

Лекция 2. Терминология. Ошибки и их причины. Корректирующая способность избыточного кода. Критерии эффективности помехоустойчивых кодов.

Общее представление вероятности ошибки в системах связи

Основные причины возникновения ошибок:

1. Последствия фильтрации сигналов в передатчике, канале и приемнике – неидеальная фильтрация приводит к появлению *межсимвольной интерференции* (ISI – intersymbol interference).
2. Электромагнитные излучения различных других источников.
3. Электрические помехи: *тепловой шум*, который аддитивно накладывается на передаваемый сигнал (аддитивный шум), *импульсные помехи*, *кратковременные перерывы* (замирания), которые представляют собой мультипликативные помехи.

Природа шумов – *искусственные и естественные* шумы.

Основная статистическая характеристика *аддитивного шума* $n(t)$ – его амплитуды x распределены по нормальному или гауссову закону распределения с нулевым средним a и дисперсией σ^2 :

$$p(x) := \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-a)^2} \quad (2.1)$$

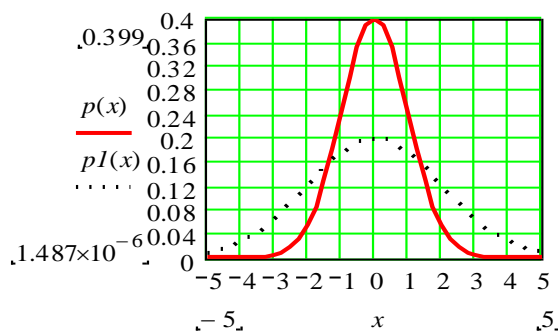


Рис. 2.1. Нормальный закон распределения амплитудных значений гауссова шума при $a=0$ и $\sigma=1$ (сплошная линия) и $\sigma=2$ (пунктирная линия).

Правило 3σ .

Основной *спектральной характеристикой* аддитивного шума является спектральная плотность мощности $G_n(f)$, которая будет равномерной для всего частотного диапазона от $-\infty$ до $+\infty$ и записываться в виде

$$G_n(f) = \frac{N_0}{2} \text{ Вт / Гц}. \quad (2.2)$$

Здесь коэффициент 2 в делителе означает, что функция $G_n(f)$ является двусторонней спектральной плотностью мощности (как в положительной, так и в отрицательной области частот). Из (2.2) следует, что спектральная плотность мощности гауссова шума одинакова на любой частоте (рис. 2.2,

а), такой шум принято называть *белым*, по аналогии с белым светом, содержащего равные доли всех частот видимого диапазона электромагнитного излучения.

Обратное преобразование Фурье спектральной плотности мощности белого шума определяет его автокорреляционную функцию, равную

$$R_n(\tau) = F^{-1} \{ G_n(f) \} = \int_{-\infty}^{\infty} G_n(f) e^{2\pi i f \tau} df = \int_{-\infty}^{\infty} \frac{N_0}{2} e^{2\pi i f \tau} df = \frac{N_0}{2} \delta(\tau) \quad (2.3)$$

и имеющую вид дельта-функции Дирака, взвешенной множителем $N_0/2$ и находящейся в точке $\tau=0$ (рис. 2.2, б). Равенство автокорреляционной функции $R_n(\tau)$ нулю для всех $\tau \neq 0$ говорит о том, что различные выборки белого шума во времени не коррелируют между собой, т.е. являются независимыми. По этой причине канал с *аддитивным белым гауссовым шумом* (АБГШ) называется *каналом без памяти*.

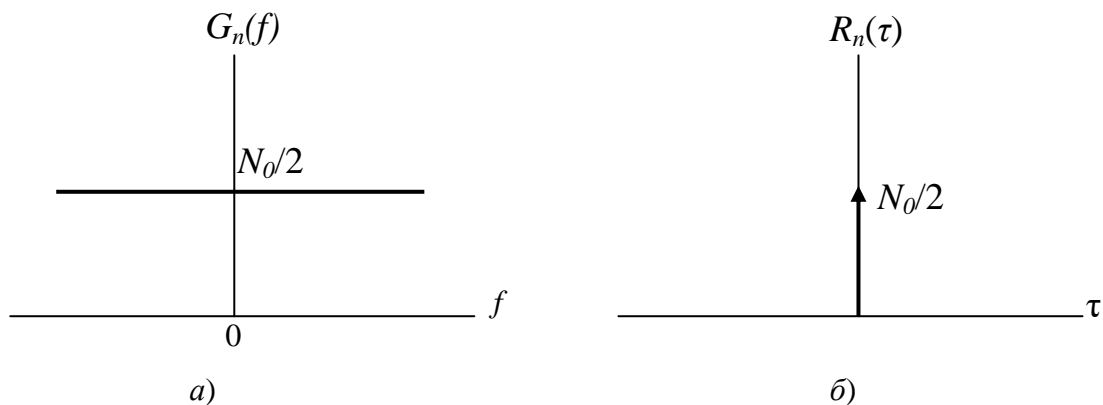


Рис. 2.2. Спектральная плотность мощности (а) и автокорреляционная функция (б) белого шума.

Заметим, что односторонняя спектральная плотность мощности считается равной N_0 [2].

Значение средней мощности шума в канале АБГШ конечна и, как доказано в [1], равна $N_0/2$ на выходе коррелятора или согласованного фильтра.

Вероятность ошибки и отношение сигнал/шум.

Общепринятым критерием качества *аналоговой связи* принято считать отношение сигнал/шум, представляющее собой отношение средней мощности сигнала к средней мощности шума S/N . В *цифровой связи* в качестве критерия качества связи как правило используется нормированное отношение сигнал/шум как E_b/N_0 , где E_b – это энергия битового сигнала и её можно определить как мощность сигнала S , умноженную на время передачи битового сигнала T_b , N_0 – спектральная плотность мощности шума, которую можно выразить как мощность шума N , деленную на ширину полосы W . Учитывая, что битовая скорость передачи R_b равна $R_b = 1/T_b$, получаем:

$$\frac{E_b}{N_0} = \frac{S T_b}{N / W} = \frac{S / R_b}{N / W} = \frac{S}{N} \left(\frac{W}{R_b} \right). \quad (2.4)$$

В дальнейшем в нашем курсе речь будет идти о цифровой связи, поэтому в качестве основного критерия связи естественно выбрано отношение E_b/N_0 .

