad \text{где } K \text{ – константа.} \tag{3}$$

Один из таких быстрых методов – метод БВЕ (быстрого вычисления Е-функций, см. [6–9]) был создан для вычисления простейших и высших трансцендентных функций со сложностью (3). Это второй после метода АГС Гаусса (см., например, [10–12]) метод быстрого вычисления классических констант π и e, а также простейших трансцендентных функций при любом аргументе. При этом, в отличие от АГС, с помощью БВЕ со сложностью, близкой к оптимальной, можно вычислить также некоторые высшие трансцендентные функции для алгебраических значений аргумента и параметров.

В [13–15] были представлены первые алгоритмы быстрого вычисления элементарных алгебраических функций, основанные на методе Ньютона. Например, самый простой алгоритм деления числа a на число b заключается в вычислении методом Ньютона обратной величины $1/b$ с точностью до n знаков с последующим "быстрым умножением" на a.

Ранее, в промежуточных вычислениях в БВЕ-алгоритмах всегда предполагалось использование метода Ньютона для вычисления элементарных алгебраических и обратных функций. В то же время, возникает вопрос о существовании других быстрых алгоритмов для вычисления этих функций. Цель настоящей статьи – построить

такие алгоритмы с применением в них БВЕ-вычислений. При этом иногда там будут использоваться также конструкции метода А.А. Карацубы от 1960 г. (см. [1–3]). Заметим, что сам А.А. Карацуба не называл свой метод каким-либо именем, зато другие пользователи этого метода называли его многочисленными именами, среди которых самыми распространенными являются “Divide and Conquer” (разделяй и властвуй) и “Binary Splitting” (бинарное разбиение).

§ 2. Алгоритм быстрого деления

Пусть a и b – натуральные числа, меньшие 2^{n+1} и $b < a$. Чтобы разделить a на b с остатком, нужно найти такие целые q и r, чтобы

$$a = bq + r, \quad 0 \leqslant r < b.$$

Если будет найдено S_n, такое что

$$\left| S_n - \frac{1}{b} \right| < 2^{-(n+1)}, \tag{4}$$

то одно из трех чисел $[aS_n]$, $[aS_n] - 1$, $[aS_n] + 1$ равно q, а именно то число $d \in \{[aS_n] - 1, [aS_n], [aS_n] + 1\}$, для которого

$$0 \leqslant a - bd < b.$$

При этом

$$r = a - bq.$$

Если $a < b$, то для вычисления частного a/b вычисляется S_n, удовлетворяющее неравенству (4), а затем с точностью $2^{-(n+1)}$ посредством быстрого алгоритма вычисляется произведение aS_n, и в результате получаем значение a/b с точностью до n знаков. Таким образом, задача сводится к построению быстрого алгоритма вычисления обратной величины $1/b$ с точностью до n знаков.

Пусть $b = z$ является $(n+1)$-значным числом, $n \geqslant 2$:

$$z = 2^n + \alpha_{n-1}2^{n-1} + \ldots + \alpha_m 2^m + \ldots + \alpha_1 2 + \alpha_0,$$

где $\alpha_m = 0$ или $\alpha_m = 1$, $m = 0, 1, 2, \ldots, n-1$. Отсюда

$$\frac{1}{z} = \frac{1}{2^n(1 + \alpha_{n-1}2^{-1} + \ldots + \alpha_m 2^{m-n} + \ldots + \alpha_1 2^{1-n} + \alpha_0 2^{-n})}, \tag{5}$$

и нужно найти S_n, такое что

$$\frac{1}{z} = 2^{-n}S_n + \theta 2^{-2n-1}, \quad |\theta| \leqslant 1. \tag{6}$$

Заметим, что число $1 + \alpha_{n-1}2^{-1} + \ldots + \alpha_m 2^{m-n} + \ldots + \alpha_1 2^{1-n} + \alpha_0 2^{-n}$ требует для записи $n+1$ позицию, т.е. должно учитываться как $(n+1)$-значное. Обозначим

$$x = \alpha_{n-1}2^{-1} + \alpha_{n-2}2^{-2} + \ldots + \alpha_m 2^{m-n} + \ldots + \alpha_1 2^{1-n} + \alpha_0 2^{-n}, \tag{7}$$

$\alpha_m = 0$ или $\alpha_m = 1$, $m = 0, 1, 2, \ldots, n-1$, и

$$\frac{1}{z} = 2^{-n}\frac{1}{1+x},$$

где с учетом (7)

$$x \leqslant \frac{1}{2}\left(1+\frac{1}{2}+\ldots+\frac{1}{2^{n-1}}\right)=1-\frac{1}{2^n}<1.$$

С другой стороны, при $x<1$

$$\frac{1}{1+x}=(1-x)(1+x^2+\ldots+x^{2m}+\ldots)=$$
$$=(1-x)(1+x^2)(1+x^4)\ldots(1+x^{2^k})\ldots \tag{8}$$

Рассмотрим равенство (7). Число α_{n-1} может принимать два значения: 0 или 1.

