А.Г.Дьячков

Теория информации

§2. Передача дискретных сообщений по двоичному каналу связи

Содержание

2.1 Определения и обозначения

а) Пусть $(0,1)^N$ — множество из всех 2^N двоичных последовательностей длины N. Элементы

$$\mathbf{x} = (x_1, \ldots, x_N) \in (0,1)^N, \quad \mathbf{y} = (y_1, \ldots, y_N) \in (0,1)^N, \quad x_i, y_i \in (0,1)$$

будем называть *словами* длины N из двоичных (0,1) букв.

б) Число единичных позиций слова $\mathbf{x}$ называется *весом Хэмминга* слова $\mathbf{x}$ и обозначается символом

$$|\mathbf{x}| = \sum_{i=1}^{N} x_i.$$

в) Сложение *по mod 2* слов **x, y** обозначим символом

$$\mathbf{x}\oplus\mathbf{y} = (x_1\oplus y_1,\dots,x_N\oplus y_N), \quad 0\oplus 0 = 1\oplus 1 = 0,\ 0\oplus 1 = 1\oplus 0 = 1.$$

г) *Расстоянием Хэмминга* называется число несовпадающих позиций в словах **x** и **y**

$$\rho(\mathbf{x},\mathbf{y}) = |\mathbf{x}\oplus\mathbf{y}|.$$

д) *Кодом* объема M, длины N называется двоичная (из 0 и 1) матрица

$$X = \|x_i(j)\|, \quad i = \overline{1,N},\ j = \overline{1,M},$$

столбцы которой $\mathbf{x}(m) = \Big(x_1(m),\dots,x_N(m)\Big)$ называются *кодовыми словами* кода X. Число

$$R := \frac{\log M}{N}$$

называется *скоростью* кода. Для фиксированного значения параметра $R>0$ код X объема

$$M = \Big\lceil \exp\{RN\}\Big\rceil, \quad (\lceil . \rceil \text{ -- символ верхней целой части})$$

длины N называется (N,R)-*блоковым кодом* с длиной блока N и скоростью R. Важный частный случай (N,R)-кодов, называемых *линейными* блоковыми кодами, описан в разделе **2.2**.

Замечание. Под символами log и exp здесь и в дальнейшем подразумеваются либо *двоичные*, либо *натуральные* логарифмы и экспоненты. Про скорость R будем говорить, что она измеряется либо в двоичных, либо в натуральных *единицах* соответственно.

2.2 Линейный код

Пусть k, $1\le k<N$, – целое число, а $M := 2^k$. Числу $m = \overline{1,M}$ поставим во взаимно-однозначное соответствие последовательность $\mathbf{u} = (u_1,\dots,u_k)$ длины k, которая называется информационным словом для сообщения m и представляет собой двоичное разложение числа $m-1$, т.е.

$$m \Leftrightarrow \mathbf{u} \quad : \quad m-1 = \sum_{i=1}^{k} u_i 2^{i-1},\ u_i = \overline{0,1}.$$

Пусть $G = \left\|g_i(j)\right\|, \quad i = \overline{1,N},\ j = \overline{1,k}$ — двоичная (из 0 и 1) матрица, состоящая из N строк и k столбцов, столбцы которой линейно независимы в линейном пространстве $(0,1)^N$ над полем двоичных чисел. *Линейный код* $X = (\mathbf{x}(1), \ldots, \mathbf{x}(M))$, объема $M = 2^k$ задается матрицей G, называемой *порождающей* матрицей кода. Сообщение $m = \overline{1,M}$ кодируется этим кодом в кодовое слово $\mathbf{x}(m)$ по правилу

$$m \quad \Leftrightarrow \quad \mathbf{u} \quad \Leftrightarrow \quad \mathbf{x}(m) = G\mathbf{u},$$

где операция произведения производится в двоичном поле. Скорость кода в двоичных единицах $R = k/N$.

