KYBERNETIKA ČÍSLO 2, ROČNÍK 2/1966

Некоторые математические вопросы теории передачи информации*

М. С. Пинскер

Этот доклад посвящен обзору вероятностных аспектов некоторых направлений теории передачи информации. В первой части обзора будут изложены отдельные результаты по обоснованию Шенноновской теории оптимального кодирования информации, носящие окончательный характер. Главное внимание уделяется проблеме построения конкретных простых способов кодирования и декодирования, поскольку основные усилия исследователей сосредоточены в настоящее время на разработке эффективных методов передачи информации. В ходе изложения сформулировано несколько актуальных, как нам кажется, проблем. Естественно, что содержание доклада тесно переплетается с работами автора.

1. ВВЕДЕНИЕ

Полное и математически четкое изложение центральных результатов теории оптимального кодирования информации весьма громоздко [3, 4]. Поэтому для того, чтобы не перегружать изложения мы в §§ 1 и 2 дадим лишь беглый обзор основных понятий и результатов (см. [3]) теории. Существо интересующих нас результатов мы будем пояснять на отдельных примерах, представляющих, однако, весьма общие ситуации передачи информации. Перенесение полученных результатов на более общие случаи не представит труда для читателя — математика.

Шенноновская модель системы передачи информации исходит из следующей схемы:

$$\xi \xrightarrow[\text{кодирование сообщений}]{} \eta \xrightarrow[\text{передача по каналу}]{} \tilde{\eta} \xrightarrow[\text{декодирование сообщений}]{} \tilde{\xi} \tag{1}$$

Здесь ξ — сообщение, генерируемое источником, есть случайная величина, при-

* Доклад, прочитанный на *Летним семинаре по теории информации и статистическим методам автоматического управления*, г. Прага 25 мая — 4 июня 1965 г.

нимающая значения x в пространстве X. Значениями сообщения ξ являются обычно реализациями случайного процесса $\xi(t)$ на некотором временном полуинтервале; этими реализациями могут быть, например, последовательности чисел $x = (x_1 \dots x_m)$; функции непрерывного аргумента $x = \{x(t); 0 < t \leqq T\}$; функции нескольких аргументов (кадры телевизионного изображения) и т.п. Переход $\xi \to \eta$ есть кодирование (преобразование) сообщения ξ в сигнал η на входе канала, принимающий значения y в пространстве Y. Значениями сигнала η могут быть также последовательности чисел, функции непрерывного аргумента и т.п. Переход $\eta \to \tilde{\eta}$ есть передача сигнала по каналу, на выходе которого образуется сигнал $\tilde{\eta}$, принимающий значения $\tilde{y}$ в пространстве $\tilde{Y}$. Переход от входного сигнала η к выходному сигналу $\tilde{\eta}$ подчиняется условному распределению вероятностей

$$P_{\tilde{\eta}|y}(\tilde{B} \mid y) = P_{\tilde{\eta}|\eta}(\tilde{B} \mid y) = P\{\tilde{\eta} \in \tilde{B} \mid \eta = y\}, \quad \tilde{B} \subset \tilde{Y} \tag{2}$$

(вероятность того, что значения сигнала на выходе $\tilde{\eta}$ принадлежат множеству $\tilde{B}$ при условии, что сигнал на входе η принял значение y). Это условное распределение характеризует канал и описывает вероятности искажений (шума), возникающих в канале. Весьма часто сигнал на выходе $\tilde{\eta}$ представляется, в определенном смысле, как сумма сигнала η и шума ζ

$$\tilde{\eta} = \eta + \zeta\,. \tag{2a}$$

Переход $\tilde{\eta} \to \tilde{\xi}$ есть декодирование (преобразование) выходного сигнала $\tilde{\eta}$ в воспроизводимое сообщение $\tilde{\xi}$, принимающего значения в пространстве $\tilde{X}$. Кодирование $\xi \to \eta$ и декодирование $\tilde{\eta} \to \tilde{\xi}$ определяются условным распределениями вероятностей

$$P_{\eta|\xi}(B \mid x) = P\{\eta \in B \mid \xi = x\}, \quad B \subset Y, \tag{3}$$

$$P_{\tilde{\xi}|\tilde{\eta}}(A \mid \tilde{y}) = P\{\tilde{\xi} \in \tilde{A} \mid \tilde{\eta} = \tilde{y}\}, \quad \tilde{A} \subset \tilde{X}\,. \tag{4}$$

В большинстве работ рассматриваются лишь так называемое нерандомизированное кодирование и декодирование, когда условные распределения (3) и (4) сосредоточены в одной точке. В этом случае кодирование и декодирование задаются однозначными функциями

$$\eta = f(\xi), \quad \tilde{\xi} = \varphi(\tilde{\eta})\,. \tag{5}$$

В дальнейшем задание правил кодирования и декодирования сообщений и сигналов мы будем отождествлять с заданием условных распределений (3) и (4) или функций (5). Пространства X и $\tilde{X}$, Y и $\tilde{Y}$ могут не совпадать. Воспроизводимое сообщение должно удовлетворять условиям верности, задаваемым получателем. Условие верности обычно сводится к ограничениям на совместные распределения вероятностей $P_{\xi\tilde{\xi}}$ сообщения на выходе источника информа-

нии ξ и воспроизводимого сообщения $\tilde{\xi}$

$$P_{\xi\tilde{\xi}} \in W, \tag{6}$$

где W – некоторое множество совместных распределений пар $(\xi, \tilde{\xi})$. Так как распределение P_ξ – фиксировано, то, в случае существования условных распределений $P_{\tilde{\xi}|\xi}$, задание множества W равносильно заданию некоторого множества условных распределений $P_{\tilde{\xi}|\xi}$.

Хорошо известными примерами условий верности воспроизведения являются

$$\text{а) } P(\xi \neq \tilde{\xi}) = 0 \quad \text{(условие полного воспроизведения)};$$

$$\text{б) } P(\xi \neq \tilde{\xi}) \leqq \varepsilon, \quad \varepsilon > 0; \tag{7}$$

$$\text{в) } E(\xi(t) - \tilde{\xi}(t))^2 \leqq \varepsilon^2, \tag{8}$$

где $\xi(t)$ и $\tilde{\xi}(t)$ – одномерные случайные величины и E – знак математического ожидания.

Могут быть также наложены ограничения на распределения сигналов на входе канала η

$$P_\eta \in V, \tag{9}$$

где V – некоторое множество распределений. Например, в случае когда $\eta = \{\eta(t); 0 < t \leqq T\}$, где $\eta(t)$ – одномерные случайные величины, это ограничение может иметь вид:

$$E\eta^2(t) \leqq P_c \tag{10}$$

– ограничение на среднюю мощность сигнала.

Основная проблема передачи информации заключается в следующем. Пусть заданы P_ξ, $P_{\tilde{\eta}|\eta}$, W, V. Когда возможно задать правила кодирования и декодирования так, чтобы в предположении, что цепь случайных величин $\xi, \eta, \tilde{\eta}, \tilde{\xi}$ марковская и пара случайных величин $(\eta, \tilde{\eta})$ подчинена условиям (2) и (9), мы имеем $P_{\xi\tilde{\xi}} \in W$. Короче эта проблема будет формулироваться так: когда возможно построить передачу так, чтобы воспроизводимое сообщение удовлетворяло заданным критериям верности.

Решение этой проблемы формулируется как теоремы Шеннона. Некоторые ее аспекты будут рассмотрены в § 2. Последующие параграфы будут посвящены построению методов кодирования и декодирования в случае, когда передача возможна.

2. ТЕОРЕМЫ ШЕННОНА. СКОРОСТЬ СОЗДАНИЯ СООБЩЕНИЙ, ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА. ВЕРОЯТНОСТИ ОШИБОК

Основными величинами с помощью которых формулируются и доказываются теоремы Шеннона являются информации и информационная плотность.

Эти величины определяются следующим образом. Пусть ξ и η случайные величины, принимающие значения в пространствах X и Y, P_η, P_ξ и $P_{\xi\eta}$ – распределения вероятностей случайных величин η, ξ и пары случайных величин (ξ, η). Логарифм плотности вероятности меры $P_{\xi\eta}$ по произведению вероятностных мер $P_\xi \,.\, P_\eta$ (если плотность существует) называется *информационной плотностью* $i_{\xi\eta}$

$$i_{\xi\eta}(x, y) = \log \frac{P_{\xi\eta}(\mathrm{d}x\, \mathrm{d}y)}{P_\xi(\mathrm{d}x)\, P_\eta(\mathrm{d}y)} .$$

Математическое ожидание случайной величины $i_{\xi\eta}(\xi, \eta)$ называется *информацией*

$$I(\xi, \eta) = E i_{\xi\eta}(\xi, \eta) = \int_{X \times Y} i_{\xi\eta}(x, y)\, P_{\xi\eta}(\mathrm{d}x\, \mathrm{d}y) . \tag{11}$$

Если информационная плотность не существует, то полагаем

$$I(\xi, \eta) = \infty . \tag{12}$$

Формулировки теорем Шеннона непосредственно выражаются через экстремальные значения информации: пропускной способности канала и скорости создания сообщений. Наибольшее значение (верхняя грань) информации $I(\eta, \tilde{\eta})$ пары случайных величин $(\eta, \tilde{\eta})$, подчиненных условному распределению вероятностей (2) и ограничению (9) называется *пропускной способностью канала*

$$C = \sup_{P_\eta \in V} I(\eta, \tilde{\eta}) . \tag{13}$$

Наименьшее значение (нижняя грань) информации пары случайных величин ξ и $\tilde{\xi}$, подчиненных условию $P_{\xi\tilde{\xi}} \in W$, называется *энтропией* $H_W(\xi)$ сообщения при верности воспроизведения W:

$$H_W(\xi) = \inf_{P_{\xi\tilde{\xi}} \in W} I(\xi, \tilde{\xi}) . \tag{14}$$

(При условии полного воспроизведения $H_W(\xi) = I(\xi, \tilde{\xi}) = H(\xi)$ – энтропия случайной величины ξ.) Часто пространства X и $\tilde{X}$ совпадают и условия верности воспроизведения выражаются через параметр ε, ограничивающий, в определенной метрике, расстояние воспроизводимых значений сообщений от сообщений, генерируемых источником. Примерами таких условий являются условия задаваемые соотношениями (7) и (8). В этих случаях вводятся обозначения $H_\varepsilon(\xi) = H_W(\xi)$ и $H_W(\xi)$ называется *ε-энтропией* сообщения ξ.

Так как при построении передачи сообщений по каналу последовательность $\xi, \eta, \tilde{\eta}, \tilde{\xi}$ образует цепь Маркова, то

$$I(\xi, \tilde{\xi}) \leqq I(\eta, \tilde{\eta}) .$$

Сравнивая это соотношение с (13) и (14) видим, что *для возможности передачи необходимо, чтобы*

$$H_W(\xi) \leqq C\,. \tag{15}$$

Это утверждение носит название обратной теоремы Шеннона.