1) Пусть $\alpha_{n-1}=0$. Тогда из (7) имеем $x \leqslant \frac{1}{2^2}\left(1+\frac{1}{2}+\ldots+\frac{1}{2^{n-2}}\right)<\frac{1}{2}$. При

$$m=n+1 \tag{9}$$

из (8) имеем

$$\frac{1}{1+x}-(1-x)\left(1+x^2+\ldots+x^{2(m-1)}\right)=$$
$$=\frac{x^{2m}}{1+x}<\frac{1}{2^{2n+2}}=\theta_0 2^{-2n-2},\quad |\theta_0|<1. \tag{10}$$

2) Пусть $\alpha_{n-1}=1$. Тогда из (5) находим

$$\frac{1}{z}=\frac{1}{2^n\left(1+\frac{1}{2}+\alpha_{n-2}2^{-2}+\ldots+\alpha_m 2^{m-n}+\ldots+\alpha_1 2^{1-n}+\alpha_0 2^{-n}\right)}=$$
$$=2^{-n-1}\left(1-\frac{1}{4}+\alpha_{n-2}2^{-3}+\ldots+\alpha_m 2^{m-n-1}+\ldots+\alpha_1 2^{-n}+\alpha_0 2^{-n-1}\right)^{-1}=$$
$$=2^{-n-1}\left(1+\frac{1}{4}\left(\alpha_0 2^{-n+1}+\alpha_1 2^{-n+2}+\ldots+\alpha_m 2^{m-n+1}+\ldots+\alpha_{n-2}2^{-1}-1\right)\right)^{-1}=$$
$$=\frac{2^{-n-1}}{1+\widetilde{x}},\quad |\widetilde{x}|<\frac{1}{8},$$

и следовательно, с учетом (9),

$$\frac{1}{1+\widetilde{x}}-(1-\widetilde{x})\left(1+\widetilde{x}^2+\ldots+\widetilde{x}^{2(m-1)}\right)=\frac{\widetilde{x}^{2m}}{1+\widetilde{x}}=\theta_1 2^{-6n-6},\quad |\theta_1|<1. \tag{11}$$

Обозначим

$$S_n=(1-x)(1+x^2+\ldots+x^{2m})=(1-x)(1+x^2)(1+x^4)\ldots(1+x^{2^k}), \tag{12}$$
$$\widetilde{S}_n=(1-\widetilde{x})(1+\widetilde{x}^2+\ldots+\widetilde{x}^{2m})=(1-\widetilde{x})(1+\widetilde{x}^2)(1+\widetilde{x}^4)\ldots(1+\widetilde{x}^{2^k}), \tag{13}$$

при этом

$$m=1+2+2^2+\ldots+2^{k-1}=2^k-1, \tag{14}$$

и $x,\widetilde{x}$ определяются формулами

$$x=\alpha_0 2^{-n}+\alpha_1 2^{1-n}+\ldots+\alpha_m 2^{m-n}+\ldots+\alpha_{n-2}2^{-2},$$
$$\widetilde{x}=\widetilde{\alpha}_0 2^{-n-1}+\widetilde{\alpha}_1 2^{-n}+\ldots+\widetilde{\alpha}_m 2^{m-n-1}+\ldots+\widetilde{\alpha}_{n-2}2^{-3}-2^{-2},$$

где коэффициенты α_i, $\widetilde{\alpha}_j$ равны 0 или 1.

Легко видеть, что произведения (12), (13) содержат $k+1$ сомножителей, в которых участвуют последовательно вычисляемые числа $x^2, x^4, \dots, x^{2^k}$, $\widetilde{x}^2, \widetilde{x}^4, \dots, \widetilde{x}^{2^k}$, каждое последующее из которых является квадратом предыдущего. При этом на первом шаге этого процесса получаем

$$y = x^2 = \\ = 2^{-3}\left(\beta_{2n-3} + \beta_{2n-4}2^{-1} + \dots + \beta_m 2^{m-2n+3} + \dots + \beta_1 2^{4-2n} + \beta_0 2^{3-2n}\right), \tag{15}$$

$$\widetilde{y} = \widetilde{x}^2 = 2^{-5}\left(\widetilde{\beta}_{2n-1} + \widetilde{\beta}_{2n-2}2^{-1} + \dots + \widetilde{\beta}_1 2^{-2n+2} + \widetilde{\beta}_0 2^{-2n+1}\right), \tag{16}$$

где β_i, $\widetilde{\beta}_j$ равны 0 или 1. Ясно, что для вычисления $y = x^2$ из (15) достаточно $O(M(n-2))$ операций, для вычисления же $\widetilde{y} = \widetilde{x}^2$, $\widetilde{x} = (\widetilde{\widetilde{x}} - 1)^2 = \widetilde{\widetilde{x}}^2 + 1 - 2\widetilde{\widetilde{x}}$ из (16) достаточно $O(M(n-1))$ операций. На втором шаге процесса вычисляются квадраты чисел (15), (16), однако вместо полученных величин

$$u = y^2 = 2^{-5}\left(\gamma_0 2^{8-4n} + \dots + \gamma_{4n-8}\right), \\ \widetilde{u} = \widetilde{y}^2 = 2^{-9}\left(\widetilde{\gamma}_0 2^{8-4n} + \dots + \widetilde{\gamma}_{4n-8}\right), \quad n \geqslant 2,$$

отнеся к остаточному члену малые слагаемые, запишем эти числа в виде

$$u = y^2 = 2^{-5}\left(\delta_0 + \delta_1 2^{-1} + \dots + \delta_r 2^{-r_1} + \theta_1 2^{-r_1}\right) = \\ = 2^{-5}\left(\delta_0 + \delta_1 2^{-1} + \dots + \delta_r 2^{-r_1}\right) + \theta_1 2^{-(r_1+5)}, \tag{17}$$

$$\widetilde{u} = \widetilde{y}^2 = 2^{-9}\left(\widetilde{\delta}_0 + \widetilde{\delta}_1 2^{-1} + \dots + \widetilde{\delta}_r 2^{-r_1} + \widetilde{\theta}_1 2^{-r_1}\right) = \\ = 2^{-9}\left(\widetilde{\delta}_0 + \widetilde{\delta}_1 2^{-1} + \dots + \widetilde{\delta}_r 2^{-r_1}\right) + \widetilde{\theta}_1 2^{-(r_1+9)}, \tag{18}$$