Одной из важнейших характеристик цифровой системы связи является зависимость вероятности битовой ошибки P_b от E_b/N_0 , имеющая характерный вид, представленный на рис. 2.3.

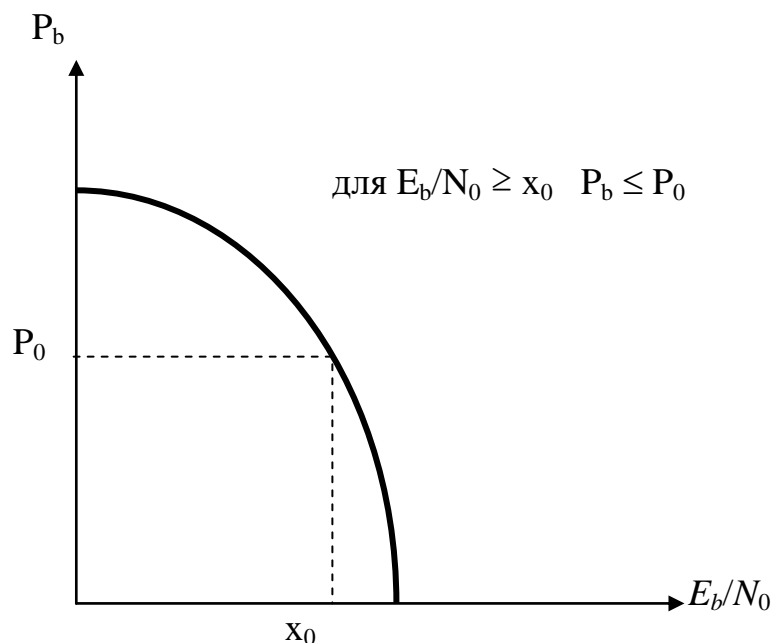


Рис. 2.3. Общий вид зависимости битовой вероятности ошибки от E_b/N_0 .

Таким образом, отношение E_b/N_0 может рассматриваться как метрика, позволяющая сравнивать различные цифровые системы связи.

Как отмечено выше, другой причиной возникновения ошибок является *межимпульсная (межсимвольная) интерференция* (intersymbol interference – ISI), которая непосредственно зависит от импульсной характеристики линейной системы, в первую очередь её от полосы пропускания.



Рис. 2.4. Импульсная характеристика линейной системы $h(t)$

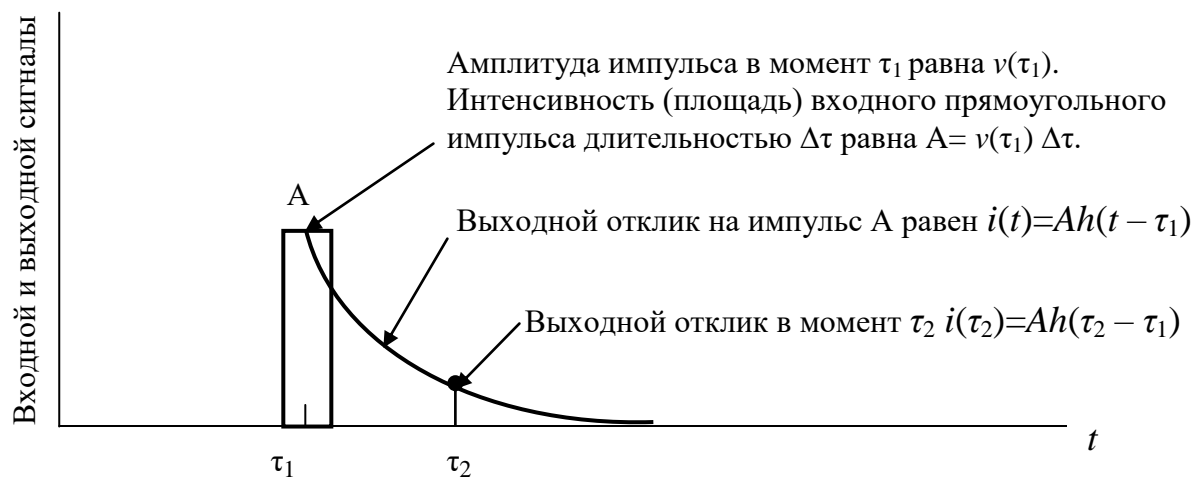


Рис. 2.5. Реакция линейной системы на прямоугольный импульс в момент времени τ_1

На рис.2.6 показаны два примера прохождения прямоугольного импульса $x(t)$ с амплитудой V_m через линейную систему (фильтр) с ограниченной полосой пропускания и выходным сигналом $y(t)$.

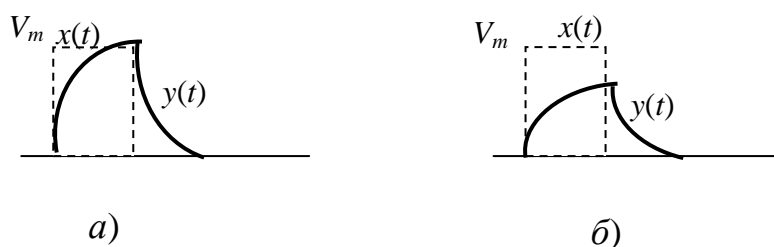


Рис. 2.6. Примеры прохождения прямоугольного импульса длительности T с полосой $W_u = 1/T$ через фильтр с полосой пропускания W_f

а) «хорошая точность воспроизведения», $W_u \approx W_f$;

б) «плохое воспроизведение», $W_u \ll W_f$.