Пример. $N = 3; \quad k = 2; \quad M = 2^2 = 4; \quad G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}; \quad R = \frac{2}{3}$

$$\begin{array}{lllll} m = 1 & \Leftrightarrow & \mathbf{u} = (0,0) & \to & \mathbf{x}(1) = (0,0,0) \\ m = 2 & \Leftrightarrow & \mathbf{u} = (0,1) & \to & \mathbf{x}(2) = (0,1,1) \\ m = 3 & \Leftrightarrow & \mathbf{u} = (1,0) & \to & \mathbf{x}(3) = (1,0,1) \\ m = 4 & \Leftrightarrow & \mathbf{u} = (1,1) & \to & \mathbf{x}(4) = (1,1,0) \end{array} \qquad X =$$

$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$ — код.

2.3 Математическая модель системы связи

a) Пусть $[M]$ — множество натуральных чисел от 1 до M.*Сообщением источника* (*передаваемым* сообщением) назовем случайную величину ξ, имеющую равномерное распределение на $[M]$, т.е.

$$\Pr\{\xi = m\} = \frac{1}{M}, \quad m = \overline{1,M}.$$

Пусть ξ приняло значение m. Для передачи сообщения m адресату оно *кодируется* кодом X в кодовое слово $\mathbf{x}(m)$, которое является *сигналом на входе канала связи.*

b) *Двоичный канал связи* создает *шум* $\mathbf{z} = (z_1, z_2, \ldots, z_N) \in (0,1)^N$ и на выходе канала наблюдается выходной сигнал (слово)

$$\mathbf{y} = \mathbf{x}(m) \oplus \mathbf{z}.$$

Будем считать выполненным естественное предположение о том, что сообщение источника ξ и шум в канале $\mathbf{z}$ являются *независи-*

мыми случайными величинами. Поэтому распределению вероятностей $\mathbf{z}$ взаимно-однозначно соответствует *условное распределение*:

$$P_N(\mathbf{y}|\mathbf{x}) \quad \Leftrightarrow \quad \underset{\text{"выход"}}{\mathbf{y}} \quad = \quad \underset{\text{"вход"}}{\mathbf{x}} \quad \oplus \quad \underset{\text{"шум"}}{\mathbf{z}},$$

где $P_N(\mathbf{y}|\mathbf{x})$ — вероятность получить на выходе канала слово $\mathbf{y}$, если на входе канала было слово $\mathbf{x}$.

Пример b1. Пусть компоненты шума $\mathbf{z} = (z_1, z_2, \ldots, z_N)$ являются независимыми случайными величинами с одинаковым распределением

$$z_i = \begin{cases} 1, & \text{с вероятностью} \quad p, \\ 0, & \text{с вероятностью} \quad 1-p, \end{cases} \qquad 0 < p < 1/2,\ i = \overline{1, N}.$$

Это означает, что для любого слова $\mathbf{e} = (e_1, \ldots, e_N)$, $\mathbf{e} \in (0,1)^N$ вероятность

$$\Pr\{\mathbf{z} = \mathbf{e}\} = \prod_{i=1}^{N} p^{e_i}(1-p)^{1-e_i} = p^{|\mathbf{e}|}(1-p)^{N-|\mathbf{e}|}.$$

Канал с таким распределением шума называется *двоичным симметричным каналом* без памяти (ДСК) с вероятностью ошибки p, $0 < p < 1/2$, при передаче одного двоичного символа. Для ДСК условное распределение

$$P_N(\mathbf{y}|\mathbf{x}) = p^{\rho(\mathbf{x},\mathbf{y})}(1-p)^{N-\rho(\mathbf{x},\mathbf{y})}.$$

Пример b2. Зафиксируем целое число t, $1 \le t < N$, и предположим, что шум $\mathbf{z} = (z_1, \ldots, z_N)$ имеет равномерное распределение на сфере Хэмминга радиуса t, т.е. для любого слова $\mathbf{e} = (e_1, \ldots, e_N) \in (0,1)^N$ вероятность

$$\Pr\{\mathbf{z} = \mathbf{e}\} = \begin{cases} (C_N^t)^{-1}, & \text{если} \quad |\mathbf{e}| = t, \\ 0, & \text{если} \quad |\mathbf{e}| \ne t. \end{cases}$$