где $|\theta_1| < 1$, $|\widetilde{\theta}_1| < 1$, а δ_ℓ, $\widetilde{\delta}_i$ равны 0 или 1. И так далее. На j-м шаге будем иметь

$$w = v^2 = 2^{-2^j-1}\left(\varepsilon_0 + \varepsilon_1 2^{-1} + \dots + \varepsilon_r 2^{-r_j} + \theta_{j-1}2^{-r_j}\right), \quad |\theta_{j-1}| < 1, \\ \widetilde{w} = \widetilde{v}^2 = 2^{-2^{j+1}-1}\left(\widetilde{\varepsilon}_0 + \widetilde{\varepsilon}_1 2^{-1} + \dots + \widetilde{\varepsilon}_r 2^{-r_j} + \widetilde{\theta}_{j-1}2^{-r_j}\right), \quad |\widetilde{\theta}_{j-1}| < 1,$$

где ε_ℓ, $\widetilde{\varepsilon}_i$ равны 0 или 1, и

$$r_j + 2^j + 1 = r \quad \text{или} \quad r_j + 2^{j+1} + 1 = r. \tag{19}$$

Принимая во внимание, что

$$(a_1+\theta_1)(a_2+\theta_2)\dots(a_n+\theta_n) = a_1a_2\dots a_n\left(1+\frac{\theta_1}{a_1}\right)\left(1+\frac{\theta_2}{a_2}\right)\dots\left(1+\frac{\theta_n}{a_n}\right), \\ (1+\alpha)(1+\beta) \approx 1+\alpha+\beta,$$

можно сделать вывод о том, что если $r = 2n + \log 2n$, то вычисляя квадраты (17), (18) и последующие квадраты получаемых значений $v = u^2$, $\widetilde{v} = \widetilde{u}^2$, $w = v^2$, $\widetilde{w} = \widetilde{v}^2$, … с точностью $2^{-2n-\log 2n}$, в результате получим произведения (12), (13) заведомо с точностью 2^{-2n-1}. В то же время, вынесенные за скобки в (15), (16) и (17), (18) и далее на каждом шаге $1, 2, \dots, j, \dots$ множители 2^{-2^j-1}, $2^{-2^{j+1}-1}$ сокращают значность возводимых в квадрат чисел. Учитывая (19), для вычисления последовательных квадратов на шаге j, $2 \leqslant j \leqslant k$, достаточно потратить

$$M(2n + \log 2n - 2 - 2^j) + O(2n + \log 2n)$$

операций. Суммируя число операций по всем шагам, получаем

$$s_{S_n}(n) = O\left(\log n(2n + \log 2n) + \sum_{j=1}^{k} M(2n + \log 2n - 2 - 2^j) + M(n)\right).$$

Ясно, что оценка сложности вычисления $\widetilde{S}_n$ будет асимптотически такая же. Отсюда и из (9)–(11), (14) и (6) следует, что доказано следующее утверждение.

Теорема 1. *Для сложности вычисления числа, обратного заданному n-значному числу z, справедлива оценка*

$$s_{1/z}(n) = O(M(n)\log n). \tag{20}$$

Замечание 1. Можно построить аналогичный быстрый алгоритм вычисления $1/z$ на основе метода А.А. Карацубы (см. [3]), разбивая z на две части "посередине":

$$\frac{1}{x_0 + 2x_1 + \ldots + 2^{n-1}x_{n-1}} = \frac{1}{\alpha_0 + 2^{n_1}\alpha_1} = 2^{-n_1}\alpha_1 \frac{1}{1 + 2^{-n_1}\dfrac{\alpha_0}{\alpha_1}} =$$
$$= 2^{-n_1}\frac{1}{\alpha_1}\left(1 - 2^{-n_1}\frac{\alpha_0}{\alpha_1} + 2^{-2n_1}\left(\frac{\alpha_0}{\alpha_1}\right)^2\theta\right), \quad |\theta| \leqslant 1,$$

$0 \leqslant \dfrac{\alpha_0}{\alpha_1} < 2$, $n - 1 = 2^k$, $n_1 = \dfrac{n-1}{2} = 2^{k-1}$. Далее таким же образом разбиваются на части числа α_1, и так далее.

Замечание 2. Если в БВЕ-алгоритмах (см. [6–9]) вместо вычисления делений методом Ньютона использовать вышепредставленный метод деления, то общая заявленная сложность алгоритмов, основанных на БВЕ, не изменится. Причина в том, что метод БВЕ – это метод быстрого суммирования рядов, большинство из которых имеет вид

$$f_1 = f_1(z) = \sum_{j=0}^{\infty} \frac{a(j)}{b(j)} z^j \tag{21}$$

при условии, что $a(j)$, $b(j)$ — целые числа, $|a(j)| + |b(j)| \leqslant (Cj)^K$, $|z| < 1$, K и C – константы, и z – алгебраическое число, причем сложность вычисления таких рядов посредством БВЕ равна

$$s_{f_1}(n) = O(M(n)\log^2 n), \tag{22}$$

т.е. превышает (20) в $\log n$ раз, и даже если деление с точностью 2^{-n} будет производиться на каждом шаге (которых в методе БВЕ всегда $\sim \log n$), это никак не поменяет оценку (22), как и вычисление на последнем шаге обратного значения числа размером $n \log n$.

Только при суммировании посредством БВЕ рядов вида

$$f_2 = f_2(z) = \sum_{j=0}^{\infty} \frac{a(j)}{b(j)} \frac{z^j}{j!},$$

при условии, что $a(j)$, $b(j)$ — целые числа, $|a(j)|+|b(j)| \leqslant (Cj)^K$, $|z| < 1$, K и C – константы, и z – алгебраическое число (примером является БВЕ-вычисление числа e), сложность вычисления будет такой же, как и в (20), а именно

$$s_{f_2}(n) = O(M(n)\log n). \tag{23}$$

Однако такая пониженная сложность достигается за счет сверхбыстрой сходимости, которая позволяет достичь нужной точности 2^{-n} при суммировании не $r \sim n$, а

$$r \sim \frac{n}{\log n} \tag{24}$$

членов заданного ряда. При этом на последнем шаге происходит вычисление обратного значения числа размера $\sim r \log r$ со сложностью $\left(r \log^3 \log \log r\right)$, но с учетом (24) это не меняет оценку (23).