На рис. 2.7 показан пример межимпульсной интерференции в тактовый момент времени τ_3 . Очевидно, что в момент τ_3 отклик линейной системы на предыдущие импульсы превышает порог и приводит к ошибке.

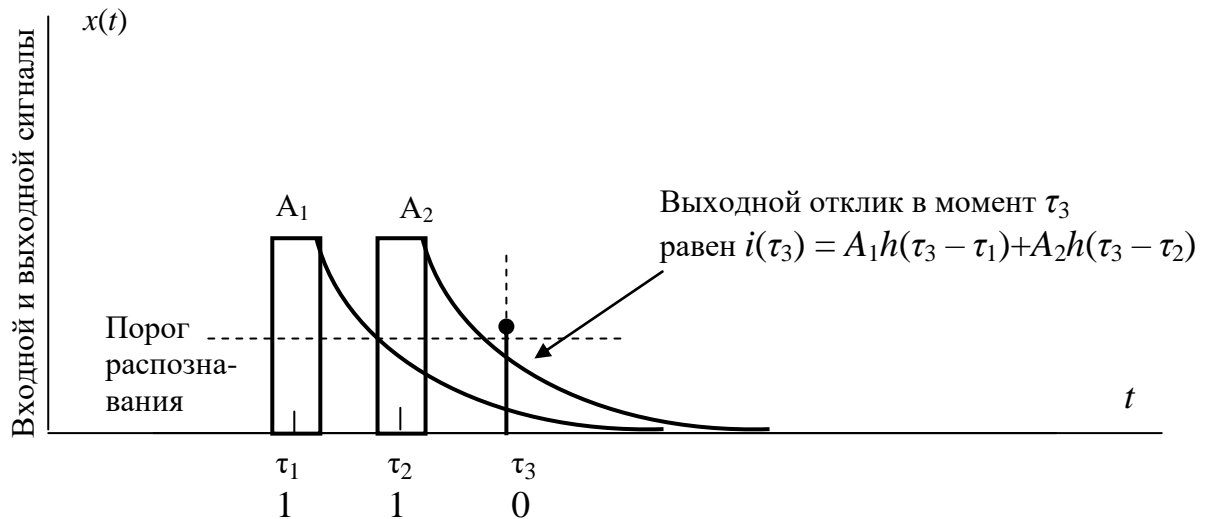


Рис. 2.7. Реакция (отклик) линейной системы на два прямоугольных импульса в момент времени τ_3 , где передается «0».

При N импульсах на входе линейной системы отклик на её выходе в последующий момент времени t будет определяться выражением

$$i(t) = A_1 h(t - \tau_1) + A_2 h(t - \tau_2) + \dots + A_N h(t - \tau_N), \quad (2.6)$$

где $\tau_1, \tau_2, \dots, \tau_N$ – моменты поступления импульсов на вход.

Для устранения влияний межимпульсной интерференции применяют специальные устройства, называемые *эквалайзерами* [1].

Другой широко используемый в настоящих системах связи метод предотвращения межимпульсной интерференции (ISI) заключается в выборе специальной формы принимаемого импульса. Над этой проблемой долгое время занимался Найквист. Он показал, что минимальная теоретическая ширина полосы частот, требуемая для передачи со скоростью $R=1/T$ символов/секунду без ISI, должна быть равна $\Delta F=1/2T$. При этом, форма импульса, не вызывающего ISI, должна описываться функцией $\frac{\sin x}{x}$. Такой

импульс называют [1] идеальным импульсом Найквиста, форма которого показана на рис.2.8. Как видно из рисунка, передача разных импульсов не создает межимпульсной интерференции в моменты t взятия выборок, $t=\pm 1, \pm 2, \pm 3, \dots$

Разумеется, что это будет иметь место в предположении идеальной синхронизации моментов взятия выборок (отсчетов) сигналов.

На практике сегодня такие сигналы применяют при ортогональном многочастотном уплотнении OFDM.

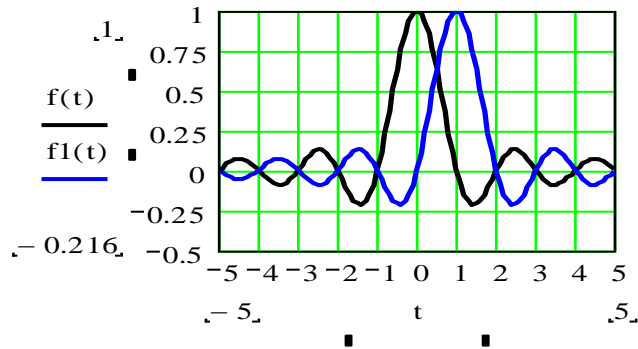


Рис. 2.8. Форма импульсов Найквиста

Вероятность ошибки при передаче двоичных сигналов в канале с гауссовым шумом (АБГШ)

Двоичный сигнал, переданный в течение временного интервала $(0, T)$, будем представлять в следующем виде:

$$s_i(t) = \begin{cases} s_1(t) & 0 \leq t \leq T \text{ для символа } 1 \\ s_2(t) & 0 \leq t \leq T \text{ для символа } 0. \end{cases} \quad (2.7)$$

Для упрощения анализа будем рассматривать влияние в канале с АБГШ только гауссова шума, обозначаемого как $n(t)$. Тогда сигнал $r(t)$ на выходе канала можно представить в виде

$$r(t) = s_i(t) + n(t) \quad i=1,2, \quad 0 \leq t \leq T. \quad (2.8)$$

Принятый сигнал в приемнике проходит демодуляцию и дискретизацию в соответствии с тактовой синхронизацией. При этом в конце интервала T на выходе устройства дискретизации (или согласованного фильтра, или коррелятора) будет формироваться выборка

$$z(T) = a_i(T) + n_0(T), \quad i=1,2. \quad (2.9)$$

Часто для упрощения записи в выражении (1.9) опускают длительность интервала T и пользуются записью $z = a_i + n_0$, где шумовая компонента n_0 является гауссовой с нулевым средним, а компонента a_i является средним значением функции $z(T)$, а именно a_1 или a_2 в зависимости от входного двоичного символа $s_1(t)$ («1») или $s_2(t)$ («0»).