Утверждение прямой теоремы Шеннона о достаточных условиях для возможности передачи носит асимптотический характер. Обычно величины $H_W(\xi) = H_{W^T}(\xi^T)$ и $C = C^T$ зависят от параметра T, являющегося временем генерирования сообщения источником и длительностью сигнала на входе и выходе канала. В этих случаях рассматриваются величины

$$\overline{C} = \lim_{T\to\infty} \frac{1}{T} C^T \tag{16}$$

и

$$\overline{H} = \overline{\lim_{T\to\infty}}\ \frac{1}{T} H_{W^T}(\xi^T)\,. \tag{17}$$

$\overline{C}$ – также называется *пропускной способностью канала*, $\overline{H}$ – *скоростью создания сообщений.*

Прямая теорема Шеннона утверждает, что в весьма общих ситуациях при выполнении соотношения

$$\overline{\lim_{T\to\infty}}\ \frac{H_{W^T}(\xi^T)}{C^T} < 1 \tag{18}$$

возможно построение передачи.

Для справедливости этого утверждения по существу достаточно выполнения следующих условий: при любых $\delta_1, \delta_2 > 0$ можно построить пары случайных величин $(\xi^T, \tilde{\xi}^T)$ и $(\eta^T, \tilde{\eta}^T)$, удовлетворяющих соотношениям (2), (6) и (9), такие, что

$$\overline{\lim_{T\to\infty}}\ \frac{I(\xi^T, \tilde{\xi}^T)}{H_{W^T}(\xi^T)} \leqq 1 + \delta_1\,, \quad \lim_{T\to\infty} \frac{I(\eta^T, \tilde{\eta}^T)}{C^T} \geqq 1 - \delta_1\,, \tag{19}$$

$$\lim_{T\to\infty} P\left\{\left|\frac{i_{\xi^T\tilde{\xi}^T}(\xi^T, \tilde{\xi}^T)}{I(\xi^T, \tilde{\xi}^T)} - 1\right| < \delta_2\right\} = 0\,, \tag{20}$$

$$\lim_{T\to\infty} P\left\{\left|\frac{i_{\eta^T\tilde{\eta}^T}(\eta^T, \tilde{\eta}^T)}{I(\eta^T, \tilde{\eta}^T)} - 1\right| < \delta_2\right\} = 0\,, \tag{21}$$

Совокупности случайных величин $(\xi^T, \tilde{\xi}^T)$ и $(\eta^T, \tilde{\eta}^T)$, для которых выполнены соотношения (20) и (21), соответственно, называются *информационно устойчивыми.*

Как известно [3, 4] выполнение соотношений (19) и (20) приводит к тому, что при любом $\delta > 0$ для достаточно больших T существует $2^{H_{W^T}(\xi^T)(1+\delta)}$ значений воспроизводимых сообщений, такие, что по ним с верностью W^T мы можем воспроизвести генерируемое источником сообщение, а выполнение условий (19) и (21) привожит к тому, что при достаточно большом T существует $2^{C^T(1-\delta)}$ значений сигнала на входе канала, которые мы можем восстановить с как угодно малой вероятностью ошибки по значениям сигналов на выходе канала.

Более того, Шеннон [5] показал, что если M возможных значений входных сигналов образуются путем независимой выборки с распределением P_{η^T}, то при восстановлении входных сигналов по значениям выходных сигналов вероятность ошибки, усредненная по ансамблю таких выборок, удовлетворяет неравенству

$$P_{\text{ош}} \leqq P\{i_{\eta^T\tilde{\eta}^T}(\eta^T, \tilde{\eta}^T) \leqq \log M + \Theta\} + \mathrm{e}^{-\Theta}, \tag{22}$$

где $\Theta > 0$ — произвольное число.

Отсюда, в частности, видно, что если число сообщений меньше $2^{C^T(1-\delta)}$, $\delta > 0$, $T \to \infty$ и выполнено (19) и (21), то мы можем передать сообщения при сколь угодно малой вероятности ошибки при воспроизведении.

Важно отметить, что в свою очередь, соотношения (19) – (21) имеют место при весьма слабых ограничениях на каналы и источники; этими ограничениями могут, например, являться эргодичность в определенном смысле источника и канала и конечность пределов $\overline{C}$ и $\overline{H}$. Разнообразные предположения, при выполнении которых доказывается теорема Шеннона, обычно оказываются частными случаями такой эргодичности. Идея доказательства справедливости соотношений (19) и (21) хорошо иллюстрируется случаем, когда сигнал на выходе $\tilde{\eta}(t)$ представляется в виде сигнала на входе и независимого его стационарного шума $\zeta(t)$, т.е.

$$\tilde{\eta}(t) = \eta(t) + \zeta(t), \tag{23}$$

а ограничения на входной сигнал сводятся к некоторым ограничениям на распределения случайных величин $\eta(t)$ (например, среднее значение квадрата $\eta(t)$ не больше P_c, $E\,\eta^2(t) \leqq P_c$). В этом случае для справедливости (19) и (21), а тем самым и теоремы Шеннона достаточно, чтобы шум $\zeta(t)$ был вполне эргодическим и

$$\overline{C} = \lim_{T \to \infty} \frac{1}{T} C^T < \infty . \tag{24}$$

Процесс $\zeta(t)$ называется *вполне эргодическим,* если при любом $T > 0$ эргодична двусторонняя последовательность случайных величин

$$\ldots, \beta_{-1}, \beta_0, \beta_1, \beta_2, \ldots, \beta_k, \ldots ,$$

где

$$\beta_k = \zeta_{kT}^{(k+1)T} = \{\zeta(t);\ kT < t \leqq (k+1)\,T\}$$

(вполне эргодическим будет, например, процесс с перемешиванием). Для получения соотношений (19) и (21) надо рассмотреть случайный процесс $\eta(t)$ такой, что случайные величины

$$v_k = \eta_{kT}^{(k+1)T} = \{\eta(t);\ kT < t \leqq (k+1)\,T\}\,,\quad k = 0, \pm 1, \pm 2, \ldots$$

независимы, имеют одно и тоже распределение вероятностей и

$$I(v_k, \tilde{v}_k) = I\left(\eta_{kT}^{(k+1)T}, \tilde{\eta}_{kT}^{(k+1)T}\right) > (\bar{C} - \delta)\,T, \tag{25}$$

где

$$\tilde{v}_k = \tilde{\eta}_{kT}^{(k+1)T} = \{\tilde{\eta}(t);\ kT < t \leqq (k+1)\,T\}\,,\quad \tilde{\eta}(t) = \eta(t) + \zeta(t)\,.$$

Как известно (см. [6]), совокупность пар $(v_k, \tilde{v}_k)$ образует эргодический процесс и

$$\lim_{k\to\infty} P\left\{\left|\frac{i_{\eta_0^{kT}\tilde{\eta}_0^{kT}}(\eta_0^{kT}, \tilde{\eta}_0^{kT})}{I(\eta_0^{kT}, \tilde{\eta}_0^{kT})} - 1\right| < \delta_2\right\} = 0\,,\quad \delta_2 > 0\,. \tag{26}$$

Кроме того, используя независимость случайных величин v_k и проведя элементарные преобразования информации, имеем

$$I(\eta_0^{kT}, \tilde{\eta}_0^{kT}) \geqq kI(\eta_0^{T}, \tilde{\eta}_0^{T}) > kT(\bar{C} - \delta)\,. \tag{27}$$

Из (24), (26) и (27) очевидно вытекает (19) и (21).

Аналогично доказывается, что соотношения (19) и (20) справедливы для источников, генерирующих стационарный вполне эргодический процесс $\xi(t)$, для которого условия верности сводятся к ограничениям на распределения вероятностей пары случайных величин $(\xi(t), \tilde{\xi}(t))$

$$P_{\xi(t)\tilde{\xi}(t)} \in W_t = W, \tag{28}$$

(W_t не зависит от t), и при некотором $\tau > 0$

$$I(\xi_0^{\tau}, \tilde{\xi}_0^{\tau}) < \infty\,. \tag{29}$$

В этом случае строится процесс $\tilde{\xi}(t)$ такой, что при любом целом k случайные величины

$$\mu_k = \xi_{kT}^{(k+1)T}\,,\qquad \tilde{\mu}_k = \tilde{\xi}_{kT}^{(k+1)T}\,,\qquad \{\xi(t), \tilde{\xi}(t);\ -\infty < t \leqq kT, (k+1)\,T < t < \infty\}$$

образуют цепь Маркова (т.е. при фиксированной реализации процесса $\xi(t)$ случайные величины $\ldots, \tilde{\xi}_{-T}^{0}, \tilde{\xi}_{0}^{T}, \tilde{\xi}_{T}^{2T}, \ldots$ становятся независимыми) и пары случайных величин $(\mu_k, \tilde{\mu}_k)$ $k = 0, \pm 1, \ldots$ имеют одно и то же распределение

вероятностей. Отличие от канала здесь будет состоять лишь в том, что соотношение (27) заменяется на

$$I(\xi_0^{kT}, \tilde{\xi}_T^{kT}) \leqq kT(\xi_0^T, \tilde{\xi}_0^T). \tag{30}$$

Из (29) и (30) вытекает, что существует предел

$$\bar{H} = \lim_{T\to\infty} \frac{1}{T} H_{W^T}(\xi^T) \leqq \frac{1}{\tau} I(\xi_0^\tau, \tilde{\xi}_0^\tau) < \infty. \tag{31}$$

Сочетание соотношений (30) и (31) приводит к выражению аналогичному (27)

$$I(\xi_0^{kT}, \tilde{\xi}_0^{kT}) \leqq kI(\xi_0^T, \tilde{\xi}_0^T) < kT(\bar{H} + \delta). \tag{32}$$

Следует отметить, что само требование эргодичности или хотя бы стационарности для выполнения (20) и (21) не является необходимым. Так, например, для гауссовских источников и каналов для выполнения (20) и (21) необходимо и достаточно, чтобы

$$\varliminf_{T\to\infty} H_{W^T}(\xi_0^T) = \infty, \quad \varliminf_{T\to\infty} C^T = \infty. \tag{33}$$

Как известно [1–3], канал называется гауссовским, если условное распределение $P_{\tilde{y}|y} = P_{\tilde{\eta}|\eta}$ гауссовское, а ограничение на распределении сигналов η сводится к ограничениям на первые и вторые моменты. Например, для каналов вида (23), будут гауссовскими, если шум $\zeta(t)$ гауссовский процесс и ограничение на входной сигнал сводится к ограничению на среднюю мощность (10). Сообщение называется гауссовским, если случайная величина ξ – гауссовская, а ограничение на совместное распределение $P_{\xi\tilde{\xi}} \in W$ сводится к ограничениям на первые и вторые моменты пар случайных величин $(\xi, \tilde{\xi})$. Например, сообщение $\xi = \xi_0^T = \{\xi(t);\ 0 \leqq t \leqq T\}$ будет гауссовским, если порождающий его случайный процесс $\xi(t)$ гауссовский, а условия верности воспроизведения сводятся к (8) или к

$$E\int_\tau^{T+\tau} a(t)\,(\xi(t-\tau) - \tilde{\xi}(t-\tau))^2\,\mathrm{d}t \leqq \varepsilon^2,\ -\infty < \tau < \infty, \tag{34}$$

где $a(t)$ – вещественная функция.