Замечание 3. Может показаться, что построенный выше БВЕ-алгоритм вычисления значения $1/z$ имеет сложность хуже, чем алгоритм Ньютона (см. [13–15]). Однако, как оказалось, в [13–15] сложность вычисления $1/z$ подсчитывалась по-другому. Рассмотрим ньютоновский процесс. Значение $S = 1/z$,

$$\frac{1}{2} \leqslant z \leqslant 1, \tag{25}$$

вычисляется по формуле

$$S_{2n} = -zS_n^2 + 2S_n, \quad S_1 = \frac{3}{2}. \tag{26}$$

При этом

$$\text{если} \quad |S - S_n| \leqslant 2^{-n}, \quad \text{то} \quad |S - S_{2n}| \leqslant 2^{-2n}. \tag{27}$$

Из (27) легко видеть, что вычисление по формуле (26) от S_1 до S_n происходит за $\log n$ шагов. При этом, если z – n-значное число, скажем,

$$z = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^7} + \frac{1}{2^{11}} + \ldots + \frac{1}{2^n},$$

то удовлетворяя, очевидно, условию (25), оно в то же время требует $\sim n$ знаков для своей записи. Начиная со второго шага, согласно формуле (26), это число будет возводиться в квадрат. Таким образом, сложность ньютоновского алгоритма вычисления значения $1/z$, где z – n-значное число, с точностью до n знаков будет иметь ту же сложность (20), что и БВЕ-алгоритм.

§ 3. Алгоритм быстрого вычисления рациональной функции

Вычисление функции $y = z^{r/m}$, $(r, m) = 1$, r, m – натуральные числа, $m \geqslant 2$, с точностью 2^{-n} сводится к извлечению корня m-й степени из числа z с точностью до n знаков с последующим перемножением r полученных значений корня с той же точностью. Поскольку r – фиксированная константа, сложность вычисления произведения r чисел с точностью 2^{-n} оценивается как $O(M(n))$. Таким образом, задача сводится к вычислению с точностью 2^{-n} значения $y = z^{1/m}$.

Для удобства считаем, что $n = 2^k$, $k \geqslant 1$. Как и по методу А.А. Карацубы (см. [1–3]), представим n-значное число z в виде

$$\begin{aligned} z &= z_0 + 2z_1 + \ldots + 2^{n_1-1} z_{n_1-1} + 2^{n_1}\left(z_{n_1} + 2z_{n_1+1} + \ldots + 2^{n_1-1} z_n\right) = \\ &= \alpha_{10} + 2^{n_1}\alpha_{11}, \quad n_1 = \frac{n}{2} = 2^{k-1}, \end{aligned} \tag{28}$$

$$z^{1/m} = \left(\alpha_{10} + 2^{n_1}\alpha_{11}\right)^{1/m} = 2^{n_1/m}\alpha_{11}^{1/m}\left(1 + 2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^{\frac{1}{m}}, \quad 0 \leqslant \frac{\alpha_{10}}{\alpha_{11}} < 2, \tag{29}$$

где α_{10}, α_{11} являются n_1-значными числами. Согласно биному Ньютона

$$\left(1+2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^{1/m}=1+\binom{1/m}{1}2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}+\binom{1/m}{2}2^{-2n_1}\left(\frac{\alpha_{10}}{\alpha_{11}}\right)^2+$$
$$+\theta_1\binom{1/m}{3}2^{-3n_1}\left(\frac{\alpha_{10}}{\alpha_{11}}\right)^3,\quad |\theta_1|\leqslant 1. \tag{30}$$

Здесь

$$\binom{1/m}{j}=\frac{\frac{1}{m}\left(\frac{1}{m}-1\right)\dots\left(\frac{1}{m}-j+1\right)}{j!},\quad \left|\binom{1/m}{j}\right|<\frac{1}{jm},\quad j\geqslant 2.$$

На первом шаге с вынесением "очевидного общего множителя", как в методе БВЕ (см. [6–9]), сумма двух чисел из (30) преобразуется следующим образом:

$$\binom{1/m}{1}2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}+\binom{1/m}{2}2^{-2n_1}\left(\frac{\alpha_{10}}{\alpha_{11}}\right)^2=\frac{1}{m}2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}-\frac{m-1}{2m^2}2^{-2n_1}\left(\frac{\alpha_{10}}{\alpha_{11}}\right)^2=$$
$$=\frac{2^{-2n_1}}{2m^2}\frac{\alpha_{10}}{\alpha_{11}^2}\left(2^{n_1+1}m\alpha_{11}-(m-1)\alpha_{10}\right)=\frac{2^{-2n_1}}{2m^2}\frac{\alpha_{10}}{\alpha_{11}^2}\beta_1,\quad n_1=2^{k-1},$$

при этом вычисляется целое число β:

$$\beta=2^{n_1+1}m\alpha_{11}-(m-1)\alpha_{10}.$$

На втором шаге преобразуем $\alpha_{11}^{1/m}$ в формуле (29), где α_{11} является n_1-значным числом ($n_1=n/2$), представив это значение в виде

$$\alpha_{11}=\alpha_{20}+2^{n_2}\alpha_{21},\quad n_2=\frac{n_1}{2}=\frac{n}{4},$$