Этот канал назовем *каналом с равновесным шумом* (КРШ). Для КРШ условное распределение

$$P_N(\mathbf{y}|\mathbf{x}) = \begin{cases} (C_N^t)^{-1}, & \text{если} \quad \rho(\mathbf{x},\mathbf{y}) = t, \\ 0, & \text{если} \quad \rho(\mathbf{x},\mathbf{y}) \ne t. \end{cases}$$

c) Для *декодирования*, т.е. для вычисления *сообщения адресата* $\tilde{\xi}$ по выходному сигналу $\mathbf{y}$, зададим разбиение множества $(0,1)^N$ слов $\mathbf{y} = (y_1, \ldots, y_N)$ на $M+1$ непересекающихся подмножеств

$$(0,1)^N = D_1 + D_2 + \cdots + D_M + D_{M+1}.$$

Другими словами, определим на $(0,1)^N$ *декодирующую* функцию

$$D(\mathbf{y}) := m \quad \text{если } \mathbf{y} \in D_m,$$

которая принимает значения в $[M+1]$. Положим, по определению, сообщение адресата

$$\tilde{\xi} := D(\mathbf{y}) = D(\mathbf{x}(\xi) \oplus \mathbf{z}),$$

где сигнал на выходе канала $\mathbf{y} = \mathbf{x}(\xi) \oplus \mathbf{z}$ определяется сообщением источника ξ, кодом X и шумом канала $\mathbf{z}$. Случай, когда $\tilde{\xi} = M+1$, называется *отказом* от принятия решения о сообщении источника ξ.

Приведем некоторые важные примеры декодирующих функций. Отметим, что при задании $D(\mathbf{y})$ можно использовать знание кода $X = \Big(\mathbf{x}(1), \ldots, \mathbf{x}(M)\Big)$ и условных вероятностей $P_N\Big(\mathbf{y}|\mathbf{x}(m)\Big), \quad m = \overline{1, M}$.

Пример 1. Декодирование по *максимуму правдоподобия* (МП-декодирование). Оно определяется следующим образом. При заданном $\mathbf{y}$ положим $D_{\text{МП}}(\mathbf{y}) = m$, где $m \in [M]$ — наименьшее среди чисел $m' = \overline{1, M}$, для которых достигается

$$\max_{m' = \overline{1,M}} P_N\Big(\mathbf{y}|\mathbf{x}(m')\Big) = P_N(\mathbf{y}|\mathbf{x}(m)\Big). \tag{$*$}$$

Отметим, что для МП-декодирования множества $D_{M+1}^{(\text{МП})}$ — пусто,а при $m = \overline{1, M}$ множество

$$D_m^{(\text{МП})} := \Big\{\mathbf{y} \quad : \quad D_{\text{МП}}(\mathbf{y}) = m\Big\}.$$

Пример 2. *Мягкое МП-декодирование кода* X. При заданном $\mathbf{y}$ положим $\hat{D}_{\text{МП}}(\mathbf{y}) = m$, если существует *единственное* значение m, $1 \le m \le M$, для которого достигается максимум в $(*)$. В противном случае положим $\hat{D}_{\text{МП}}(\mathbf{y}) = M+1$, т.е. *откажемся* от принятия решения.

Пример 3. Декодирование по *минимуму расстояния* (МР-декодирование). Это правило декодирования *универсально* в том смысле, что не зависит от условных вероятностей канала $P_N(\mathbf{y}|\mathbf{x})$. При заданном $\mathbf{y}$ положим $D_{\text{МР}}(\mathbf{y}) = m$, где $m = \overline{1, M}$ — *наименьшее* среди чисел $m' = \overline{1, \text{M}}$, для которых достигается

$$\min_{m' = \overline{1,M}} \rho(\mathbf{x}(m'), \mathbf{y}) = \rho(\mathbf{x}(m), \mathbf{y}). \tag{$**$}$$

Здесь, как и в **примере 1**, множество $D_{M+1}^{(\mathrm{MP})}$ — пусто.

Пример 4. *Мягкое декодирование кода X по минимуму расстояния.* При заданном $\mathbf{y}$ положим $\hat{D}_{\mathrm{MP}}(\mathbf{y}) = m$, если существует *единственное* значение m, $1 \le m \le M$, для которого достигается минимум в $(**)$. В противном случае положим $\hat{D}_{\mathrm{MP}}(\mathbf{y}) = M+1$, т.е. *откажемся* от принятия решения.

Замечание. Отметим, что для ДСК МР-декодирование совпадает с МП-декодированием, а для КРШ это не так.