Указанные результаты для гауссовских источников и каналов вытекают из того, что для информационной устойчивости гауссовских пар случайных величин $(\alpha^T, \tilde{\alpha}^T)$ необходимо и достаточно выполнения условия

$$\varliminf_{T\to\infty} I(\alpha^T, \tilde{\alpha}^T) = \infty. \tag{35}$$

Неравенство (22) полезно не только для обоснования теоремы Шеннона, но и для получения весьма общих оценок вероятности ошибок. Проиллюстрируем

это на примере гауссовских каналов. Информационная плотность пары $(m+n)$-мерных гауссовских случайных величин $(\alpha, \tilde{\alpha}) = (\alpha_1, \ldots, \alpha_m, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_n)$ допускает представление

$$i_{\alpha\tilde{\alpha}}(\alpha, \tilde{\alpha}) = \sum_{j=1}^{k} (\beta_j^2 - \tilde{\beta}_j^2) \sqrt{\lambda_j} - \tfrac{1}{2} \sum_{j=1}^{k} \log (1 - \lambda_j), \tag{36}$$

где $0 \leqq \lambda_j < 1$; $k \leqq \min(m, n)$; $\beta_j, \tilde{\beta}_j$, $j = 1, \ldots, k$ – независимые гауссовские случайные величины с нулевым средним и единичной дисперсией и

$$-\tfrac{1}{2} \sum_{i=1}^{k} \log (1 - \lambda_i) = I(\alpha, \tilde{\alpha}). \tag{37}$$

Выражение (36) может быть получено следующим образом. Как известно [2], после надлежащих невырожденных линейных преобразований случайных величин $\alpha_1, \ldots, \alpha_m$ и $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n$ полученные случайные величины $\alpha'_1, \ldots, \alpha'_m$ и $\tilde{\alpha}'_1, \ldots, \tilde{\alpha}'_n$ за исключением пар $(\alpha'_j, \tilde{\alpha}'_j)$ независимы, а тогда

$$i_{\alpha\tilde{\alpha}}(\alpha, \tilde{\alpha}) = i_{\alpha'\tilde{\alpha}'}(\alpha', \tilde{\alpha}') = \sum_{j=1}^{k} i_{\alpha_j'\tilde{\alpha}_j'}(\alpha'_j, \tilde{\alpha}'_j). \tag{38}$$

Но для гауссовской пары $(\alpha'_j, \tilde{\alpha}'_j)$ одномерных случайных величин, выражение для информационной плотности находится непосредственно

$$i_{\alpha_j'\tilde{\alpha}_j'}(\alpha'_j, \tilde{\alpha}'_j) = (\beta_j^2 - \tilde{\beta}_j^2) \sqrt{\lambda_j} - \tfrac{1}{2} \log (1 - \lambda_j) = (\beta_j^2 - \tilde{\beta}_j^2) \sqrt{\lambda_j} + I(\alpha'_j, \tilde{\alpha}'_j), \tag{39}$$

где $0 \leqq \lambda_j < 1$ – коэффициенты корреляции случайных величин α'_j и $\tilde{\alpha}'_j$. Нетрудно показать, что λ_j при $\det N \neq 0$ суть корни уравнений

$$\det (\lambda Q - SN^{-1}S') = 0, \tag{40}$$

где $\begin{pmatrix} Q & S' \\ S & N \end{pmatrix}$, Q, N – матрицы ковариаций случайных величин $(\alpha, \tilde{\alpha})$, α и $\tilde{\alpha}$, соответственно.

Вообще для любой гауссовской пары $(\alpha, \tilde{\alpha})$ случайных величин (не обязательно конечномерных) имеют место две альтернативы:

1. вероятностная мера $P_{\alpha\tilde{\alpha}}$ сингулярна относительно произведения мер $P_\alpha \cdot P_{\tilde{\alpha}}$;

2. вероятностная мера $P_{\alpha\tilde{\alpha}}$ абсолютно непрерывна относительно меры $P_\alpha \cdot P_{\tilde{\alpha}}$ и информационная плотность выражается в виде (36) (где k может принимать как конечные так и бесконечные значения).

Сочетая (22) и (36) можно легко получить следующую оценку для оптималь-

ной вероятности ошибки для произвольного гауссовского канала

$$P_{\text{ош}} \leqq \gamma e^{-(C-\log M)^2/4C}, \tag{41}$$

где M – число значений сигналов, в которые кодируется передаваемое сообщение, γ – некоторое число, не зависящее от канала и от числа M. В частности, если $C = T\bar{C}$ и $\log M = \bar{R}T$, то формула (41) принимает вид

$$P_{\text{ош}} \leqq \gamma e^{-T(\bar{C}-\bar{R})^2/4\bar{C}}. \tag{42}$$

Знание асимптотики распределения чисел λ_j позволяет, естественно, получить более сильные оценки для $P_{\text{ош}}$, чем (41).

В ряде случаев такую асимптотику нетрудно получить; например, пусть $(\alpha(t), \tilde{\alpha}(t))$ – двумерный гауссовский стационарный процесс дискретного аргумента;

$$A(\omega) = \begin{pmatrix} f_{\alpha\alpha}(\omega) & f_{\alpha\tilde{\alpha}}(\omega) \\ f_{\tilde{\alpha}\alpha}(\omega) & f_{\tilde{\alpha}\tilde{\alpha}}(\omega) \end{pmatrix}$$

матрица производных спектральных функций; тогда если

$$\int |\log f_{\alpha\alpha}(\omega)| \, d\omega < \infty$$

(т.е. процесс $\alpha(t)$ несингулярен), то в точках непрерывности функции

$$r^2(\omega) = \frac{|f_{\alpha\tilde{\alpha}}(\omega)|^2}{f_{\alpha\alpha}(\omega) f_{\tilde{\alpha}\tilde{\alpha}}(\omega)}$$

для пары случайных величин $(\alpha_0^T, \tilde{\alpha}_0^T)$

$$\varkappa(\lambda) = \frac{T}{2\pi} \mu(\omega : r^2(\omega) > \lambda) + O(T), \tag{43}$$

где $\varkappa(\lambda)$ – число λ_j больших λ, а $\mu(\omega : .)$ – мера Лебега множества точек ω.

Этот результат может быть использован для оценки асимптотики вероятности ошибки для гауссовского канала со стационарным шумом вида (23), когда $\zeta(t)$ – стационарный гауссовский процесс.

Доказательство (43) проводится аналогично теоремам Сеге [21] об асимптотике собственных значений матриц Теплица. На этом мы здесь останавливаться не будем.

3. КОДИРОВАНИЕ И ДЕКОДИРОВАНИЕ ДЛЯ ДИСКРЕТНЫХ СООБЩЕНИЙ КАНАЛОВ

Обсуждение проблемы кодирования и декодирования для дискретных систем связи мы проведем на примере двоичного симметричного канала без памяти

и сообщения образованного последовательностями двоичных символов 0,1, появляющихся независимо и с вероятностями $p(0) = p(1) = \frac{1}{2}$; критерием верности является вероятность неправильного воспроизведения информационных символов.

Ограничение рассмотрений таким примером вызвано только удобством и простотой изложения. Приведенные ниже факты легко переносятся на общий случай передачи информации.

Сигналы на входе и выходе двоичного симметричного канала без памяти образуются последовательностями двоичных символов 0, 1

$$\text{вход: } 1010\ldots = y_1 y_2 y_3 y_4 \ldots$$

$$\text{выход: } 0010\ldots = \tilde{y}_1 \tilde{y}_2 \tilde{y}_3 \tilde{y}_4 \ldots,$$

а условные вероятности для такого канала представляются в виде

$$p(\tilde{y}_1, \ldots, \tilde{y}_n \mid y_1, \ldots, y_n) = p(\tilde{y}_1 \mid y_1) \ldots p(\tilde{y}_n \mid y_n), \tag{44}$$

$$p(1 \mid 1) = p(0 \mid 0) = q\,; \quad p(1 \mid 0) = p(0 \mid 1) = p = 1 - q\,,$$

p – вероятность ошибки.

Таким образом ошибки в канале возникают независимо и с вероятностью p. Сигнал на выходе $\tilde{y} = (\tilde{y}_1 \ldots \tilde{y}_n)$ такого канала можно представить как покомпонентную сумму по mod 2 значения сигнала на входе $y = (y_1 \ldots y_n)$ и независимого от него шума $z = (z_1 \ldots z_n)$

$$\tilde{y}_i = y_i + z_i\,,$$

где символы $z_1, \ldots, z_n$ появляются независимо с вероятностями

$$p(z_i) = \begin{cases} q & \text{при } z_i = 0\,, \\ p & \text{при } z_i = 1\,. \end{cases}$$

Для этого канала

$$\bar{C} = 1 + p \log_2 p + q \log_2 q\,. \tag{45}$$

Задачу кодирования мы будем формулировать следующим образом: пусть в единицу времени по каналу передается один символ, а символы генерируемые источником поступают через равные промежутки времени $1/\bar{R} \geqq 1$, так что в момент времени k будет известно $[k\bar{R}] + 1$ информационных символов. Кодирование задает закон, по которому по символам, генерированным источником, к k-му моменту времени строится k-ый символ y_k сигнала на входе канала. Математически это сводится к заданию функции

$$y_k = f_k(x_1, \ldots, x_{[k\bar{R}]+1}), \quad k = 1, 2, \ldots \tag{46}$$

от $[k\bar{R}] + 1$ аргументов. После передачи по каналу последовательность $y_1y_2 \ldots$ переходит в последовательность $\tilde{y}_1\tilde{y}_2 \ldots$ Декодирование задает закон, по которому по выходной последовательности $\tilde{y}_1\tilde{y}_2 \ldots$ строится выходное сообщение. Математически это сводится к заданию функции

$$\tilde{x}_l = \varphi_l(\tilde{y}_1, \tilde{y}_2, \ldots). \tag{47}$$

При этом говорят, что декодирование производится с равномерно ограниченной задержкой s, если время между поступлением l-го символа сообщения $l = [k\bar{R}] + 1$ и его декодированием, не более чем s, т.е.

$$\tilde{x}_l = \varphi_l(\tilde{y}_1, \ldots, \tilde{y}_k, \ldots, \tilde{y}_{k+s}). \tag{48}$$

Теорема Шеннона утверждает, что при $\bar{R} < \bar{C}$, существует кодирование и декодирование, такие, что с вероятностью $P = 1 - \bar{P}_{\text{ош}}$, сколь угодно близкой к единице, символ $\tilde{x}_l$ совпадает с x_l.

Проблема заключается в построении простых функций f_k и φ_l, при этом желательно, чтобы с увеличением задержки s вероятность ошибки $\bar{P}_{\text{ош}}$ стремилась к нулю оптимальным или близким к оптимальному образом.