где α_{20}, α_{21} – n_2-значные числа. Следовательно,

$$\alpha_{11}^{1/m}=\left(\alpha_{20}+2^{n_2}\alpha_{21}\right)^{1/m}=2^{n_2/m}\alpha_{21}^{1/m}\left(1+2^{-n_2}\frac{\alpha_{20}}{\alpha_{21}}\right)^{1/m},$$

и учитывая, что $0\leqslant\dfrac{\alpha_{20}}{\alpha_{21}}<2$, отсюда получаем

$$\left(1+2^{-n_2}\frac{\alpha_{20}}{\alpha_{21}}\right)^{1/m}=1+\binom{1/m}{1}2^{-n_2}\frac{\alpha_{20}}{\alpha_{21}}+\binom{1/m}{2}2^{-2n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^2+$$
$$+\binom{1/m}{3}2^{-3n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^3+\binom{1/m}{4}2^{-4n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^4+$$
$$+\theta_2 2^{-5n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^5,\quad |\theta_2|\leqslant 1. \tag{31}$$

На втором шаге методом БВЕ [6] преобразуем сумму четырех слагаемых из (31):

$$\binom{1/m}{1}2^{-n_2}\frac{\alpha_{20}}{\alpha_{21}}+\binom{1/m}{2}2^{-2n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^2+\binom{1/m}{3}2^{-3n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^3+$$
$$+\binom{1/m}{4}2^{-4n_2}\left(\frac{\alpha_{20}}{\alpha_{21}}\right)^4=\frac{2^{-2n_2}\alpha_{20}}{2m^2\alpha_{21}^2}\gamma_1+\frac{2^{-4n_2}(m-1)(2m-1)\alpha_{20}^3}{4!\,m^4\alpha_{21}^4}\gamma_2=$$
$$=\frac{2^{-4n_2}\alpha_{20}}{4!\,m^4\alpha_{21}^4}\beta_2,$$

и вычисляем целые числа

$$\gamma_1 = 2^{n_2+1}m\alpha_{21} - (m-1)\alpha_{20}, \quad \gamma_2 = 2^{n_2+2}m\alpha_{21} - (3m-1)\alpha_{20},$$
$$\beta_2 = 3\cdot 4\cdot 2^{2n_2}m^2\alpha_{21}^2\gamma_1 + (m-1)(2m-1)\alpha_{20}^2\gamma_2.$$

И так далее. На j-м шаге ($j \leqslant k$) имеем

$$\alpha_{j-1,j-1}^{1/m} = (\alpha_{j0} + 2^{n_j}\alpha_{j1})^{1/m} = 2^{n_j/m}\alpha_{j1}^{1/m}\left(1 + 2^{-n_j}\frac{\alpha_{j0}}{\alpha_{j1}}\right)^{1/m}, \quad 0 \leqslant \frac{\alpha_{j0}}{\alpha_{j1}} < 2, \tag{32}$$

и отсюда

$$\left(1 + 2^{-n_j}\frac{\alpha_{j0}}{\alpha_{j1}}\right)^{1/m} = 1 + \binom{1/m}{1}2^{-n_j}\frac{\alpha_{j0}}{\alpha_{j1}} + \binom{1/m}{2}2^{-2n_j}\left(\frac{\alpha_{j0}}{\alpha_{j1}}\right)^2 +$$
$$+ \binom{1/m}{3}2^{-3n_j}\left(\frac{\alpha_{j0}}{\alpha_{j1}}\right)^3 + \ldots + \binom{1/m}{2^{2^j}}2^{-2^j n_j}\left(\frac{\alpha_{j0}}{\alpha_{j1}}\right)^{2^j} +$$
$$+ \theta_j 2^{-(2^j+1)n_j}\left(\frac{\alpha_{j0}}{\alpha_{j1}}\right)^{2^j+1}, \quad |\theta_j| \leqslant 1. \tag{33}$$

На j-м шаге ($j \leqslant k$) методом БВЕ вычисляем сумму 2^j слагаемых из (33). И так далее. Процесс завершается на k-м шаге ($n = 2^k$, $n_k = 1$). Таким образом, общая конструкция такова:

$$z^{1/m} = 2^{\frac{n_1+n_2+\ldots+n_k}{m}}\left(1 + 2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^{1/m}\left(1 + 2^{-n_2}\frac{\alpha_{20}}{\alpha_{21}}\right)^{1/m}\ldots$$
$$\ldots\left(1 + 2^{-n_j}\frac{\alpha_{j0}}{\alpha_{j1}}\right)^{1/m}\ldots\left(1 + 2^{-n_k}\frac{\alpha_{k0}}{\alpha_{k1}}\right)^{1/m} =$$
$$= 2^{\frac{n_1+n_2+\ldots+n_k}{m}}\left(1 + \binom{1/m}{1}2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}} + \ldots\right)\times$$
$$\times\left(1 + \binom{1/m}{1}2^{-n_2}\frac{\alpha_{20}}{\alpha_{21}} + \ldots\right)\ldots\left(1 + \binom{1/m}{1}2^{-1}\frac{\alpha_{k0}}{\alpha_{k1}} + \ldots\right) =$$
$$= 2^{\frac{n_1+n_2+\ldots+n_k}{m}}\left(1 + \frac{2^{-2n_1}}{2!\,m^2}\frac{\alpha_{10}}{\alpha_{11}^2}\beta_1\right)\left(1 + \frac{2^{-4n_2}\alpha_{20}}{4!\,m^4\alpha_{21}^4}\beta_2\right)\ldots, \tag{34}$$