г) Событие $\left\{\xi \ne \tilde{\xi}\right\}$ называется *ошибкой при блоковой передаче дискретного сообщения* ξ по двоичному каналу связи. Противоположное событие $\left\{\xi = \tilde{\xi}\right\}$ называется *правильной* передачей. В качестве *критерия* такой передачи рассмотрим *вероятность ошибки*

$$\mathcal{P}(X) := \Pr\left\{\xi \ne \tilde{\xi}\right\} = \frac{1}{M}\sum_{m=1}^{M} \mathcal{P}_m(X), \qquad (1)$$

где *условная вероятность ошибки*

$$\mathcal{P}_m(X) := \Pr\left\{\xi \ne \tilde{\xi}\middle|\xi = m\right\} = \sum_{\mathbf{y}\in\overline{D}_m} P_N\Big(\mathbf{y}|\mathbf{x}(m)\Big), \qquad (2)$$

а символом $\overline{D}_m$ обозначено дополненеие к D_m.

Заметим, что используя символ декодирующей функции $D(\mathbf{y})$, *вероятность правильного решения* $1 - \mathcal{P}(X) = \Pr\left\{\xi = \tilde{\xi}\right\} = \frac{1}{M}\sum_{m=1}^{M} \Pr\left\{\xi = \tilde{\xi}\middle|\xi = m\right\}$ можно записать в виде

$$1 - \mathcal{P}(X) = \tfrac{1}{M}\textstyle\sum_{m=1}^{M}\sum_{\mathbf{y}\in D_m} P_N\Big(\mathbf{y}|\mathbf{x}(m)\Big) = \tfrac{1}{M}\sum_{\mathbf{y}} P_N\left(\mathbf{y}\middle|\mathbf{x}\Big(D(\mathbf{y})\Big)\right) \le$$
$$\le \tfrac{1}{M}\textstyle\sum_{\mathbf{y}} P_N\left(\mathbf{y}\middle|\mathbf{x}\Big(D_{\text{МП}}(\mathbf{y})\Big)\right) = 1 - \mathcal{P}_{\text{МП}}(X),$$

где неравенство следует из определения МП-декодирования, а символом $\mathcal{P}_{\text{МП}}(X)$ обозначена вероятность ошибки кода X при использовании алгоритма МП-декодирования. Следовательно, при заданном коде X МП-декодирование *максимизирует* вероятность правильного решения $1 - \mathcal{P}(X)$, т.е. *минимизирует* вероятность ошибки $\mathcal{P}(X)$.

Для кода X и соответствующего этому коду мягкого МП-декодирования обозначим через $\chi_m(X,\mathbf{y})$, $m=\overline{1,M}$ *характеристическую* функцию множества

$$\overline{\hat{D}_m^{(\text{МП})}} = \Big\{\mathbf{y} \quad : \hat{D}_{\text{МП}}(\mathbf{y}) \neq m\Big\}.$$

Согласно определению мягкого МП-декодирования, можно написать

$$\chi_m(X,\mathbf{y}) := \begin{cases} 1, & m' \neq m, \ , \ \text{p} \\ & P_N\Big(\mathbf{y}|\mathbf{x}(m')\Big) \geq P_N\Big(\mathbf{y}|\mathbf{x}(m)\Big), \\ 0, & m' \neq m \\ & P_N\Big(\mathbf{y}|\mathbf{x}(m')\Big) < P_N\Big(\mathbf{y}|\mathbf{x}(m)\Big), \end{cases} \tag{3}$$

а формула (2) для условной вероятности ошибки примет вид

$$\mathcal{P}_m(X) = \sum_{\mathbf{y}} P_N\Big(\mathbf{y}|\mathbf{x}(m)\Big)\chi_m(X,\mathbf{y}). \tag{2}$$

Непосредственное вычисление $\mathcal{P}_m(X)$ по формулам (2) — (3) для оптимального, (т.е. минимизирующего вероятность ошибки) кода X в общем случае является *нереальным.* В разделе **2.4** мы рассмотрим частный случай задачи вычисления $\mathcal{P}_m(X)$, $m=\overline{1,M}$, для ДСК И КРШ, когда $M=2$. В разделе **2.5** будет описан предложенный Шенноном *метод случайного кодирования* для построения верхней оценки вероятности ошибки $\mathcal{P}(X)$ оптимального кода X на примере частного случая КРШ.