Тогда, с учетом (2.1), плотности условных вероятностей будут определяться выражениями:

$$\begin{aligned} p(z|s_1) &= \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z - a_1}{\sigma_0} \right)^2 \right]; \\ p(z|s_2) &= \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z - a_2}{\sigma_0} \right)^2 \right]. \end{aligned} \quad (2.10)$$

Эти плотности вероятностей показаны на рис.2.9, где сплошной линией отображена плотность условной вероятности $p(z, s_1) = p(z|s_1)$ при $s_1 = a_1 = 2$, что соответствует логической «1», а пунктирной линией отображена плотность условной вероятности $p(z, s_2) = p(z|s_2)$ при $s_2 = a_2 = -2$, что соответствует

логическому «0». Плотность $p(z/s_1)$ справа от оси ординат называют правдоподобием s_1 . Напомним, что эта кривая есть плотность вероятности случайной переменной z при условии передачи символа s_1 . Аналогично, плотность $p(z/s_2)$ слева от оси ординат называют правдоподобием s_2 . Эта пунктирная кривая есть плотность вероятности случайной переменной z при условии передачи символа s_2 .

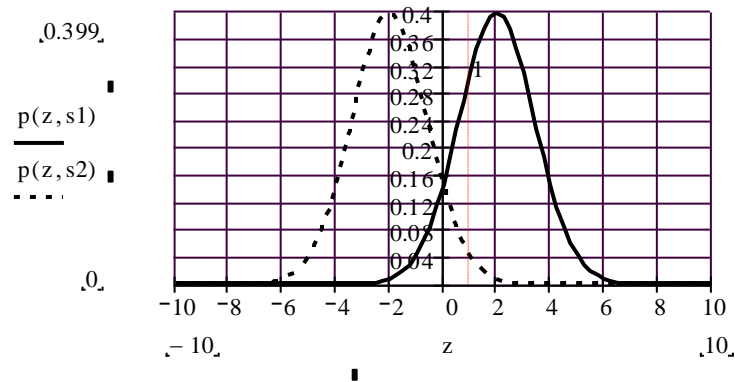


Рис. 2.9. Плотности условных вероятностей передачи двоичных сигналов в канале с АБГШ.

Жесткое и мягкое принятие решений.

На втором этапе после модуляции и дискретизации на приемной стороне принимается решение о переданном сигнале. Решение может быть *жестким* или *мягким*.

Суть жестких решений состоит в сравнении значений принятого сигнала $z(T)$ в момент времени T с некоторыми оптимальными порогами. Для двоичных сигналов $s_1(t)=a_1$ и $s_2(t)=a_2$ в канале с АБГШ порог будет один и его оптимальное значение будет равно

$$\gamma = \frac{a_1 + a_2}{2}. \quad (2.11)$$

В случае противоположных сигналов с одинаковыми абсолютными значениями, $a_2 = -a_1$, оптимальный порог, как видно из рис. 2.10, равен $\gamma=0$, т.е. совпадает с осью ординат. Следовательно, если сигнал $z(T)$ находится правее порога γ , т.е. попадает в зону правдоподобия s_1 , то будет принято решение о переданном символе «1». Если же сигнал $z(T)$ попадает в зону правдоподобия s_2 , т.е. левее порога (оси ординат), то будет выдано решение о приеме символа «0». Таким образом, при передаче двоичных сигналов и жестких решениях формируется решение о приеме конкретного двоичного символа на основе двухуровневой схемы (рис. 2.10).

Рисунки 2.9 и 2.10 демонстрируют возможные ошибки. Так, площадь под кривой $p(z/s_1)$ левее порога (оси ординат) определяет вероятность ошибки при передаче символа «1». Аналогично, площадь под кривой $p(z/s_2)$ правее оси ординат определяет вероятность ошибки при передаче символа «0».

Таким образом, алгоритм жесткого принятия решений при передаче двоичных сигналов по каналу с АБГШ может быть записан в виде:

$$z(T) \begin{cases} > \gamma \rightarrow H_1 \\ < \gamma \rightarrow H_2 \end{cases}, \quad (2.12)$$

где H_1 и H_2 – две возможные гипотезы.

Из приведенной записи следует, если $z(T) > \gamma$, то выбирается гипотеза H_1 , которая говорит о том, что был передан сигнал $s_1(t)$, т.е. символ «1». В противном случае при $z(T) < \gamma$ принимается гипотеза H_2 о том, что был передан «0».

Вероятность ошибки e при передаче сигнала $s_1(t)=a_1$ может возникнуть с вероятностью

$$P(e|s_1) = P(H_2|s_1) = \int_{-\infty}^{\gamma} p(z|s_1) dz. \quad (2.13)$$

В случае передачи сигнала $s_2(t)=a_2$ вероятность возникновения ошибки аналогично будет равна

$$P(e|s_2) = P(H_1|s_2) = \int_{\gamma}^{\infty} p(z|s_2) dz. \quad (2.14)$$

Суммарная битовая вероятность ошибки при передаче двоичных сигналов в канале с АБГШ определяется выражением:

$$P_b = \sum_{i=1}^2 P(e, s_i) = \sum_{i=1}^2 P(e|s_i) P(s_i). \quad (2.15)$$

Учитывая (2.13) и (2.14) вероятность битовой ошибки (2.15) можно записать в виде

$$P_b = P(e|s_1)P(s_1) + P(e|s_2)P(s_2) = P(H_2|s_1)P(s_1) + P(H_1|s_2)P(s_2). \quad (2.16)$$