и произведение скобок вычисляется методом БВЕ. Кроме того, нужно вычислить быстро первый множитель в формуле (34) – число $2^{\frac{n_1+n_2+\ldots+n_k}{m}}$. Это можно сделать двумя способами. Можно подобрать значение n таким образом, чтобы оно превышало заданную точность вычисления, и помимо того что $n = 2^k$, число $2^{k-1} + 2^{k-2} + \ldots + 2 = 2^k - 2$ было бы кратно m. Сложность операции возведения двойки в степень n составляет $O(n)$. Другой способ – быстрое вычисление значения $2^{1/m}$, $m \geqslant 2$. Представим сначала число $a = \left(\frac{1}{2}\right)^{1/m}$ в виде ряда

$$a = \left(1 - \frac{1}{2}\right)^{1/m} = 1 + \sum_{j=1}^{\infty}(-1)^j\frac{\frac{1}{m}\left(\frac{1}{m}-1\right)\ldots\left(\frac{1}{m}-j+1\right)}{j!}\left(\frac{1}{2}\right)^j \tag{35}$$

и просуммируем (35) с точностью 2^{-n-1} с помощью БВЕ-процесса со сложностью вычисления

$$s_{(1/2)^{1/m}} = O(M(n)\log n).$$

Затем вычислим с точностью до $n+1$ знаков значение, обратное к a, посредством БВЕ-алгоритма из предыдущего параграфа со сложностью (20). С учетом сложности деления (20) для чисел α_{ji}, α_{jl} и их степеней и роста размера чисел, общая сложность вычисления равна

$$s_{z^{1/m}} = O(M(n)\log^2 n).$$

Тем самым, доказана

Теорема 2. *Для сложности вычисления рациональной функции* $y = z^{r/m}$, $(r,m)=1$, r, m – *натуральные числа*, $m \geqslant 2$, *справедлива оценка*

$$s_{z^{r/m}}(n) = O(M(n)\log^2 n). \tag{36}$$

Замечание 4. Если в БВЕ-алгоритмах (см. [6–9]) вместо вычисления рациональных функций методом Ньютона использовать вышепредставленный метод вычисления этих функций, то общая заявленная сложность алгоритмов, основанных на БВЕ, не изменится. В алгоритмах со сложностью (23) рациональная функция не участвует, в алгоритмах со сложностью (22), если и встречается вычисление рациональной функции, то конечное количество раз, что не меняет окончательную оценку сложности вычисления.

Замечание 5. Построенный алгоритм быстрого вычисления рациональных функций является "гибридным", поскольку совмещает в себе две идеи – метода А.А. Карацубы 1960 г. и метода БВЕ 1990 г., но далее мы будем называть его просто БВЕ.

Замечание 6. Как и выше (см. замечание 3) сравним представленный БВЕ-алгоритм вычисления значения $z^{1/m}$ и алгоритм Ньютона (см. [13–15]). Как оказалось, в [13–15] сложность вычисления значения $z^{1/m}$ определялась не так, как в настоящей статье. Рассмотрим ньютоновский процесс. Значение $S = z^{1/m}$, $m \geqslant 2$, при $z \geqslant (m+1)^m$ вычисляется по формуле

$$S_{2n} = \frac{m+1}{m}S_n - \frac{S_n^{m+1}}{mz}, \tag{37}$$

причем

$$\text{если} \quad |S - S_n| \leqslant 2^{-n}, \quad \text{то} \quad |S - S_{2n}| \leqslant 2^{-2n}. \tag{38}$$

Из (38) легко видеть, что вычисление по формуле (37) от S_1 до S_n происходит за $\log n$ шагов. Если на каждом шаге вычислять соответствующие обратные значения n-значных чисел, то общая сложность алгоритма будет та же, что и для БВЕ, т.е. (36). Однако для ньютоновского процесса (37) обратное значение n-значного числа z, а вернее, значение $\alpha = 1/mz$, $m = \text{const}$, можно вычислить один раз (на первом шаге), а затем вычислять только произведения $\alpha, \alpha^2, \ldots, C\alpha$, $C = \text{const}$ (с нужной точностью). За S_1 берут одно из двух целых чисел M или $M+1$, таких что $M^m < z \leqslant (M+1)^m$, чтобы либо $|M - z^{1/m}|$, либо $|M+1-z^{1/m}|$ не превосходило $1/2$. Согласно (37) на каждом шаге j вычисляется произведение n-значных чисел (во втором слагаемом из (37)), одно из которых $1/mz$, а другое –

$$\left(\frac{m+1}{m}S_{j-1} - S_{j-1}^{m+1}\frac{1}{mz}\right)^{m+1},$$

которое представляет собой n-значное число, возведенное в степень $m+1$. Поскольку $m+1$ – постоянная, то сложность всех таких произведений на каждом шаге $j = 2, 3, \ldots$ составляет $O(M(n))$, и следовательно, сложность вычисления рациональной функции с помощью алгоритма Ньютона есть $O(M(n)\log n)$, т.е. в $\log n$ раз лучше, чем сложность вышепредставленного БВЕ-алгоритма вычисления $z^{1/m}$.

Замечание 7. В дальнейшем планируется разработать простую БВЕ-конструкцию вычисления рациональной функции с оценкой сложности $O(M(n)\log n)$.

§ 4. Быстрое вычисление обратных функций

Заметим, что алгоритмы быстрого деления и быстрого вычисления рациональной функции, представленные выше, обобщаются на случай комплексного аргумента вычисляемой функции. В этом случае аппроксимации должны быть построены как для вещественной, так и для комплексной частей аргумента, и произведут, соответственно, вещественную и комплексную части значения функции. Сложность вычисления удвоится, т.е. оценки сложности (20), (36) не изменятся.