В **§§5, 6** будут описаны методы Шеннона для получения границ вероятности ошибки дискретного канада без памяти (ДКБП), частным случаем которого является ДСК.

2.4 Передача двух сообщений

2.4.1 Случай ДСК, M=2, $0<p<1/2$

Пусть при передаче сообщения источника ξ, принимающего два значения, используется код $X = \Big(\mathbf{x}(1),\mathbf{x}(2)\Big)$ со словами $\mathbf{x}(1) = (0,0,\ldots,0)$, $\mathbf{x}(2)=(1,1,\ldots,1)$ длины N. Если при декодировании по минимуму расстояния произошла ошибка, то это означает, что в канале произошло событие $\sum_{i=1}^N z_i = |\mathbf{z}| \geq N/2$. Следовательно, для любого

$m = \overline{1,2}$ условная вероятность ошибки

$$\mathcal{P}_m(X) = \Pr\left(\sum_{i=1}^{N} z_i \geq N/2\right) = \sum_{i=N/2}^{N} C_N^i p^i (1-p)^{N-i}.$$

Задача 2.1. Воспользовавшись законом больших чисел, докажите, что

$$\lim_{N\to\infty} \mathcal{P}_m(X) = 0,$$

если $0 < p < 1/2$.

Задача 2.2. Пусть s, $1 \leq s \leq N$, — целое число, и в коде $X = \Big(\mathbf{x}(1), \mathbf{x}(2)\Big)$, состоящем из двух слов, расстояние между словами $\rho\Big(\mathbf{x}(1); \mathbf{x}(2)\Big) = s$. Докажите, что при передаче по ДСК с использованием МР-декодирования условная вероятность ошибки

$$\mathcal{P}_m(X) = \sum_{i=\lfloor s/2 \rfloor}^{s} C_s^i p^i (1-p)^{s-i}.$$

Введем функцию параметра x, $0 < x < 1$,

$$h(x) = -x \log x - (1-x)\log(1-x),$$

которую в дальнейшем будем называть *двоичной энтропией*.

Задача 2.3. Проверьте, что $h(x)$ — выпукла вверх, симметрична относительно $x = 1/2$, достигает максимума при $x = 1/2$, $h(1/2) = \log 2$ и с помощью известной *формулы Стирлинга*

$$\log N! = N \log \frac{N}{e} + \frac{1}{2}\log(2\pi N) + o(N),$$

где $0 < o(N) < \log e/12N$, докажите, что при любом фиксированном x, $0 < x < 1$,

$$C_N^{\lfloor Nx \rfloor} \leq \exp\Big\{N h(x)\Big\}$$

и существует

$$\lim_{N\to\infty} \frac{\log C_N^{\lfloor Nx \rfloor}}{N} = h(x).$$

Замечание. Более точные оценки, устанавливаемые с помощью неравенств Стирлинга, имеют вид

$$\sqrt{\frac{n}{8j(n-j)}} \leq C_n^j \exp\left\{-nh\left(\frac{j}{n}\right)\right\} \leq \sqrt{\frac{n}{2\pi j(n-j)}}.$$

Для фиксированных $0 < p < x < 1$ определим величину

$$Q_N(p,x) = \Pr\left\{\sum_{i=1}^{N} z_i > xN\right\} = \sum_{i=\lfloor xN \rfloor}^{N} C_N^i p^i (1-p)^{N-i},$$

которая называется *вероятностью больших уклонений для испытаний Бернулли.*

Задача 2.4. Используя результат **задачи 2.2**, докажите следующую теорему, которая называется *теоремой о больших уклонениях* (ТБУ) для испытаний Бернулли:

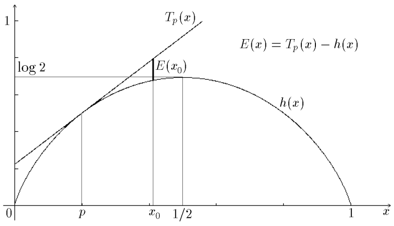
ТБУ. *Если* $0 < p < x < 1$, *то существует*

$$\lim_{N\to\infty} \frac{-\log Q_N(p,x)}{N} = T_p(x) - h(x),$$

где функция

$$T_p(x) = -x\log p - (1-x)\log(1-p), \quad 0 < x < 1,$$

описывает уравнение касательной к выпуклой вверх функции $h(x)$ *в точке* $x = p$.