Для равновероятных двоичных сигналов суммарная вероятность ошибки равна

$$P_b = \frac{1}{2} P(H_2|s_1) + \frac{1}{2} P(H_1|s_2). \quad (2.17)$$

Учитывая симметричность плотностей условных вероятностей, получаем

$$P_b = P(H_1|s_2) = P(H_2|s_1). \quad (2.18)$$

Тогда вероятность битовой ошибки можно определить как следующую интегральную функцию, например, от плотности вероятности $p(z/s_2)$

$$P_b = \int_{\gamma=\frac{a_1+a_2}{2}}^{\infty} p(z|s_2) dz. \quad (2.19)$$

Заменяя плотность вероятности $p(z/s_2)$ формулой (1.10), получим

$$P_b = \int_{\gamma=\frac{a_1+a_2}{2}}^{\infty} \frac{1}{\sigma_0 \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{z-a_2}{\sigma_0} \right)^2 \right] dz. \quad (2.20)$$

Проведя замену переменной $\frac{z-a_2}{2} = u$, получим

$$P_b = \int_{u=\frac{a_1-a_2}{2\sigma_0}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = Q\left(\frac{a_1-a_2}{2\sigma_0}\right), \quad (2.21)$$

где $Q(x)$ – гауссов интеграл ошибки.

Тенденция развития цифровых систем связи свидетельствует, что в большинстве современных систем стали применять для принятия решений более двух уровней квантования. В этих случаях схемы принятия решений называют *мягкими*. На рис. 2.10 показан пример мягких решений с восьмиуровневой схемой квантования. Уровни квантования условно пронумерованы от нуля до семи в двоичном коде. Если уровень сигнала $z(T)$ попал в зону трех единиц {111}, это свидетельствует о том, что с большой достоверностью переданным символом была «1», а выделение уровня {100} свидетельствует, что была передана «1», но с низкой достоверностью. Для более простой и понятной интерпретации мягких решений уровни в данном примере могут быть обозначены [1] как $-7, -5, -3, -1, 1, 3, 5, 7$. Чем больше цифра уровня, тем больше достоверность принятия мягких решений.

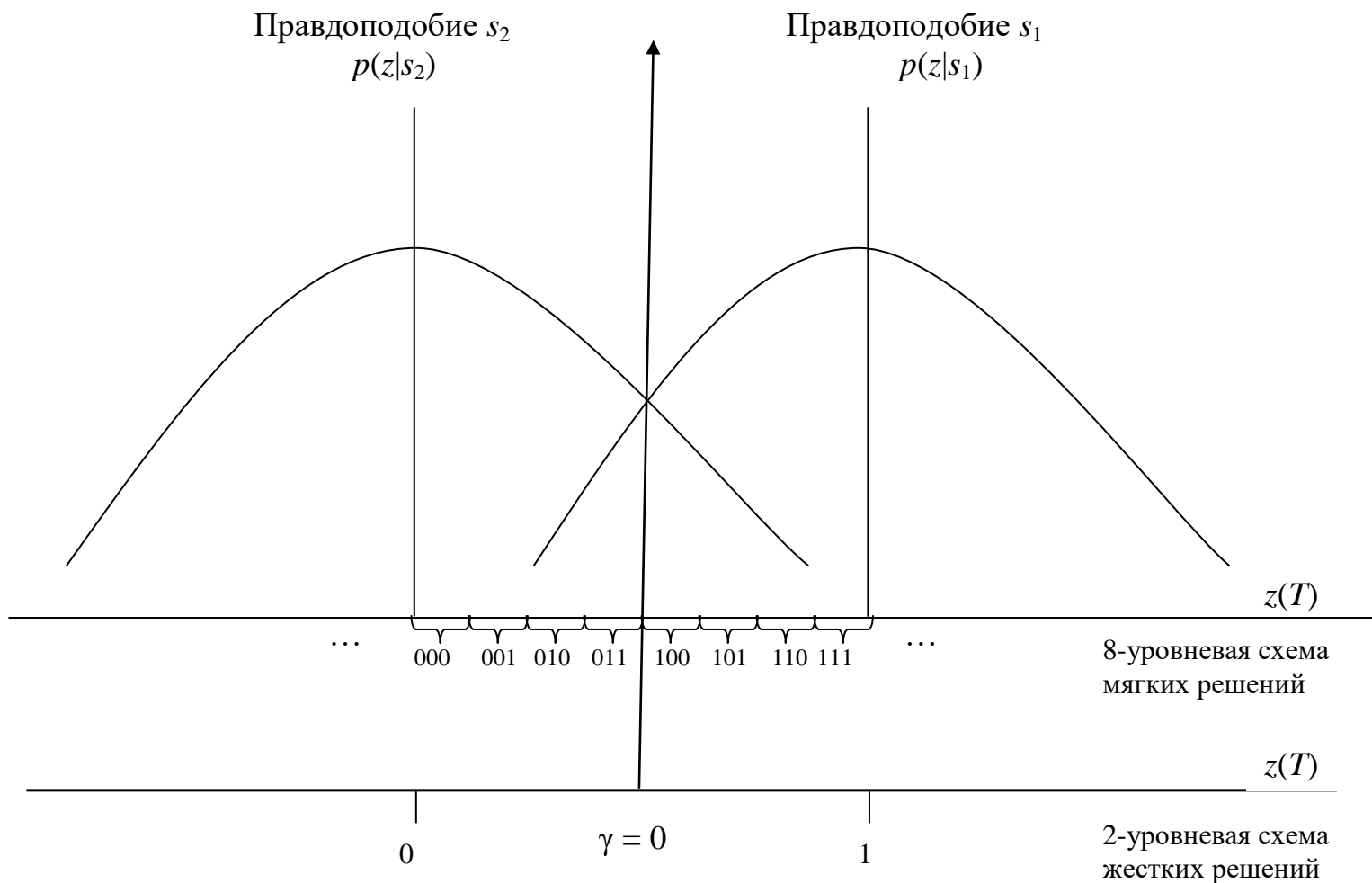


Рис. 2.10. Жесткая и мягкая схемы принятия решений

Знак у цифр характеризует тот или иной двоичный сигнал, например, положительный знак соответствует сигналу s_1 («1»), а отрицательный – сигналу s_2 («0»).