Поскольку все обратные гиперболические и обратные тригонометрические функции выражаются через логарифм, в последнем случае комплексный и с комплексным аргументом, как, например,

$$\operatorname{arctg} z = \frac{1}{2i}\ln\frac{1+iz}{1-iz} = \frac{1}{2i}\ln\left(\frac{1-z^2}{1+z^2} + i\frac{2z}{1+z^2}\right),$$

то важно построить быстрый алгоритм для вычисления логарифма, а вычисления других обратных тригонометрических и гиперболических функций можно свести к вычислению логарифма, подобно тому как в [6] вычисление тригонометрических функций косинуса и синуса сводится к вычислению вещественной и мнимой частей экспоненциальной функции с мнимым аргументом. Заметим, что вычисление логарифма с комплексным аргументом использует известные значения $\ln(i)$, $\ln(-1)$, $\ln(-i)$ и разложение комплексного логарифма в ряд вблизи известных значений.

Пусть нужно вычислить натуральный логарифм $y = \log z$, где z – n-значное число, $z = z_0 + 2z_1 + \ldots + 2^{n-1}z_{n-1}$, с точностью 2^{-n}. Предполагая для простоты, что $n = 2^k$, подобно тому как это делается по методу А.А. Карацубы, разобьем аргумент логарифма z на две части, как в формуле (28). Тогда

$$\log z = \log 2^{n_1}\alpha_{11} + \log\left(1 + 2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right) = n_1\log 2 + \log\alpha_{11} + a_1, \tag{39}$$

где

$$a_1 = \log\left(1 + 2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right), \tag{40}$$

а α_{10}, α_{11} являются n_1-значными числами. Поскольку

$$0 \leqslant \frac{\alpha_{10}}{\alpha_{11}} < 2, \quad n_1 = \frac{n}{2} \geqslant 1, \quad \left|2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right| < 1,$$

то справедливо следующее разложение в ряд Тейлора:

$$a_1 = 2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}} - \frac{1}{2}\left(2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^2 + \frac{1}{3}\left(2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^3 - \frac{1}{4}\left(2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^4 +$$
$$+ \theta_1\frac{1}{5}\left(2^{-n_1}\frac{\alpha_{10}}{\alpha_{11}}\right)^5, \quad |\theta_1| \leqslant 1. \tag{41}$$

На первом шаге методом БВЕ вычисляется сумма первых четырех слагаемых из (41). На втором шаге аналогично разбиваем на две части $n_1 = \dfrac{n}{2} = 2^{k-1}$-значное

число α_{11}, получая для $\log \alpha_{11}$ выражение

$$\log \alpha_{11} = n_2 \log 2 + \log \alpha_{21} + a_2, \quad a_2 = \log \left(1 + 2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right), \tag{42}$$

где α_{20}, α_{21} являются n_2-значными числами, $n_2 = \frac{n_1}{2} = \frac{n}{4} = 2^{k-2}$. Поскольку

$$0 \leqslant \frac{\alpha_{20}}{\alpha_{21}} < 2, \quad n_2 = \frac{n}{4} \geqslant 1, \quad \left|2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right| < 1,$$

то справедливо следующее разложение в ряд Тейлора:

$$\begin{aligned} a_2 = {} & 2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}} - \frac{1}{2} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^2 + \frac{1}{3} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^3 - \frac{1}{4} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^4 + \\ & + \frac{1}{5} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^5 - \frac{1}{6} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^6 + \frac{1}{7} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^7 - \frac{1}{8} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^8 + \\ & + \theta_2 \frac{1}{9} \left(2^{-n_2} \frac{\alpha_{20}}{\alpha_{21}}\right)^9, \quad |\theta_2| \leqslant 1. \end{aligned} \tag{43}$$

На втором шаге методом БВЕ вычисляется сумма первых восьми слагаемых из (43). И так далее. На j-м шаге ($1 \leqslant j \leqslant k$) имеем

$$\log \alpha_{j-1,1} = n_j \log 2 + \log \alpha_{j1} + a_j, \tag{44}$$

и 2^{j+1} первых слагаемых a_j суммы

$$\begin{aligned} a_j = {} & 2^{-n_j} \frac{\alpha_{j0}}{\alpha_{j1}} - \frac{1}{2} \left(2^{-n_j} \frac{\alpha_{j0}}{\alpha_{j1}}\right)^2 + \frac{1}{3} \left(2^{-n_j} \frac{\alpha_{j0}}{\alpha_{j1}}\right)^3 - \ldots - \frac{1}{2^{j+1}} \left(2^{-n_j} \frac{\alpha_{j0}}{\alpha_{j1}}\right)^{2^{j+1}} + \\ & + \frac{\theta_j}{2^{j+1}+1} \left(2^{-n_j} \frac{\alpha_{j0}}{\alpha_{j1}}\right)^{2^{j+1}+1}, \quad |\theta_j| \leqslant 1, \end{aligned} \tag{45}$$

суммируются методом БВЕ. Из (45) на вычисление всех слагаемых a_j по всем шагам j, $1 \leqslant j \leqslant k$, посредством БВЕ, достаточно

$$O(M(n) \log^2 n) \tag{46}$$

операций. Из (39), (42) и (44) следует, что кроме a_j, $1 \leqslant j \leqslant k$, нужно также вычислить

$$b = \sum_{j=1}^{k} n_j \log 2 = \log 2 \sum_{j=1}^{k} 2^{k-j} = (n-1) \log 2,$$

для чего нужно вычислить с точностью до n знаков значение $\log 2$. Воспользовавшись известным разложением (см., например, [16])

$$\log p = \log q + 2 \left(\frac{p-q}{p+q} + \frac{1}{3} \left(\frac{p-q}{p+q}\right)^3 + \frac{1}{5} \left(\frac{p-q}{p+q}\right)^5 + \ldots\right),$$

где p, q – натуральные числа, получаем

$$\log 2 = \frac{2}{3} \left(1 + \frac{1}{3} \left(\frac{1}{3}\right)^2 + \frac{1}{5} \left(\frac{1}{3}\right)^4 + \ldots\right),$$

т.е. ряд вида (21), для сложности суммирования которого справедлива оценка (22). Учитывая (39), (40), (44), (45), а также (22) и (46), получаем доказательство следующего утверждения.