Как следствие **ТБУ** получаем свойство экспоненциальной скорости сходимости к нулю условной вероятности ошибки $\mathcal{P}_m(X)$ в **задаче 2.1**, а именно: существует

$$\lim_{N\to\infty} \frac{-\log \mathcal{P}_m(X)}{N} = T_p(1/2) - h(1/2) = \frac{1}{2}\log\frac{1}{4p(1-p)}.$$

2.4.2 Случай КРШ, M=2, $1 \le t \le N/2$

Пусть $X = \Big(\mathbf{x}(1), \mathbf{x}(2)\Big)$ — код с двумя словами с расстоянием $\rho(\mathbf{x}(1), \mathbf{x}(2)) = s$, где $1 \le s \le N$.

Задача 2.5. Докажите, что для КРШ при МП-декодировании условная вероятность ошибки

$$\mathcal{P}_m^{(t)}(X) = \frac{C_s^{s/2} C_{N-s}^{t-s/2}}{C_N^t}, \quad m = \overline{1,2}, \quad C_a^b := \begin{cases} \frac{a!}{b!(a-b)!}, & 0 \le b \le a - \;, \\ 0, & \;, \end{cases}$$

а условная вероятность ошибки при МР-декодировании

$$\mathcal{P}_m(X) = \begin{cases} \frac{\sum_{i \ge s/2}^{\min(t;s)} C_s^i C_{N-s}^{t-i}}{C_N^t}, & t \ge \frac{s}{2}, \\ 0, & t < \frac{s}{2}, \end{cases} \qquad m = \overline{1,2}.$$

Указание. Не нарушая общности можно считать, что

$$\mathbf{x}(1) = (0, 0, \ldots, 0); \quad \mathbf{x}(2) = (\underbrace{1, 1, \ldots, 1}_{s}, \underbrace{0, ..., 0}_{N-s}).$$

2.5 Теорема Шеннона для КРШ

Легко понять, что для канала с равновесным шумом (КРШ), задаваемого параметром t, $1 \le t < N$, формулы (1) — (3), описывающие вероятность ошибки $\mathcal{P}^{(t)}(X)$, для мягкого МП-декодирования могут быть записаны в виде

$$\mathcal{P}^{(t)}(X) = \frac{1}{M} \sum_{m=1}^{M} \mathcal{P}_m^{(t)}(X), \tag{1}$$

$$\mathcal{P}_m^{(t)}(X) = \sum_{\mathbf{y}} P_N^{(t)}\Big(\mathbf{y}|\mathbf{x}(m)\Big) \chi_m^{(t)}(X, \mathbf{y}), \tag{2}$$

$$\chi_m^{(t)}(X, \mathbf{y}) = \begin{cases} 1, & m' \ne m \;\text{р}\; \rho\Big(\mathbf{x}(m'), \mathbf{y}\Big) = t, \\ 0, & m' \ne m \;\text{р}\; \rho\Big(\mathbf{x}(m'), \mathbf{y}\Big) \ne t. \end{cases} \tag{3}$$

$$P_N^{(t)}(\mathbf{y}|\mathbf{x}) = \begin{cases} (C_N^t)^{-1}, & \rho(\mathbf{x}, \mathbf{y}) = t, \\ 0, & \rho(\mathbf{x}, \mathbf{y}) \ne t. \end{cases} \tag{4}$$

Пусть $R > 0$, $0 < p < 1/2$, — фиксированные параметры, такие, что при заданной длине кода N

$$M := \Big\lceil \exp\{RN\} \Big\rceil, \quad t := \lfloor pN \rfloor. \tag{5}$$

Введем
$$\mathcal{E}_N(p,R) = \min_X \mathcal{P}^{(t)}(X), \tag{6}$$
где минимум вероятности ошибки $\mathcal{P}^{(t)}(X)$ берется по всевозможным (N,R)-кодам X длины N. Число $\mathcal{E}_N(p,R)$ называется *вероятностью ошибки оптимального (наилучшего)* (N,R)*-кода*, когда шум в КРШ определяется параметром p, $0 < p < 1/2$.