Возможна и другая интерпретация мягких решений для двоичных сигналов в канале с АБГШ. Например, указанные выше цифры в абсолютных значениях более целесообразно соотнести с вероятностями попадания в тот или иной уровень, а именно: $-7/8, -5/8, -3/8, -1/8, 1/8, 3/8, 5/8, 7/8$. То есть зоны, ближайšie к математическим ожиданиям сигналов s_1 и s_2 , имеют наибольшую (близкую к 1) вероятность.

Как утверждается в [1], переход на восьмиуровневую схему мягких решений по сравнению с двухуровневой схемой жестких решений дает выигрыш отношения сигнал/шум в 2 дБ. Это означает, что восьмиуровневая схема мягких решений обеспечивает такую же вероятность битовой ошибки как и двухуровневая схема жестких решений, но при отношении сигнал/шум E_b/N_0 на 2 дБ ниже. Однако за этот выигрыш от применения восьмиуровневой схемы мягких решений приходится платить повышением быстродействия процессора как минимум в 3 раза.

Оптимизация вероятности ошибки.

Смысл оптимизации вероятности ошибки заключается в её минимизации применительно к используемому каналу связи.

Как было показано первым шагом оптимизации вероятности ошибки для двоичных сигналов являлся выбор оптимального порога γ в соответствии с выражением (2.11). Для канала с АБГШ вероятность ошибки двоичных сигналов P_b определяется выражением (2.21), из которого следует, что для её минимизации необходимо обеспечить максимальное значение аргумента в функции $Q(x)$ в выражении (2.21). Очевидно, что этому будет соответствовать

максимум величины $\frac{(a_1 - a_2)^2}{\sigma_0^2}$, где числитель представляет энергию разностного

сигнала на входе согласованного фильтра в момент времени $t=T$, а $(\sigma_0)^2$ – двусторонняя спектральная плотность мощности гауссова шума, равная $N_0/2$. Известно, что если фильтр согласовывает разностный сигнал $[s_1(t) - s_2(t)]$, то соотношение сигнал/шум на его выходе в момент $t=T$ можно записать как

$$\left(\frac{S}{N}\right)_T = \frac{(a_1 - a_2)^2}{\sigma_0^2} = \frac{2E_d}{N_0}, \quad (2.22)$$

где

$$E_d = \int_0^T [s_1(t) - s_2(t)]^2 dt \quad (2.23)$$

энергия разностного сигнала.

Найдя из (2.22) $\frac{(a_1 - a_2)}{2\sigma_0} = \sqrt{\frac{E_d}{2N_0}}$ и подставив это значение в (2.21), получим вероятность битовой ошибки

$$P_b = Q\left(\sqrt{\frac{E_d}{2N_0}}\right). \quad (2.24)$$

В случае равных энергий обоих сигналов можно получить общее выражение битовой вероятности ошибки двоичного сигнала, учитывающее корреляцию сигналов $s_1(t)$ и $s_2(t)$. Действительно, интеграл (2.23) можно переписать в виде

$$E_d = \int_0^T s_1^2(t)dt + \int_0^T s_2^2(t)dt - 2 \int_0^T s_1(t)s_2(t)dt = 2E_b(1 - \rho), \quad (2.25)$$

где E_b – энергия битового сигнала

$$E_b = \int_0^T s_1^2(t)dt = \int_0^T s_2^2(t)dt,$$

а ρ – коэффициент взаимной корреляции, равный

$$\rho = \frac{1}{E_b} \int_0^T s_1(t)s_2(t)dt. \quad (2.26)$$

Подставив (2.25) в (2.24), получим

$$P_b = Q\left(\sqrt{\frac{E_b(1 - \rho)}{N_0}}\right). \quad (2.27)$$

Учитывая, что $-1 \leq \rho \leq 1$, более удобной для анализа формой записи коэффициента корреляции является запись $\rho = \cos\theta$, где θ – угол между векторами s_1 и s_2 .

При противоположных или *антиподных* сигналах, для которых $\theta = 180^\circ$, (Рис. 2.11, а) коэффициент взаимной корреляции равен $\rho = -1$. Тогда вероятность битовой ошибки для таких сигналов, как следует из (2.27), будет равна

$$P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right). \quad (2.28)$$

Двоичные сигналы с нулевой корреляцией, у которых $\theta = 90^\circ$ (Рис. 2.11, б) и коэффициент взаимной корреляции равен $\rho = 0$, называются *ортогональными*. Вероятность битовой ошибки для таких сигналов, соответственно, будет равна

$$P_b = Q\left(\sqrt{\frac{E_b}{N_0}}\right). \quad (2.29)$$

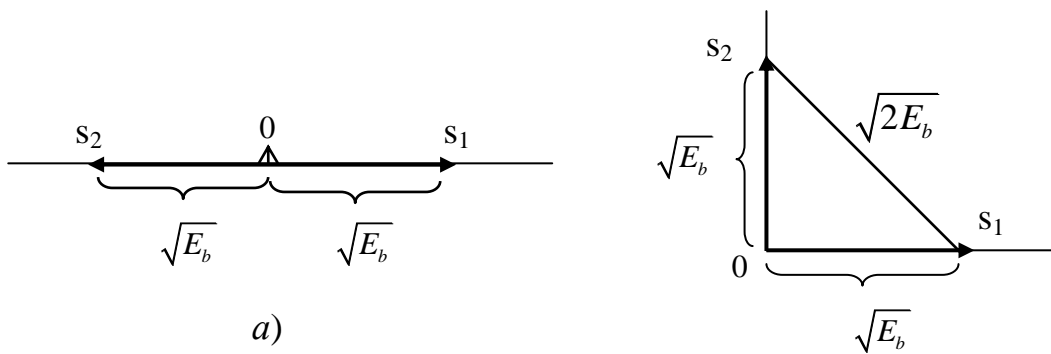


Рис. 2.11. Векторное представление двоичных сигналов s_1 и s_2 :
а) антиподных; б) ортогональных.