Открытие „групповых кодов" (см. [11]) позволило весьма эффективно решать проблему кодирования. Поясним это на примере блокового кодирования. При блоковом кодировании подлежащая передаче информационная последовательность двоичных символов $x_1x_2 \ldots$ разбивается на блоки длины $m = n\bar{R}$, а последовательность символов на входе канала на блоки длины n. Среди 2^n блоков-возможных значений сигналов длины n, выбираем 2^m различных блоков, в которые кодируем (сопоставляем) блоки информационной последовательности. Кодирование является групповым линейным, если информационная последовательность $x = (x_1 \ldots x_m)$ образованная покомпонентным сложением по mod 2 двух других последовательностей

$$x^1 = (x_1^1 \ldots x_m^1) \quad \text{и} \quad x^2 = (x_1^2 \ldots x_m^2),$$

кодируется в блок на входе канала $y = (y_1 \ldots y_n)$, являющимся покомпонентной суммой по mod 2 блоков на входе канала, соответствующих x^1 и x^2. Известно, что в этом случае кодирование можно представить в виде произведения матрицы A на вектор – столбец x'

$$y' = Ax' = \begin{pmatrix} a_{11} \ldots a_{1m} \\ \ldots\ldots\ldots \\ a_{n1} \ldots a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \tag{49}$$

$$x' = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \quad y' = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

где элементы a_{ij} матрицы A принимают значения 0 и 1: суммирование ведется по mod 2. Задержка при декодировании ограничена числом n. При этом если матрица A невырождена, то не меняя вероятности ошибки декодирования блока, мы можем матрицу A преобразовать к виду

$$a_{ij} = \begin{cases} 1 \text{ при } i = j; \\ 0 \text{ при } i \neq j;\ i, j \leqq m. \end{cases}$$

В этом случае результат умножения Ax' представится в виде

$$y' = Ax' = \begin{pmatrix} x_1 \\ \vdots \\ x_m \\ \sum_{i=1}^{m} a_{(m+1)i} x_i \\ \vdots \end{pmatrix}, \tag{50}$$

т.е. первые m символов кодового блока $y = (y_1 \ldots y_n)$ совпадают с символами сообщения, а остальные являются их линейными комбинациями. Эти первые m символов называются информационными, а остальные проверочными. Кодирование, допускающее разделение кодовых символов на информационные и проверочные называются систематическим.

Заметим, что (при фиксированном $m/n = \bar{R}$) при случайной и независимой выборке элементов a_{ij}, $i > m$, с вероятностями $p(0) = p(1) = \frac{1}{2}$ вероятность неправильного декодирования кодовых блоков, осредненная по всему ансамблю матриц A, убывает по экспоненте с ростом n. Эта экспонента совпадает с экспонентой вероятности ошибки при случайном независимом выборе (с вероятностями $p(y_1, \ldots, y_n) = 2^{-n}$) 2^m кодовых блоков из n символов.

Декодирование проводится по критерию максимума апостериорной вероятности.

Среди групповых кодов выделяют подклассы кодов столь же оптимальных с точки зрения вероятности ошибки, но обладающих еще большой простотой кодирования.

Нетрудно убедиться, что умножение Ax' требует числа арифметических операций, растущего степенным образом с ростом n. Однако, оптимальное декодирование (по критерию минимума апостериорной вероятности) требует, вообще говоря, сравнения последовательности на выходе канала $\tilde{y}_1 \ldots \tilde{y}_n$ со всеми возможными входными последовательностями. За переданную принимается та, которая отличается от принятой в наименьшем числе символов. Число всех возможных кодовых последовательностей $y_1 \ldots y_n$ равно $2^m = 2^{Rn}$ и число сравнений экспоненциально растет с ростом n. При больших n такой перебор при декодировании становится практически невозможным.

Одной из наиболее интересных и перспективных процедур декодирования является метод последовательного декодирования, развитый Возенкрафтом и Райфеном [10]. Изложенная ими процедура позволила разработать такие методы группового кодирования для которых число операций при декодировании одного информационного символа растет не быстрее, чем степенным образом с ростом задержки s, а вероятность неправильного декодирования $P_{\text{ош}}$ экспоненциально убывает с ростом s. Под числом операций при декодировании мы будем в дальнейшем понимать среднее число сравнений последовательностей на выходе канала с кодовыми комбинациями на входе канала. Определенное таким образом число операций может быть выражено через среднее число арифметических операций, выполняемых при декодировании. Алгоритм декодирования, подробно разобранный в книге Возенкрафта и Райфена [10], требует все же экспоненциального роста числа операций с s, однако небольшое видоизменение этого алгоритма приводит к желаемому степенному росту. Результаты моделирования алгоритма, указанного в [10], оказались весьма благоприятными.

Следует заметить, что в публикациях, посвященных последовательному декодированию [10, 12–14] (а также другим методам декодирования [15]), такой рост числа операций с ростом s был получен для скоростей передачи $\bar{R}$ меньших некоторого $R_{\text{выч}} < \bar{C}$. Для дискретного канала без памяти значения этого порога $R_{\text{выч}}$, рассматриваемые в указанных работах, оцениваются сверху выражением

$$R_{\text{выч}} \leqq \bar{R}_{\text{выч}} = \max_{p_i} \{ -\log \sum_j [\sum_i p_i \sqrt{p_{j|i}}]^2 \}, \tag{51}$$

где p_i – вероятности символов на входе канала, $p_{j|i}$ – вероятности перехода. Для двоичного симметричного канала формула (51) принимает вид

$$\bar{R}_{\text{выч}} = \bar{R}_{\text{выч}}(p) = -\log \tfrac{1}{2}(1 + 2\sqrt{[p(1-p)]}). \tag{52}$$

Легко убедиться, что при $p \to \frac{1}{2}$ значения $\bar{R}_{\text{выч}}$ сходятся к $\frac{1}{2}\bar{C}$.

Мы покажем ниже, что *для любых скоростей передачи $\bar{R}$, меньших пропускной способности $\bar{C}$, можно построить методы кодирования такие, что число операций при декодировании одного информационного символа будет меньше некоторой константы, не зависящей от вероятности ошибки $\varepsilon > 0$ (однако зависящей от скорости передачи $\bar{R}$).*

Искомое кодирование будет образовано итерацией группового блокового кодирования и последовательного (сверточного) кодирования, а декодирование итерацией блокового и последовательного декодирования [10, 11, 17].

Для описания этой конструкции нам понадобятся некоторые сведения по последовательному декодированию Возенкрафта и итеративных кодах Элайеса.

Последовательное декодирование предполагает наличие некоторого специального вида кодирования — последовательного кодирования, при котором кодовые последовательности образуют так называемый древовидный код. Поясним построение этого кода, для случая, когда скорость передачи имеет вид $\bar{R} = 1/r$, где r — целое число. Каждой последовательности информационных символов $x_1 x_2 \dots x_m$ ставим в соответствие последовательность кодовых символов $y_1 y_2 \dots y_{mr}$ следующим образом. Разбиваем последовательность $y_1 y_2 \dots y_{mr}$ на „отрезки" из r символов $y_1 \dots y_r$; $y_{r+1} \dots y_{2r}$; $y_{(m-1)r+1} \dots$ $\dots y_{mr}$. Значения кодовых символов из m-го отрезка $y_{(m-1)r+1} \dots y_{mr}$ определяются значениями m-информационных символов $x_1 \dots x_m$

$$y_{(m-1)r+i} = f_{mi}(x_1, \dots, x_m), \quad i = 1, \dots, r, \tag{53}$$

т.е. из каждого информационного символа x_m „исходит" r кодовых символов $y_{(m-1)r+1} \dots y_{mr}$.

Построение такого кода для случая $\bar{R} = \frac{1}{3}$, $m = 4$; $r = 3$; $n = rm = 12$ хорошо иллюстрируется на рис. 1. Здесь каждому информационному символу соответствует три кодовых. Кодовые последовательности служат как бы ветвями дерева; информационные символы в узлах указывают путь следования; ребро этого дерева соответствует трем кодовым символам, написанным сверху этого ребра. Например, информационной последовательности: 1011 соответствует кодовая последовательность: 111 011 101 101. Для такого кода характерно, что две информационные последовательности, имеющие одни и те же k начальных информационных символов имеют в соответствующих кодовых последовательностях kr общих начальных символов, а в кодовом дереве k общих начальных ребер.

Оптимальное декодирование (по критерию максимума апостериорной вероятности) с задержкой s производится аналогично блоковому декодированию. Для декодирования первого информационного символа x_1 полученная на выходе канала последовательность $\tilde{y}_1 \dots \tilde{y}_s$ сравнивается со всеми кодовыми последовательностями (ветвями кодового дерева) длины s. Если ближайшая ветвь (т.е. ветвь, отличающаяся от принятой в минимальном числе символов) соответствует информационной последовательности с первым символом $\tilde{x}_1$, то мы принимаем, что $\tilde{x}_1$ — первый информационный символ; для декодирования второго информационного символа x_2 рассматриваем последовательность $\tilde{y}_{r+1} \dots \tilde{y}_{r+s}$ и сравниваем ее со всеми кодовыми последовательностями $\tilde{y}_{r+1} \dots \tilde{y}_{r+s}$ исходящими из узла соответствующему информационному символу $\tilde{x}_1$, и т.д.

Имеет место следующая теорема:

Если значения функций $y_{(m-1)r+i} = f_{mi}(x_1, \dots, x_m)$ *получаются при независимой случайной выборке двух элементов* 0 *и* 1 *с вероятностями* $p(0) = p(1) = \frac{1}{2}$, *то*

вероятность неправильного декодирования символа (усредненная по ансамблю кодов) экспоненциально убывает с задержкой s и эта экспонента не больше экспоненты при случайном блоковом кодировании.

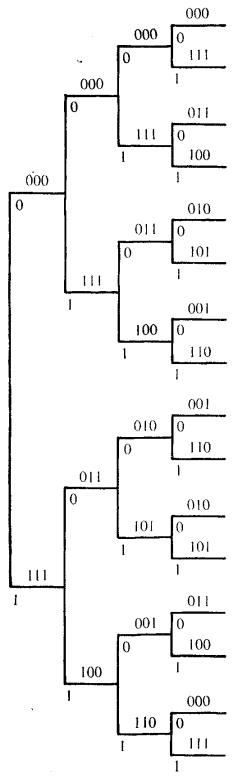

Рис. 1.