Положим
$$C = \log 2 - h(p). \tag{7}$$
Заметим, что число $C > 0$ при $0 < p < 1/2$. Верхнюю оценку величины $\mathcal{E}_N(p,R)$ устанавливает

Теорема Шеннона. *Для любого* $R > 0$, $0 < p < 1/2$
$$\mathcal{E}_N(p,R) \le \exp\Big\{-N(C-R)\Big\}. \tag{8}$$

Из оценки (8) вытекает

Следствие. (Прямая теорема Шеннона). *Если* $0 < R < C$, *то существует*
$$\lim_{N\to\infty} \mathcal{E}_N(p,R) = 0.$$

Замечание. Далее (см. **§5**) будет доказана обратная теорема Шеннона, которая утверждает, что
$$\underline{\lim}_{N\to\infty}\mathcal{E}_N(p,R) \ge 1 - \frac{C}{R}\ ,$$
а потому при $R > C$ величина $\mathcal{E}_N(p,R) \not\to 0$ при $N \to \infty$. Прямая и обратная теоремы Шеннона оправдывают название числа C — *пропускной способностью* КРШ, удовлетворяющего (5).

Доказательство теоремы Шеннона. Будем использовать метод доказательства теорем существования, который придумал Шеннон и назвал методом *случайного кодирования*. Рассмотрим *ансамбль*, состоящий из всех 2^{NM} кодов X длины N и объема M. Введем равномерное распределение вероятностей $\mathbf{Q}(X)$ на этом ансамбле, т.е. для любого X определим
$$\mathbf{Q}(X) := 2^{-NM}, \quad X = \Big(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(M)\Big).$$

При этом кодовые слова $\mathbf{x}(m)$ будут независимыми случайными величинами с одним и тем же распределением
$$Q_N(\mathbf{x}) := 2^{-N}, \quad \mathbf{x} = (x_1, x_2, \dots, x_N) \in (0,1)^N$$

$$\mathbf{Q}(X) = \prod_{m=1}^{M} Q_N\Big(\mathbf{x}(m)\Big). \tag{9}$$

Во введенном ансамбле кодов условные вероятности ошибки $\mathcal{P}_m^{(t)}(X)$, $m = \overline{1, M}$, и вероятности ошибки $\mathcal{P}^{(t)}(X)$ являются *случайными величинами* (функциями, определенными на элементах X ансамбля). Обозначим чертой сверху $\overline{\mathcal{P}^{(t)}(X)}$, $\overline{\mathcal{P}_m^{(t)}(X)}$, $m = \overline{1, M}$, усреднение (математическое ожидание) по ансамблю. Из (1) следует, что

$$\overline{\mathcal{P}^{(t)}(X)} = \frac{1}{M} \sum_{m=1}^{M} \overline{\mathcal{P}_m^{(t)}(X)}. \tag{10}$$

Имеет место

Лемма. *Для любого* $m = \overline{1, M}$

$$\overline{\mathcal{P}_m^{t}(X)} = 1 - \left(1 - \frac{C_N^t}{2^N}\right)^{M-1}. \tag{11}$$

Прежде, чем доказывать лемму, покажем, как с ее помощью устанавливается граница (8) теоремы Шеннона. Из (10) и (11) следует

$$\overline{\mathcal{P}^{(t)}(X)} = 1 - \left(1 - \frac{C_N^t}{2^N}\right)^{M-1} \le (M-1)\frac{C_N^t}{2^N}, \tag{12}$$

где последнее неравенство вытекает из известного неравенства Бернулли $(1+x)^n \ge 1 + nx$, если $x \ge -1$. Неравенство (12) означает существование кода X объема M, длины N, вероятность ошибки которого

$$\mathcal{P}^{(t)}(X) \le (M-1)\frac{C_N^t}{2^N}.$$

Поэтому, согласно определению (6), оптимальная вероятность ошибки

$$\mathcal{E}_N(p, R) \le (M-1)\frac{C_N^t}{2^N}, \tag{13}$$

где $M = \Big\lceil \exp\{NR\} \Big\rceil$, $t = \lfloor pN \rfloor$. В **задаче 2.2** показано, что

$$C_N^{\lfloor pN \rfloor} \le \exp\Big\{Nh(p)\Big\}. \tag{14}$$

Поскольку $M - 1 \le \exp\{NR\}$, то из (13) и (14) вытекает доказываемое неравенство (8).