М-ичная передача сигналов и вероятность битовой ошибки

При М-ичной передаче сигналов количество возможных сигналов равно $M=2^k$, где показатель k определяет количество бит информации, передаваемых одним М-арным символом. Ошибочный прием одного М-ичного символа приводит к возникновению от одного до k ошибочных бит. Поэтому вероятность битовой ошибки P_b в случае М-ичной передачи зависит от вероятности символьной ошибки $P_E(M)$, которая в свою очередь зависит от режима передачи сигналов и от способов модуляции. Подробно об этом описано в [1]. К примеру, вероятности битовой ошибки и вероятности символьной ошибки для ортогональных М-ичных сигналов связаны соотношением

$$\frac{P_b}{P_E(M)} = \frac{2^{k-1}}{2^k - 1} = \frac{M}{2(M-1)}. \quad (2.30)$$

Очевидно, что при увеличении k получаем в пределе следующее

$$\lim_{k \rightarrow \infty} \frac{P_b}{P_E} = \frac{1}{2}.$$

Другим примером М-ичной передачи является многофазная передача MPSK, для которой при использовании кода Грея справедливо соотношение

$$P_b = \frac{P_E}{\log_2 M} \quad (\text{для } P_E \ll 1). \quad (2.31)$$

Вес и расстояние Хемминга между кодовыми векторами.

Число ненулевых элементов в кодовом слове U называют его *весом* и обозначают $w(U)$.

Обобщённой характеристикой кода, увязывающей избыточность (скорость) и корректирующие способности кода, является *расстояние Хемминга* между векторами U и V , которое обозначается как $d(U,V)$ и определяется количеством одноименных позиций с отличающимися друг от друга кодовыми элементами. Для двоичных кодов расстояние Хемминга между векторами U и V равно весу их поэлементной суммы по модулю 2, т.е.

$$d(U,V) = w(U+V).$$

Блочный линейный код может быть *групповым*, который среди своих кодовых комбинаций содержит также комбинацию, состоящую из одних нулевых элементов. Для такого группового кода характерно, что минимальный вес w_{\min} ненулевой кодовой комбинации равен минимальному расстоянию Хемминга d_{\min} .

Минимальное расстояние и его влияние на корректирующие способности линейного кода.

Параметром блочного помехоустойчивого кода является наименьшее значение Хеммингового расстояния d для всех сравниваемых пар кодовых комбинаций. Этот параметр называется минимальным кодовым расстоянием

Хемминга и обозначается d_{\min} . Поэтому часто равномерный блочный (n,k) -код, имеющий параметр d_{\min} , записывают как (n,k,d_{\min}) или просто (n,k,d) .

Рассмотрим как d_{\min} влияет на корректирующие способности помехоустойчивого кода.

Для большинства блочных помехоустойчивых кодов декодер принимает решение по принципу *максимального правдоподобия*, суть которого состоит в том, что декодер декодирует принятую n -элементную комбинацию в ближайшую к ней разрешенную кодовую комбинацию по расстоянию Хемминга. Другими словами, декодер определяет расстояние Хемминга между принятым вектором r и всеми возможными разрешенными кодовыми векторами U_j и выбирает наиболее правдоподобное кодовое слово U_i , для которого единственного выполняется условие, что

$$d(r, U_i) < d(r, U_j) \text{ для всех } j \neq i.$$

Если мы хотим построить помехоустойчивый код, исправляющий все сочетания из t или менее ошибочных элементов в любой принятой комбинации, то необходимо обеспечить значение минимального кодового расстояния, удовлетворяющее равенству

$$t = \left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor, \quad (2.32)$$

где $\lfloor \cdot \rfloor$ обозначает целую часть дроби.

Таким образом, для исправления всех сочетаний ошибок из t или менее ошибочных символов необходимо и достаточно, чтобы каждая разрешенная комбинация отличалась от любой другой разрешенной комбинации кода не менее, чем в $(2t+1)$ позициях. Формулу (2.32) можно довольно наглядно проиллюстрировать условными графическими представлениями, показанными на рис.2.2 для минимальных кодовых расстояний 4 и 5.

Как видно из демонстрационного рис.2.12,а, при $d_{\min}=4$ переданная разрешенная кодовая комбинация A_i вследствие однократной ошибки ($t=1$) перейдет в некоторую запрещенную комбинацию $ЗК_1$, которая по расстоянию Хемминга находится ближе всего к разрешенной комбинации A_i и поэтому декодер примет правильное решение, т.е. будет декодирована переданная комбинация. Если же в принятой комбинации возникнут 3 ошибки, которые переведут комбинацию A_i в запрещенную комбинацию $ЗК_3$, то эта комбинация окажется ближе всего к другой разрешенной комбинации A_j и декодер примет решение, что была передана именно эта комбинация, т.е. произойдет неправильное декодирование. Наконец, если возникнет двукратная ошибка, которая переведет разрешенную комбинацию A_i в запрещенную комбинацию $ЗК_2$, то декодер сформирует отказ от декодирования принятой комбинации в силу её равноудаленности и от A_i , и от A_j .

Таким образом, при $d_{\min}=4$ декодер гарантированно исправляет однократные ошибки и может обнаружить, но не исправить, двукратные ошибки. Тем самым мы показали, что в общем виде при d_{\min} чётном декодер сможет гарантированно исправить все ошибки до кратности t включительно, где значение t определяется выражением (2.16).

Анализируя рис.2.12,б, придём к заключению, что при $d_{\min}=5$ декодер, используя критерий наиболее близкого расстояния к какой-либо разрешенной комбинации, сможет гарантированно исправить все однократные и двукратные ошибки, тогда как ошибки кратности 3, 4 и 5 приведут к неправильному декодированию. Таким образом, для любого нечетного d_{\min} декодер гарантированно исправляет все ошибки до кратности $t = \frac{d_{\min}-1}{2}$ включительно.