Экспонента не изменится, если функция f_{mi} будет зависеть лишь от $s/r + 1$ последних информационных символов $x_{m-s/r} \dots x_m$. Более того (с сохранением экспоненты вероятности ошибки не большей, чем при блоковом декодировании), кодовое дерево можно построить следующим образом (см. рис. 1). Рассматривается последовательность из s символов $u = u_1 u_2 \dots u_s$, называе-

мая генерирующей, такая, что $u_1 = 1$; $u_{rj+1} = 0$; $j = 1, \ldots, s/r - 1$, остальные символы получаются путем независимого и случайного выбора 0 и 1 с вероятностями $p(0) = p(1) = \frac{1}{2}$. Далее полагаем

$$\mathfrak{D}^0 u = u;\ \mathfrak{D}u = \underbrace{0 \ldots 0}_{r \text{ нулей}}\ u_1 \ldots u_s;\ \ldots;\ \mathfrak{D}^m u = \underbrace{0 \ldots 0}_{mr \text{ нулей}}\ u_1 \ldots u_s,$$

т.е. знак $\mathfrak{D}$ означает сдвиг вправо последовательности u на r символов. Символ $y_{(m-1)r+i}$, $i = 1, 2, \ldots, r$ определяется как $(m-1)r + i$-ый символ в сумме

$$\sum_{j=1}^{m} x_j \mathfrak{D}^{j-1} u \pmod 2. \tag{54}$$

Здесь

$$x_j \mathfrak{D}^j u = \begin{cases} \mathfrak{D}^j u & \text{при } x_j = 1, \\ 0 & \text{при } x_j = 0. \end{cases}$$

Очевидно, что не более чем в s/r последних членах суммы значения $(m-1)r + i$-го символа отлично от нуля. Построенный таким образом древовидный код будет систематическим и $y_{mr+1} = x_{m+1}$. Код, изображенный на рис. 1, получен именно таким образом с генерирующей последовательностью $u = 111\,011\,010\,001$ и $r = 3$.

Древовидная структура кода позволяет построить пороговую процедуру декодирования, резко сокращающую число операций (перебор) при сохранении экспоненты вероятности неправильного декодирования. Мы ограничимся набросками одного из всевозможных подходов к таким процедурам (см. [10]).

На некоторой длине $s_1 < s$ принятая последовательность $\tilde{y}_1 \ldots \tilde{y}_{s_1}$ сравнивается с возможными кодовыми последовательностями $y_1 \ldots y_{s_1}$ и отбрасываются те кодовые последовательности, которые находятся от $\tilde{y}_1 \ldots \tilde{y}_{s_1}$ на расстоянии Хемминга,* большем некоторого порога $p_1 s_1$, $0 < p < p_1 < \frac{1}{2}$. Далее оставшиеся кодовые последовательности рассматриваются на длине s_2 и из них выбрасываются те, которые находятся от принятой последовательности на расстоянии $p_2 s_2$, $0 < p < p_2 < \frac{1}{2}$ и т.д. Последнее сравнение производится на всей длине задержки s. Если на длине s останется непустое множество кодовых последовательностей, то за $\tilde{x}_1$ принимается информационный символ одной последовательности из этого множества. Если же такой кодовой последовательности не оказалось, то процедура повторяется с другими, более слабыми, порогами и т.д. Для декодирования второго информационного символа эту процедуру повторяем для выходной последовательности $\tilde{y}_{r+1} \ldots \tilde{y}_{r+s}$, причем сравнения производятся только с теми кодовыми последовательностями, которые начинаются с символов $\tilde{y}_1 \ldots \tilde{y}_r$, соответствующих первому информа-

* Расстояние Хемминга между двумя двоичными последовательностями — есть число символов, в которых эти последовательности отличны друг от друга.

ционному символу $\tilde{x}_1$. Декодирование третьего и т.д. информационного символов производится аналогичным образом.

Как указано в [14] существуют такие модификации последовательного декодирования, что для $\bar{R} < \bar{R}_{выч}(p)$ число операций при декодировании ограничено сверху числом, которое не зависит от задержки s при экспоненциальном убывании с s вероятности неправильного декодирования.

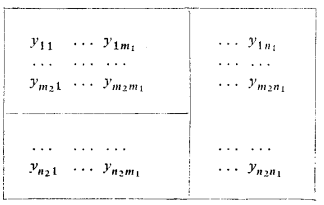

Рис. 2.

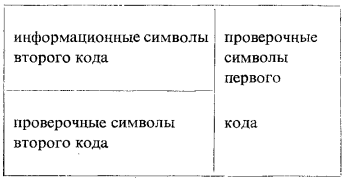

Рис. 3.

Остановимся теперь на кратном описании систематических итеративных кодах Элайеса [11], составленных из двух кодов и являющихся как бы их произведением. Пусть мы имеем два блоковых систематических кода длин n_1 и n_2 с m_1 и m_2 информационными символами. Итеративный код, длины $n = n_1n_2$ с m_1m_2 информационными символами можно представить с помощью прямоугольной матрицы рис. 2 и рис. 3. Строки этой матрицы образуют кодовые последовательности длины n_1 первого кода, а столбцы — кодовые последовательности длины n_2 второго кода; первые m_1 элементов в строке являются информационными символами первого кода, а первые m_2 элементов столбца — информационными элементами второго кода; m_1m_2 символов прямоугольной матрицы в левом верхнем углу являются информационными символами итеративного кода.

Очевидно, что скорость передачи итеративного кода равна

$$R = \frac{m}{n} = \frac{m_1m_2}{n_1n_2} = \frac{m_1}{n_1}\frac{m_2}{n_2} = R_1R_2, \tag{55}$$

т.е. равна произведению скоростей передачи составляющих кодов.

Рис. 2 нам удобно представить схематическим рис. 3. Если второй код — древовидный (сверточный), в котором информационные символы перемежаются с проверочными, то вместо рис. 3 удобно пользоваться рис. 4.

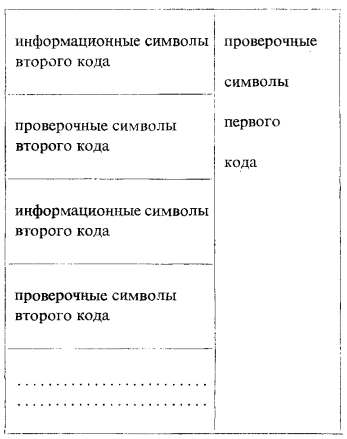

Рис. 4.

Приступим теперь к описанию конструкции кодирования и декодирования обеспечивающей при любых $\bar{R} < \bar{C}$, ограниченность числа операций при декодировании независимо от вероятности неправильного декодирования ε.

Рассмотрим отношение

$$\frac{\bar{R}_{\text{выч}}(p)}{\bar{C}} = \frac{-\log \frac{1}{2}(1 + 2\sqrt{(p(1-p))})}{1 + p\log_2 p + (1-p)\log_2(1-p)}. \tag{56}$$

График зависимости этого отношения от p имеет вид, указанный на рис. 5.

Пусть задана скорость $\bar{R} < \bar{C}$. Выберем $\bar{R}_1 < \bar{C}$ и $\bar{R}_2 < 1$ так, чтобы $\bar{R}_1\bar{R}_2 = \bar{R}$ и целое число n_1 столь большим, чтобы блоковый код длины n_1 с $m_1 = [n_1\bar{R}_1] + 1$ информационными символами обеспечивал при декодировании вероятность ошибки, меньшую некоторого числа p'.

Далее выберем p' столь малым, чтобы для двоичного симметричного канала с переходной вероятностью ошибки p' скорость $\bar{R}_2$ была бы меньше $\bar{R}_{\text{выч}}(p')$.

В соответствии с рис. 4 построим итеративный код так, чтобы кодовые слова блокового кода, образованные строками матрицы рис. 4 содержали n_1 элементов, в которых $m_1 = [n_1\bar{R}_1] + 1$ информационные и обеспечивали вероятность ошибки на символ меньшую p'; а столбцы матрицы образовывались элемента-

ми древовидного кода со скоростью передачи $\bar{R}_2$. Для построенного кода

$$\bar{R} = \lim_{n_2 \to \infty} \frac{m_1 m_2}{n_1 n_2} = \lim_{n_2 \to \infty} \frac{([n_1 \bar{R}_1] + 1) n_2 \bar{R}_2}{n_1 n_2} \geqq \bar{R}_1 \bar{R}_2 . \tag{57}$$

Декодирование осуществляется по этапам. Сначала исходя из блокового кода производится декодирование по строкам, так что информационные и проверочные символы второго кода с вероятностью $1 - p'$ восстанавливаются без

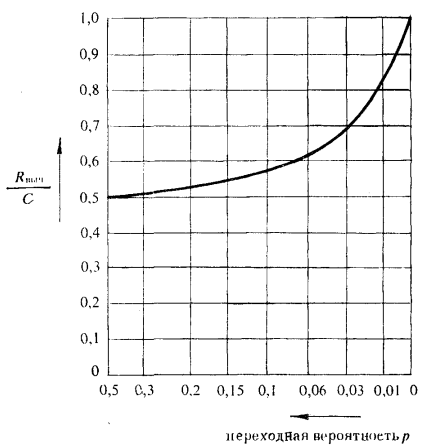

Рис. 5.

ошибки; при этом ошибки в разных символах одного столбца будут независимы. Далее исходя из второго — древовидного кода производится декодирование по столбцам; восстанавливаются информационные символы. Поскольку $\bar{R}_2 < \bar{R}_{выч}(p')$ то можно считать, что число операций N_2 при декодировании на втором этапе не зависит от ε. Первый этап (декодирование блокового кода) требует не более чем N_1 операций (N_1 — мощность блокового кода), не зависящем от ε.

Таким образом, общее число операций при декодировании на обоих этапах не зависит от ε и не превосходит $N \leqq N_1 + N_2$.

Нетрудно вывести асимптотические оценки порядка дополнительного числа операций N_1, вносимых блоковым кодом при $\bar{R} \to \bar{C}$. Имеем при $\bar{R} \to \bar{C}$:

$$\bar{R}_2 \to 1, \ \bar{R}_1 \to \bar{C} ; \tag{58}$$

$$\bar{R}_2 \leqq \bar{R}_{выч}(p') = -\log \tfrac{1}{2}(1 + 2\sqrt{(p'(1 - p'))}) \approx 1 - 2\sqrt{p'} ;$$

$$\bar{C} > \bar{R}_1 = \frac{\bar{R}}{\bar{R}_2} \geqq \frac{\bar{R}}{1 - 2\sqrt{p'}} \approx \bar{R}(1 + 2\sqrt{p'}) .$$

Как известно [9], существуют блоковые систематические коды, такие, что p' будет удовлетворять неравенству

$$p' \leqq 2^{-m_1 \varkappa (\bar{C} - \bar{R}_1)^2} , \tag{59}$$

где $\varkappa > 0$.

Отсюда

$$\log p' \leqq -m_1 \varkappa (\bar{C} - \bar{R}_1)^2$$

и следовательно, в соответствии с (58)

$$N_1 = 2^{m_1} \leqq 2^{-\log p' / [\varkappa (\bar{C} - \bar{R}_1)^2]} \approx 2^{-\log p' / [\varkappa (\bar{C} - (\bar{R} + 2\sqrt{p'}))^2]} . \tag{60}$$

Положим $2\sqrt{p'} = \beta(\bar{C} - \bar{R})$, $0 < \beta < 1$. Из (60) получаем

$$N_1 \approx 2^{-\{\log[\beta(\bar{C} - \bar{R})/2]^2\} / [\varkappa (1 - \beta)^2 (\bar{C} - \bar{R})^2]} . \tag{61}$$

Надо выбрать $0 < \beta < 1$ так, чтобы правая часть (61) была бы минимальной.