Для завершения вывода теоремы Шеннона приведем

Доказательство леммы. Пусть X_m обозначает подкод кода X, состоящий из всех слов X, за исключением $\mathbf{x}(m)$. Число слов X_m равно $M-1$. Согласно определению (2), имеем

$$\overline{\mathcal{P}_m^{(t)}(X)} = \sum_{\mathbf{y}} \sum_X \mathbf{Q}(X) P_N^{(t)}\Big(\mathbf{y}|\mathbf{x}(m)\Big) \chi_m^{(t)}(X,\mathbf{y}) =$$
$$= \sum_{\mathbf{x}(m)} \sum_{\mathbf{y}} Q_N\Big(\mathbf{x}(m)\Big) P_N^{(t)}\Big(\mathbf{y}|\mathbf{x}(m)\Big) \sum_{X_m} \prod_{m'\neq m} Q_N\Big(\mathbf{x}(m')\Big) \chi_m^{(t)}(X,\mathbf{y}) \tag{15}$$

Для фиксированного $\mathbf{y} \in (0,1)^N$ введем события

$$\mathcal{A}_{\mathbf{y}}^{(t)}(m') = \Big\{\mathbf{x}(m') \ : \ \rho\Big(\mathbf{x}(m'),\mathbf{y}\Big) = t\Big\}, \tag{16}$$

где $m' = \overline{1,M}$, $m' \neq m$. Поскольку характеристическая функция (3) принимает всего два значения 0 и 1, то ее усреднение

$$\sum_{X_m} \prod_{m'\neq m} Q_N\Big(\mathbf{x}(m')\Big) \chi_m^{(t)}(X,\mathbf{y}) = \mathbf{Q}\Big\{ \bigcup_{m'\neq m} \mathcal{A}_{\mathbf{y}}^{(t)}(m')\Big\}, \tag{17}$$

где воспользовались тем, что характеристическая функция $\chi_m^{(t)}(X,y)$ является, по определению, индикатором события $\bigcup_{m'\neq m} \mathcal{A}_y^{(t)}(m')$. Символ $\mathbf{Q}\{.\}$ используется для обозначения вероятности события в ансамбле.

В силу независимости кодовых слов $\mathbf{x}(m')$, события (16) являются взаимно независимыми и, кроме того, для любого $m' \neq m$ вероятность

$$\mathbf{Q}\Big\{\mathcal{A}_{\mathbf{y}}^{(t)}(m')\Big\} = \frac{C_N^t}{2^N}$$

для любого $\mathbf{y} = (y_1, \ldots, y_N) \in (0,1)^N$. Поэтому для вероятности объединения $M-1$ независимых событий имеем

$$\mathbf{Q}\Big\{ \bigcup_{m'\neq m} \mathcal{A}_{\mathbf{y}}^{(t)}(m')\Big\} = 1 - \left(1 - \frac{C_N^t}{2^N}\right)^{M-1} \tag{18}$$

для любого $\mathbf{y} \in (0,1)^N$. Утверждение леммы, т.е. равенство (11), является следствием (15), (17) и (18).

Лемма доказана.

Задача 2.6. Рассмотрим модификацию КРШ, когда при передаче длины N переходная вероятность

$$P_N(\mathbf{y}|\mathbf{x}) = \begin{cases} 1/\sum_{i=0}^t C_N^i, & \rho(\mathbf{x};\mathbf{y}) \leq t, \\ 0, & \rho(\mathbf{x};\mathbf{y}) > t. \end{cases} \tag{19}$$

где t, $1 \le t < N/2$, — целое число. Докажите, что для такого канала среднее по ансамблю кодов объема M и длины N вероятность ошибки

$$\overline{\mathcal{P}(X)} = 1 - \left(1 - \frac{\sum_{i=0}^{t} C_N^i}{2^N}\right)^{M-1} \le (M-1)\sum_{i=0}^{t} C_N^i 2^{-N}$$

С помощью этого неравенства и результата **задачи 2.4** докажите, что для канала (19) справедлива теорема Шеннона, т.е. имеет место неравенство (8) и асимптотическое следствие из него.