Теорема 3. *Для сложности вычисления логарифмической функции* $y = \log z$ *справедлива оценка*

$$s_{\log z}(n) = O(M(n)\log^2 n).$$

Замечание 8. Построенный алгоритм быстрого вычисления логарифмической функции также является "гибридным", совмещая особенности метода А.А. Карацубы 1960 г. и метода БВЕ 1990 г.

Замечание 9. Первый "гибридный" алгоритм быстрого вычисления логарифмической функции построил С.В. Яхонтов (см. [17]), при этом метод А.А. Карацубы 1960 г. упоминается в [17] как "Binary Splitting".

§ 5. Заключение

Метод БВЕ получил известность (см. [18]) раньше, чем получил свое наименование. Настоящая статья показывает, что метод БВЕ (или БВЕ совместно с методом А.А. Карацубы) подобно методу Ньютона пригоден также для построения быстрых алгоритмов вычисления элементарных алгебраических функций, а также обратных функций. При этом в построенных алгоритмах сохраняется возможность частичного распараллеливания, что является особой характеристикой метода БВЕ.

СПИСОК ЛИТЕРАТУРЫ

1. *Карацуба А., Офман Ю.* Умножение многозначных чисел на автоматах // ДАН СССР. 1962. Т. 145. № 2. С. 293–294. http://mi.mathnet.ru/dan26729
2. *Dynkin E.B., Kolmogorov A.N., Kostrikin A.I., Pjateckiĭ-Šapiro I.I., Šafarevič I.R., Sinaĭ Ja.G.* Six Lectures Delivered at the International Congress of Mathematicians in Stockholm, 1962. Providence, R.I.: Amer. Math. Soc., 1963.
3. *Карацуба А.А.* Сложность вычислений // Оптимальное управление и дифференциальные уравнения. Сб. статей. Тр. МИАН. 1995. Т. 211. С. 186–202. http://mi.mathnet.ru/tm1120
4. *Schönhage A., Strassen V.* Schnelle Multiplikation großer Zahlen // Computing. 1971. V. 7. № 3–4. P. 281–292. https://doi.org/10.1007/BF02242355
5. *Fürer M.* Faster Integer Multiplication // SIAM J. Comput. 2009. V. 39. № 3. P. 979–1005. https://doi.org/10.1137/070711761
6. *Карацуба Е.А.* Быстрые вычисления трансцендентных функций // Пробл. передачи информ. 1991. Т. 27. № 4. С. 76–99. http://mi.mathnet.ru/ppi584
7. *Карацуба Е.А.* Быстрое вычисление дзета-функции Римана $\zeta(s)$ при целых значениях аргумента s // Пробл. передачи информ. 1995. Т. 31. № 4. С. 69–80. http://mi.mathnet.ru/ppi294
8. *Karatsuba E.A.* Fast Computation of Some Special Integrals of Mathematical Physics // Scientific Computing, Validated Numerics, Interval Methods. Boston: Springer, 2001. P. 29–41. https://doi.org/10.1007/978-1-4757-6484-0_3
9. *Карацуба Е.А.* О вычислении функции Бесселя путем суммирования рядов // Сиб. журн. вычисл. матем. 2019. Т. 22. № 4. С. 453–472. https://doi.org/10.15372/SJNM20190405
10. *Salamin E.* Computation of π Using Arithmetic-Geometric Mean // Math. Comp. 1976. V. 30. № 135. P. 565–570. https://doi.org/10.2307/2005327
11. *Carlson B.C.* Algorithms Involving Arithmetic and Geometric Means // Amer. Math. Monthly. 1971. V. 78. № 5. P. 496–505. https://doi.org/10.2307/2317754

12. *Borwein J.M., Borwein P.B.* Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity. New York: Wiley, 1987.
13. *Cook S.A.* On the Minimum Computation Time of Functions. Doct. Thesis. Harvard Univ., Cambridge, MA, USA, 1966.
14. *Бендерский Ю.В.* Быстрые вычисления // ДАН СССР. 1975. Т. 223. № 5. С. 1041–1043. http://mi.mathnet.ru/dan39204
15. *Brent R.P.* Fast Multiple-Precision Evaluation of Elementary Functions // J. ACM. 1976. V. 23. № 2. P. 242–251. https://doi.org/10.1145/321941.321944
16. *Фихтенгольц Г.М.* Курс дифференциального и интегрального исчисления. Т. 2. М.: Физматгиз, 1959.
17. *Яхонтов С.В.* Эффективное по времени и памяти вычисление логарифмической функции вещественного аргумента на машине Шёнхаге // Прикл. дискр. матем. 2013. № 2 (20). С. 101–114. http://mi.mathnet.ru/pdm414
18. *Lozier D.W., Olver F.W.J.* Numerical Evaluation of Special Functions // Mathematics of Computation 1943–1993: A Half-Century of Computational Mathematics (Proc. 48th Symp. in Applied Mathematics. Aug. 9–13, 1993. Vancouver, B.C., Canada). Providence, R.I.: Amer. Math. Soc., 1995. P. 79–125. https://doi.org/10.1090/psapm/048/1314844

Карацуба Екатерина Анатольевна
Вычислительный центр им. А.А. Дородницына
Федерального исследовательского центра
"Информатика и управление" РАН, Москва
ekaratsuba@gmail.com

Поступила в редакцию
13.06.2022
После доработки
27.07.2022
Принята к публикации
27.07.2022

Адрес редакции: 127051, Москва, Б. Каретный пер., 19, стр. 1, тел. (495) 650-47-39

Оригинал-макет подготовил *Д.Ю. Ногин*
по контракту с ООО «Тематическая редакция»

Москва
ООО «Тематическая редакция»