Проблема математически точного и полного определения сложности декодирования, а также получения точных оценок констант, выражающих сложность декодирования, требует дальнейших исследований.

До сих пор мы акцентрировали свое внимание на вопросах сложности кодирования и декодирования на входе и выходе канала. Аналогичная проблема возникает в задаче (двойственной кодированию) квантования (кодирования) сообщений источника при критерии верности воспроизведения, отличном от полного воспроизведения. Рассмотрим эту проблему на примере сообщения $\xi = (\xi_1, \ldots, \xi_n)$ образованного последовательностями независимых равновероятных двоичных символов $x = (x_1 \ldots x_n)$ с совпадающими пространствами $X = \tilde{X}$ (значений сообщения на выходе источника ξ и воспроизводимого сообщения $\tilde{\xi}$). Критерий верности воспроизведения задается условием

$$\frac{1}{n} \sum_{i=1}^{n} P(\xi_i \neq \tilde{\xi}_i) \leqq \varepsilon . \tag{62}$$

Скорость создания сообщений (ε-энтропия) в этом случае равна

$$\bar{H}_\varepsilon = \min \frac{1}{n} I(\xi, \tilde{\xi}) = 1 + \varepsilon \log \varepsilon + (1 - \varepsilon) \log (1 - \varepsilon) , \tag{63}$$

где минимум берется по парам случайных величин $(\xi, \tilde{\xi})$, удовлетворяющих (62).

Квантованием $\mathfrak{M}$ объема K сообщения ξ назовем любое множество $\mathfrak{M}$ последовательностей $\tilde{x}^i = (\tilde{x}_1^i \ldots \tilde{x}_n^i)$, $i = 1, \ldots, K$. Кодированием (нерандо-

мизированным) с квантованием $\mathfrak{M}$ называется отображение множества X (всех двоичных последовательностей), на множество $\mathfrak{M}$: кодирование задается функцией, определенной на множестве X и принимающей значения из множества $\mathfrak{M}$

$$\tilde{x} = f(x); \quad x \in X, \quad \tilde{x} \in \mathfrak{M}. \tag{64}$$

Задача кодирования может быть поставлена следующим образом: построить квантование $\mathfrak{M}$ наименьшего объема (т.е. множество $\mathfrak{M}$ с наименьшим числом точек) и функцию f таким образом, чтобы для пары случайных величин $(\xi, \tilde{\xi})$, где $\tilde{\xi} = f(\xi)$, было выполнено условие (62).

Имеет место следующая теорема. *Существует квантование объема* $K = 2^{n\bar{H}_\varepsilon + O(\log n)}$, *т.е.*

$$\log K = n\bar{H}_\varepsilon + O(\log n), \tag{65}$$

и кодирование, задаваемое (64), *такие, что для пары* $(\xi, \tilde{\xi})$ *выполнены условия верности воспроизведения* (62). *Множество* $\mathfrak{M}$ *двоичных последовательностей* $\tilde{x} = (\tilde{x}_1 \dots \tilde{x}_n)$ *можно выбрать так, чтобы оно образовало группу при покомпонентном сложении по* mod 2. *Обратно, по теореме Шеннона, не существует квантования объема* $K < 2^{n\bar{H}_\varepsilon}$ *удовлетворяющего* (62).

При заданном $\mathfrak{M}$ оптимальное кодирование (т.е. кодирование с минимальным значением $\sum_{i=1}^{n} P(\xi_i + \tilde{\xi}_i)$) задается функцией

$$\tilde{x} = f(x), \quad |\tilde{x} - x| = \min_{\tilde{x}' \in \mathfrak{M}} |\tilde{x}' - x|, \tag{66}$$

где $|\tilde{x} - x| = \sum_{i=1}^{n} |\tilde{x}_i - x_i|$ – расстояние Хемминга между последовательностями $\tilde{x}$ и x, т.е. для того, чтобы закодировать последовательность $x = (x_1 \dots x_n) \in X$ надо сопоставить ей точку множества $\mathfrak{M}$, находящуюся от нее на минимальном расстоянии Хемминга. Кодирование может быть осуществлено путем сравнения x со всеми точками $\tilde{x}' \in \mathfrak{M}$, и оставлением той точки $\tilde{x}$, до которой расстояние минимальное. Так как число элементов множества $\mathfrak{M}$ экспоненциально растет с ростом n, то при таком кодировании число операций (сравнений x с точками $\tilde{x}$ множества $\mathfrak{M}$) экспоненциально растет с n. Проблема построения простых методов кодирования сообщений остается открытой.

Интересно отметить, что в некоторых случаях канал как бы „согласован" с сообщением и передача сообщений не требует специального квантования и кодирования на входе канала. Например, пусть условие верности воспроизведения задается условием (62), а канал двоичный симметричный с $p = \varepsilon$. Посимвольная передача сообщения ξ по каналу, т.е. передача, при которой $y_i = x_i$, $\tilde{y}_i = \tilde{x}_i$, $i = 1, \dots, n$ обеспечивает выполнение условия (62). Отметим, что в этом

случае $\overline{H}_\varepsilon = \overline{C}$, так что никакой иной способ передачи, включающий кодирование и квантование, не может дать меньшего значения ε.

Некоторые подходы к такому согласованию будут обсуждены в следующем параграфе.

4. КОДИРОВАНИЯ И ДЕКОДИРОВАНИЯ ДЛЯ НЕПРЕРЫВНЫХ СООБЩЕНИЙ И СИГНАЛОВ

К настоящему времени определились некоторые эффективные подходы к решению проблемы передачи гауссовских сообщений по гауссовскому каналу.

В этом параграфе мы ограничимся рассмотрением проблемы кодирования и декодирования для следующей системы передачи информации. Будем считать, что источник генерирует гауссовский стационарный случайный процесс $\xi(t)$ непрерывного аргумента со спектральной плотностью $f_{\xi\xi}(\omega)$, а верность воспроизведения задается среднеквадратичным критерием (8), где $\tilde{\xi}(t)$ – значения воспроизводимых сообщений в момент времени t. Канал представляется в виде (23), где $\zeta(t)$ – гауссовский стационарный случайный процесс со спектральной плотностью $f_{\zeta\zeta}(\omega)$. Ограничения на входной сигнал $\eta(t)$ состоят в ограничении его средней мощности (10).

Передачу можно представить себе следующим образом: процесс (сообщение) $\xi(t)$ кодируется в процесс (сигнал) на входе канала $\eta(t)$, в канале $\eta(t)$ подвергается действию гауссовского шума $\zeta(t)$ независимого от $\eta(t)$. На выходе канала мы получаем процесс (сигнал) $\tilde{\eta}(t) = \eta(t) + \zeta(t)$, который декодируется в процесс (сообщения) $\tilde{\xi}(t)$. Среди методов кодирования и декодирования для непрерывных сообщений и каналов особое место занимают так называемые линейные методы; их теория проста и достаточно хорошо разработана. В рассматриваемом случае линейное кодирование и декодирование можно определить с помощью вещественных функций $A(\tau)$ и $B(\tau)$

$$\eta(t) = \int_{-\infty}^{\infty} A(\tau)\, \xi(t-\tau)\, \mathrm{d}\tau\,, \tag{67}$$

$$\tilde{\xi}(t) = \int_{-\infty}^{\infty} B(\tau)\, \tilde{\eta}(t-\tau)\, \mathrm{d}\tau\,, \tag{68}$$

или, переходя к спектральным представлениям,

$$\eta(t) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t\omega}\, \Phi_\eta(\mathrm{d}\omega) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t\omega}\, a(\omega)\, \Phi_\xi(\mathrm{d}\omega)\,, \tag{69}$$

$$\tilde{\xi}(t) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t\omega}\, \Phi_{\tilde{\xi}}(\mathrm{d}\omega) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t\omega}\, b(\omega)\, \Phi_{\tilde{\eta}}(\mathrm{d}\omega)\,, \tag{70}$$

где $\Phi_\alpha(\omega)$ – спектральная случайная мера стационарного процесса $\alpha(t)$. Очевид-

но $(\xi(t), \eta(t), \tilde{\eta}(t), \tilde{\xi}(t))$ – многомерный стационарный гауссовский случайный процесс. Из соотношений (67)–(70) и (10) следует, что

$$2\int_0^\infty (f_{\xi\xi}(\omega) + f_{\tilde{\xi}\tilde{\xi}}(\omega) - 2\,\mathrm{Re}\, f_{\tilde{\xi}\xi}(\omega))\,\mathrm{d}\omega \leqq \varepsilon^2\,, \tag{71}$$

$$2\int_0^\infty f_{\eta\eta}(\omega)\,\mathrm{d}\omega \leqq P_c\,; \tag{72}$$

$$\begin{aligned} &f_{\xi\eta}(\omega) = a(\omega) f_{\xi\xi}(\omega)\,, \quad f_{\eta\eta}(\omega) = |a(\omega)|^2 f_{\xi\xi}(\omega)\,, \\ &f_{\tilde{\eta}\tilde{\eta}}(\omega) = f_{\eta\eta}(\omega) + f_{\zeta\zeta}(\omega)\,, \\ &f_{\tilde{\eta}\tilde{\xi}}(\omega) = b(\omega) f_{\tilde{\eta}\tilde{\eta}}(\omega)\,, \quad f_{\tilde{\xi}\tilde{\xi}}(\omega) = |b(\omega)|^2 f_{\tilde{\eta}\tilde{\eta}}(\omega)\,. \end{aligned} \tag{73}$$

Задача оптимального линейного кодирования и декодирования заключается в подборе таких $a(\omega)$ и $b(\omega)$, которые минимизируют ε^2 в неравенстве (8). Как известно, при оптимальном линейном кодировании

$$|a(\omega)|^2 = \max\left(\frac{\lambda\sqrt{(f_{\xi\xi}(\omega) f_{\zeta\zeta}(\omega))} - f_{\zeta\zeta}(\omega)}{f_{\xi\xi}(\omega)},\ 0\right), \tag{74}$$

$$b(\omega) = \frac{\overline{a(\omega)} f_{\xi\xi}(\omega)}{|a(\omega)|^2 f_{\xi\xi}(\omega) + f_{\zeta\zeta}(\omega)} \tag{75}$$

и наименьшее среднеквадратичное отклонение равно

$$\varepsilon^2 = 2\int_0^\infty \min\left(\frac{1}{\lambda}\sqrt{(f_{\xi\xi}(\omega) f_{\zeta\zeta}(\omega))},\ f_{\xi\xi}(\omega)\right)\mathrm{d}\omega\,. \tag{76}$$

Постоянное λ в (74) и (76) в соответствии с условием (72) определяется из соотношения

$$2\int_0^\infty \max\left(\lambda\sqrt{(f_{\xi\xi}(\omega) f_{\zeta\zeta}(\omega))} - f_{\zeta\zeta}(\omega),\ 0\right)\mathrm{d}\omega = P_c\,. \tag{77}$$

Если при оптимальном линейном кодировании и декодировании для ε^2, определяемом из (77) выполнено условие

$$\bar{H}_\varepsilon = \bar{C}\,, \tag{78}$$

то согласно обратной теореме Шеннона линейные методы будут наилучшими среди любых методов кодирования, то есть при оптимальном линейном методе кодирования значения среднеквадратичного отклонения ε^2 будут не больше, чем при любом другом методе кодирования. Методы кодирования, для которых выполнено (78), будем для краткости называть *оптимальными по Шеннону*.

Исходим из выражений для $\overline{H}_\varepsilon$ и $\overline{C}$

$$\overline{H}_\varepsilon = \frac{1}{2\pi}\int_0^\infty \log \max\left(\frac{f_{\xi\xi}(\omega)}{\Theta^2}, 1\right) \mathrm{d}\omega\,, \tag{79}$$

$$\overline{C} = \frac{1}{2\pi}\int_0^\infty \log \max\left(\frac{\vartheta^2}{f_{\zeta\zeta}(\omega)}, 1\right) \mathrm{d}\omega\,, \tag{80}$$

где постоянные Θ^2 и ϑ^2 определяются из условий

$$2\int_0^\infty \min\left(\Theta^2, f_{\xi\xi}(\omega)\right) \mathrm{d}\omega = \varepsilon^2\,, \tag{81}$$

$$2\int_0^\infty f_{\eta\eta}(\omega)\,\mathrm{d}\omega = 2\int_0^\infty \max\left(\vartheta^2 - f_{\zeta\zeta}(\omega), 0\right) \mathrm{d}\omega = P_c\,. \tag{82}$$

Обозначим через Ω множество значений $\omega > 0$, в которых подинтегральное выражение в (77) отлично от нуля. Легко показать [19, 20], что для выполнения соотношения (78), при оптимальных линейных методах кодирования и декодирования необходимо и достаточно, чтобы

$$f_{\xi\xi}(\omega) f_{\zeta\zeta}(\omega) = \mathrm{const}\,, \tag{83}$$

$$f_{\zeta\zeta}(\omega) \leqq f_{\zeta\zeta}(\omega')\,, \quad \omega \in \Omega\,, \quad \omega' \notin \Omega\,, \quad \omega, \omega' > 0\,.$$

Множество Ω является полосой пропускания кодирующего устройства. Из выражений (80) и (82) видно, что при $\overline{C} < \infty$ процесс $\zeta(t)$ обобщенный.

Во многих исследованиях помимо ограничения на среднюю мощность сигнала (10) вводят ограничение на полосу $(\overline{\omega}, \overline{\omega} + W)$ пропускания сигналов в канале. В нашем случае это можно свести к следующему ограничению. Спектральная плотность процесса на входе канала $\eta(t)$ отлична от нуля лишь на интервалах $(-W - \overline{\omega}, -\overline{\omega})$ и $(\overline{\omega}, \overline{\omega} + W)$.

При ограниченной полосе пропускания формулы (69), (80), (77) принимают вид

$$\eta(t) = \int_{-\overline{\omega}-W}^{-\overline{\omega}} \mathrm{e}^{\mathrm{i}t\omega}\, a(\omega)\, \Phi_\xi(\mathrm{d}\omega) + \int_{\overline{\omega}}^{\overline{\omega}+W} \mathrm{e}^{\mathrm{i}t\omega}\, a(\omega)\, \Phi_\xi(\mathrm{d}\omega)\,, \tag{84}$$

$$\overline{C} = \frac{1}{2\pi}\int_{\overline{\omega}}^{\overline{\omega}+W} \log \max\left(\frac{\vartheta^2}{f_{\zeta\zeta}(\omega)}, 1\right) \mathrm{d}\omega\,, \tag{85}$$

$$2\int_{\overline{\omega}}^{\overline{\omega}+W} \max\left(\lambda\sqrt{\left(f_{\xi\xi}(\omega) f_{\zeta\zeta}(\omega)\right)} - f_{\zeta\zeta}(\omega), 0\right) \mathrm{d}\omega = P_c\,. \tag{86}$$

Для каналов с конечной полосой пропускания $(\overline{\omega}, \overline{\omega} + W)$ для выполнения (78) кроме (83) необходимо и достаточно следующего добавочного условия: мно-

жество Ω из (83) содержится в сумме интервалов $\Omega \subset (-\bar{\omega} - W, -\bar{\omega}) \cup (\bar{\omega}, \bar{\omega} + W)$, в частности, когда

$$f_{\xi\xi}(-\omega) = f_{\xi\xi}(\omega) = \begin{cases} a_\xi = \text{const} & \text{при } \omega \in (\bar{\omega}, \bar{\omega} + W), \\ 0 & \text{при других значениях } \omega \end{cases}$$

и

$$f_{\zeta\zeta}(-\omega) = f_{\zeta\zeta}(\omega) = a_\zeta = \text{const} \qquad \text{при} \quad \omega \in (\bar{\omega}, \bar{\omega} + W),$$

то выполнено (78) и оптимальные линейные методы кодирования сводятся к умножению процессов $\xi(t)$ и $\tilde{\eta}(t)$ на некоторые константы. В некоторых случаях, когда условия (83) не выполнены, мы можем уменьшить значения ε^2 и даже добиться выполнения (78) с помощью процедуры, называемой перестановкой спектра процесса $\xi(t)$ [20]. Суть этой процедуры мы рассмотрим сначала на одном частном случае.

Пусть гауссовские процессы $\xi(t)$ и $\eta(t)$ представляются в виде

$$\xi(t) = \int_{-\bar{\omega}-W}^{-\bar{\omega}} e^{it\omega}\, \Phi_\xi(d\omega) + \int_{\bar{\omega}}^{\bar{\omega}+W} e^{it\omega}\, \Phi_\xi(d\omega)$$

и

$$f_{\xi\xi}(-\omega) = f_{\xi\xi}(\omega) = \begin{cases} a'_\xi & \text{при } \omega \in (\bar{\omega}, \bar{\omega} + W/2), \\ a''_\xi & \text{при } \omega \in (\bar{\omega} + W/2, \bar{\omega} + W),\ a'_\xi > a''_\xi, \end{cases} \tag{87}$$

$$f_{\zeta\zeta}(-\omega) = f_{\zeta\zeta}(\omega) = \begin{cases} a'_\zeta & \text{при } \omega \in (\bar{\omega}, \bar{\omega} + W/2), \\ a''_\zeta & \text{при } \omega \in (\bar{\omega} + W/2, \bar{\omega} + W),\ a'_\xi > a''_\zeta. \end{cases} \tag{88}$$

Из (87) и (88) видно, что из $f_{\zeta\zeta}(\omega') \geqq f_{\zeta\zeta}(\omega'')$ следует

$$f_{\xi\xi}(\omega') \geqq f_{\xi\xi}(\omega''),$$

$$\omega', \omega'' \in (-\bar{\omega} - W, -\bar{\omega}) \cup (\bar{\omega}, \bar{\omega} + W). \tag{89}$$

Процесс

$$\begin{aligned} \xi^*(t) = {} & e^{itW/2} \int_{-\bar{\omega}-W}^{-\bar{\omega}-W/2} e^{it\omega}\, \Phi_\xi(d\omega) + e^{-itW/2} \int_{-\bar{\omega}-W/2}^{-\bar{\omega}} e^{it\omega}\, \Phi_\xi(d\omega) + \\ & + e^{itW/2} \int_{\bar{\omega}}^{\bar{\omega}+W/2} e^{it\omega}\, \Phi_\xi(d\omega) + e^{-itW/2} \int_{\bar{\omega}+W/2}^{\bar{\omega}+W} e^{it\omega}\, \Phi_\xi(d\omega) \end{aligned} \tag{90}$$

получается „перестановкой спектра" из процесса $\xi(t)$. Очевидно

$$f_{\xi^*\xi^*}(-\omega) = f_{\xi^*\xi^*}(\omega) = \begin{cases} a''_\xi & \text{при } \omega \in (\bar{\omega}, \bar{\omega} + W/2), \\ a'_\xi & \text{при } \omega \in (\bar{\omega} + W/2, \bar{\omega} + W) \end{cases}$$

и соотношения (89) переходят в следующие: из

$$f_{\zeta\zeta}(\omega') \geqq f_{\zeta\zeta}(\omega'')$$

следует

$$f_{\xi^*\xi^*}(\omega') \leqq f_{\xi\xi}(\omega'') \,. \tag{91}$$

Для процесса $\xi^*(t)$ характерно, что в полосе $(\bar{\omega} + W/2, \bar{\omega} + W)$ где его интенсивность (т.е. $f_{\xi^*\xi^*}(\omega)$) больше, там интенсивность шума ($f_{\zeta\zeta}(\omega)$) меньше и наоборот.

Нетрудно убедиться, что оптимальное линейное кодирование и декодирование для $\xi^*(t)$ приводят к процессу $\tilde{\xi}^*(t)$ с среднеквадратичным отклонением $\varepsilon^2 = E(\xi^*(t) - \tilde{\xi}^*(t))^2$ меньшем, чем для $\xi(t)$, в частности, если $a''_\xi = a'_\zeta = 0$, то $\varepsilon^2 = 0$ и процесс $\xi^*(t)$ восстанавливается без искажений, если же $a''_\xi a'_\zeta = a'_\xi a''_\zeta$, то выполнено соотношение (78).

Рассмотрим теперь перестановку спектра процесса $\tilde{\xi}^*(t)$, заменяя в (90) ζ, ξ^* на $\tilde{\xi}^*$, $\tilde{\xi}$ и $e^{itW/2}$ на $e^{-itW/2}$, очевидно

$$E(\xi^*(t) - \tilde{\xi}^*(t))^2 = E(\xi(t) - \tilde{\xi}(t))^2 \,.$$

Таким образом, сочетание перестановок спектра с оптимальным линейным кодированием и декодированием приводит к уменьшению среднеквадратичного отклонения воспроизводимого сообщения от сообщения на выходе источника.

Вообще говоря, процесс $\xi^*(t)$, получаемый в результате конечного числа преобразований (86) или являющийся пределом таких процессов в среднеквадратичном*, мы также будем называть процессом с перестановкой спектра $\xi(t)$.

Легко убедиться, что процесс $\xi^*(t)$, являющийся „перестановкой" процесса $\xi(t)$, представляется в виде

$$\xi^*(t) = \int_{-\infty}^{\infty} e^{i\varphi(\omega)t} \, \Phi_\xi(d\omega) \,. \tag{92}$$

Здесь $\varphi(\omega) = -\varphi(-\omega)$ – измеримая функция и при любых $\omega_1 < \omega_2$

$$\mu(\omega' : \omega_1 \leqq \varphi(\omega') \leqq \omega_2) = \omega_2 - \omega_1 \,, \tag{93}$$

где $\mu(.)$ – мера Лебега, и, наоборот, если процесс $\xi^*(t)$ допускает представление (92), то он является перестановкой спектра $\xi(t)$. Соотношение (93) выражает тот факт, что переход от $\xi(t)$ к $\xi^*(t)$ не меняет „ширины спектра". Можно показать, что существует такая перестановка спектра процесса $\xi(t)$, что из $f_{\zeta\zeta}(\omega') \geqq f_{\zeta\zeta}(\omega'')$ следует

$$f_{\xi^*\xi^*}(\omega') \leqq f_{\xi^*\xi^*}(\omega'') \,. \tag{94}$$

* Последовательность процессов $\alpha_1(t), \alpha_2(t), \ldots$ сходится к процессу $\alpha(t)$ в среднеквадратичном, если $\lim\limits_{n\to\infty} \alpha_n(t) = \alpha(t)$ т.е. $\lim\limits_{t\to\infty} E(\alpha(t) - \alpha_n(t))^2 = 0$.

Легко видеть, что среднеквадратичное отклонение при оптимальном линейном кодировании и декодировании процесса $\xi^*(t)$ будет не больше, чем среднеквадратичное отклонение при любых других перестановках спектра $\xi(t)$, линейном кодировании и декодировании.

Отсюда вытекает, что последовательность операций: *перестановка спектра* (94), *оптимальное линейное кодирование и декодирование, „обратная" перестановка спектра дает наименьшее значение*

$$\varepsilon^2 = E(\xi(t) - \tilde{\xi}(t))^2$$

по сравнению с любыми другими перестановками, линейном кодировании и декодировании.

Однако, следует отметить, что перестановка спектра и линейное кодирование не приводит к уменьшению ε^2 в случае, когда

$$E\,\xi^2(t) = 2\int_0^\infty f_{\xi\xi}(\omega)\,\mathrm{d}\omega = P_c\,,$$

$$f_{\xi\xi}(-\omega) = f_{\xi\xi}(\omega) = \begin{cases} a_\xi = \text{const} & \text{при } \omega \in (0, W_1)\,, \\ 0 & \text{при } \omega \notin (0, W_1);\ \omega,\ W > 0\,; \end{cases}$$

а полоса пропускания канала $(0, W_2)$ не совпадает с полосой $(0, W_1)$ (т.е. $W_1 \neq W_2$) и $f_{\zeta\zeta}(\omega) = a_\zeta = \text{const}$. Построение оптимальных или сколь угодно близких к оптимальным (т.е. таких, что разность $\overline{C} - \overline{H}_\varepsilon$ сколь угодно мала) эффективных методов кодирования и декодирования является здесь открытой проблемой.

Для решения этой проблемы достаточно решить следующую задачу (называемую иногда задачей сокращения спектра).

Пусть $\xi(t)$ и $\zeta(t)$ гауссовские процессы с

$$f_{\xi\xi}(-\omega) = f_{\xi\xi}(\omega) = \begin{cases} a_\xi \neq 0 & \text{при } \omega \in (0, W_1)\,, \\ 0 & \text{при } \omega \notin (0, W_1);\ W_1,\ \omega > 0\,; \end{cases}$$

$$f_{\zeta\zeta}(-\omega) = f_{\zeta\zeta}(\omega) = \begin{cases} a_\zeta \neq 0 & \text{при } \omega \in (0, W_2)\,, \\ 0 & \text{при } \omega \notin (0, W_2);\ W_2,\ \omega > 0\,; \end{cases}$$

$$W_1 \neq W_2\,.$$

При любом $\delta > 0$, построить функционалы $\varphi(.)$ и $\psi(.)$, определенные на пространстве $L^2(-T, T)$ вещественных функций с суммированием квадратом, заданных на интервале $(-T, T)$ (T зависит от δ) такие, что

$$E(\zeta(t) - \tilde{\zeta}(t))^2 \leqq \varepsilon^2 + \delta\,;$$

$$\frac{1}{2\pi} W_1 \log \frac{a_\xi W_1}{\varepsilon^2} = \frac{1}{2\pi} W_2 \log \left(1 + \frac{a_\xi}{a_\zeta}\right)$$

(левая часть последнего равенства равна ε-энтропии $\bar{H}_\varepsilon$ источника определяемого процессом $\xi(t)$, а правая часть – пропускной способности канала с $P_c = 2a_\zeta W_2$, $f_{\zeta\zeta}(\omega) = a_\zeta$ и полосой пропускания $(0, W_2)$). Случайная величина $\tilde{\xi}(t)$ удовлетворяет цепочке равенств

$$\tilde{\xi}(t) = \psi_t(\tilde{\eta}_{t-T}^{t+T}), \quad \tilde{\eta}(t) = \eta(t) + \zeta(t), \quad \eta(t) = \varphi_t(\xi_{t-T}^{t+T}).$$

Здесь

$$\varphi_t(f(\tau)) = \varphi(f(\tau - t)), \quad \psi_t(f(\tau)) = \psi(f(\tau - t)),$$

где $f(\tau)$ – функция с суммированным квадратом, заданная на интервале $(t - T, t + T)$. Значения случайных величин

$$\xi_{t-T}^{t+T} = \{\xi(\tau);\ t - T \leqq \tau \leqq t + T\}$$

– функции заданные на интервале $(t - T, t + T)$.

Нахождение простых методов построения функций φ и ψ в сочетании с перестановкой спектра и линейным кодированием дало бы возможность для произвольных гауссовских источников и каналов построить простые методы кодирования и декодирования, сколь угодно близкие к оптимальным.

(Поступило 9 июня 1965 г.)

ЛИТЕРАТУРА

[1] Shannon C.: A mathematical theory of communication. Bell. System Techn. J. *27* (1948), 3, 379—423; *27* (1948), 4, 623—656. (Руск. пер. Шеннон К.: Математическая теория связи. Сб. Шеннон К.: Работы по теории информации и кибернетике. Изд. ИЛ., М. 1963, 243—332.)

[2] Колмогоров А. Н.: Теория передачи информации. Сб. Сессия АН СССР по науч. пробл. автом. пр-ва. 1956. Пленарное заседание. Изд. АН СССР 1957, 66—99. Дискуссия 148—161.

[3] Добрушин Р. Л.: Математические вопросы шенноновской теории оптимального кодирова ния информации. Сб. Проблемы передачи информации. Изд. АН СССР, М. *10* (1961), 63—107.

[4] Добрушин Р. Л.: Общая формулировка основной теоремы Шеннона в теории информации. Усп. мат. наук *14* (1959), 6, 3—104.

[5] Shannon C.: Certain results in coding theory for noisy channels. Information and Control *1* (1957), 6—25. (Русск. пер. Шеннон К.: Некоторые результаты теории кодирования для каналов с шумами. Сб. Шеннон К.: Работы по теории информации и кибернетике. Изд. ИЛ., М. 1963, 509—531.)

[6] Пинскер М. С.: Информация и информационная устойчивость случайных величин и процессов. Изд. АН СССР, М. 1960.

[7] Пинскер М. С.: Источники сообщений. Сб. Проблемы передачи информации. Изд. АН СССР, М. *14* (1963), 5—20.

[8] Пинскер М. С.: Гауссовские источники. Сб. Проблемы передачи информации. Изд. АН СССР, М. *14* (1963), 50—100.

[9] Fano R. M.: Transmission of information. MIT Press and J. Wiley, New York - London 1961.

[10] Wozencraft J. M., Reiffen B.: Sequential decoding. MIT Press and J. Wiley, New York — London 1961. (Русс. пер. в книге Возенкрафт Дж., Райфен Б.: Последовательное декодирование. ИЛ, 1963.)

[11] Peterson W.: Error-correcting codes. MIT Press — J. Wiley, New York — London 1961. (Русск. пер. Питерсон У.: Коды исправляющие ошибки. Изд. Мир. 1964.)

[12] Ziv J.: Coding and decoding for time-discrete amplitude continuous memoryless channels. IRE Trans. Inform. Theory *8* (September 1962).

[13] Ziv J.: Successive decoding scheme for memoryless channels. IEEE Trans. Inform. Theory *9* (1963), 2, 97—105.

[14] Fano R. M.: A heuristic discussion of probabilistic decoding. IEEE Trans. Inform. Theory *9* (1963), 2, 64—74. (Русск. пер. Фано Р. М.: Эвристическое обсуждение вероятностного декодирования. Сб. Теория кодирования. Изд. МИР, М. 1964, 166—198.)

[15] Callager R. C.: Low density parity-check codes. IRE Trans. Inform. Theory *8* (1962), 1, 21—28. (Русск. пер. Галлагер Р. Г.: Коды с малой плотностью проверок на четность, Сб. Теория кодирования. Изд. МИР, М. 1964, 139—166.)

[16] Добрушин Р. Л.: По поводу последовательного декодирования методом Возенкрафта-Рейффена. Сб. Проблемы кибернетики *12* (1964), 114—123.

[17] Пинскер М. С.: О сложности декодирования. Проблемы передачи информации *1* (1965), 115—118.

[18] Цыбаков Б. С.: Шенноновская схема для гауссовского сообщения с равномерным спектром и флуктуационным шумом. Радиотехн. и электроника *6* (1964), 4, 649.

[19] Овсеевич И. А. и Пинскер М. С.: Оптимальное линейное предыскажение и корректирование. Изв. АН СССР — Техническая кибернетика (1963), 5, 54—60.

[20] Овсеевич И. А. и Пинскер М. С.: Согласование источника сообщений с каналом методом перестановки спектров. Изв. АН СССР — Техническая кибернетика (1965), 2, 81—86.

[21] Grenander U., Szego G.: Toeplitz forms and their application. University of California Press, Berkeley and Los Angeles 1958. (Русск. пер. Гренандер У., Сеге Г.: Теплицевы формы и их приложения. Изд. ИЛ, М. 1961.)

Některé matematické problémy teorie přenosu informace

M. S. PINSKER

Článek je přehledem pravděpodobnostních hledisek některých směrů teorie přenosu informace.

Pro širokou třídu zdrojů a kanálů jsou vyloženy jednotlivé úplné výsledky vycházející ze Shannonových vět. Pro libovolný Gaussův kanál jsou uvedeny horní odhady pravděpodobnosti chyby při optimálním kódování a dekódování.

Dále následuje diskuse o problému složitosti kódování a dekódování pro zdroje a kanály.

Dokazuje se, že se pro binární symetrický kanál dá sestrojit takové skupinové kódování a dekódování, že pro libovolnou rychlost přenosu, menší než je kapacita kanálu, bude počet aritmetických operací omezen nějakou konstantou nezávisející na pravděpodobnosti chyby. Pro Gaussovy zdroje a kanály jsou uvedeny nutné a postačující podmínky, aby lineární předskreslení a náprava související s permutací spektra byly nejlepší ze všech možných metod kódování a dekódování, a tedy byly optimální v Shannonově smyslu. Nakonec je probrán problém konstrukce optimálních metod kódování a dekódování pro libovolné Gaussovy zdroje a kanály.

Марк Семенович Пинскер, Институт проблем передачи информации АН СССР, Авиамоторная ул. 8а, корп. 2, Москва Е-24, СССР.