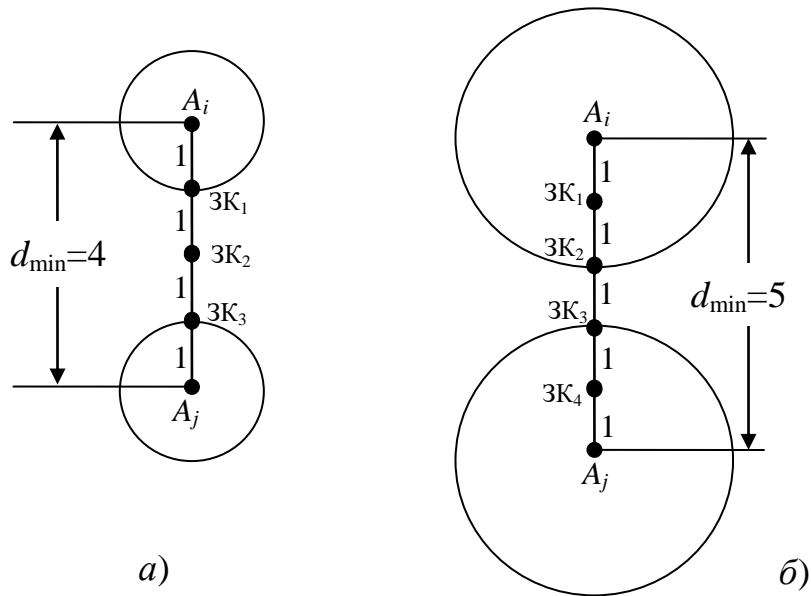


Рис.2.12. Связь минимального кодового расстояния d_{\min} и корректирующих свойств помехоустойчивого кода для $d_{\min}=4$ (а) и $d_{\min}=5$ (б) («1» на рисунке – ошибка)

Из рассмотренного выше следует вывод, что по критерию наименьшего расстояния Хемминга блочный код гарантированно исправляет все ошибки до половины минимального расстояния Хемминга (так называемого конструктивного расстояния).

Если помехоустойчивый код работает только в режиме обнаружения ошибок, то, анализируя рис. 2.12, можно заключить, что в этом режиме помехоустойчивый код гарантированно обнаруживает все ошибки до кратности σ включительно, равной

$$\sigma \leq (d_{\min} - 1). \quad (2.33)$$

Кроме того, код будет обнаруживать и определенную долю ошибок большей кратности, которые переводят разрешенную кодовую комбинацию в запрещенную. Поэтому блочный (n,k) -код в режиме обнаружения ошибок может обнаружить $2^n - 2^k$ ошибочных комбинаций. Кроме того, из-за ошибок переданная кодовая комбинация U_i может перейти в любую другую разрешенную комбинацию, количество которых равно $2^k - 1$. В этом случае синдром ошибочной комбинации будет равен 0 и, следовательно, ошибки будут *необнаруживаемые*. Вероятность необнаруживаемых ошибок зависит от весового спектра кода. Весовой спектр представляет собой распределение так называемых весовых коэффициентов $A_0, A_1, A_2, \dots, A_n$, где A_j – количество кодовых слов с весом j . Таким образом, если двоичный блочный (n,k) -код

используется только для обнаружения ошибок, то в двоичном симметричном канале вероятность поступления в декодер кодового слова с необнаруженными ошибками будет равна

$$P_{no} = \sum_{j=1}^n A_j p^j (1-p)^{n-j}, \quad (2.34)$$

где p – вероятность битовой ошибки в канале.

Кроме рассмотренных выше двух режимов блочный (n,k) -код может работать в режиме гарантированного исправления ошибок до кратности t включительно и дополнительно обнаруживать все ошибки до кратности σ включительно. При этом граничное максимальное значение σ зависит от четности минимального кодового расстояния d_{\min} . Пусть d_{\min} **нечетное**. Тогда при заданной кратности исправляемых ошибок t от 0 до $\left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$ блочный (n,k) -код может ещё дополнительно обнаруживать также все ошибки кратностью от $(t+1)$ до $[d_{\min} - (t+1)]$ включительно. Если же кратность исправляемых ошибок задана до $t = \left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$, то в этом случае код не может еще дополнительно обнаруживать ошибки большей кратности в рамках конструктивного расстояния d_{\min} .

Пример для $d_{\min}=7$ приведен в таблице 2.1.

Для **четного** d_{\min} максимальная кратность исправляемых ошибок равна $\left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$, где $\lfloor \cdot \rfloor$ – целая часть дроби. Аналогично, при заданной кратности исправляемых ошибок t блочный (n,k) -код может ещё дополнительно обнаруживать также все ошибки кратностью от $(t+1)$ до $[d_{\min} - (t+1)]$ включительно. Пример для $d_{\min}=6$ приведен в таблице 2.2.

Таблица 2.1 ($d_{\min}=7$)

Кратности исправляемых ошибок, t $0 \leq t \leq \left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$	Кратности дополнительно обнаруживаемых ошибок, σ	
	от $(t+1)$	до $[d_{\min} - (t+1)]$ включительно
0 (режим обнаружения)	1	6
1	2	5
≤ 2	3	4
≤ 3	-	-

Таблица 2.2 ($d_{\min}=6$)

Кратности исправляемых ошибок, t $0 \leq t \leq \left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$	Кратности дополнительно обнаруживаемых ошибок, σ	
	от $(t+1)$	до $[d_{\min} - (t+1)]$ включительно
0 (режим обнаружения)	1	5
1	2	4
≤ 2	3	3

Теперь покажем на примере двоичных кодов (основание кода 2), как связана избыточность блочного (n,k) -кода с кратностью исправляемых ошибок. В случае исправления всех ошибок до кратности t включительно подмножество комбинаций длины n , соответствующее некоторой разрешенной комбинации A_i , будет содержать саму разрешенную комбинацию A_i (безошибочный прием) и $(C_n^1 + C_n^2 + \dots + C_n^t)$ запрещенных комбинаций длины n , где C_n^i число сочетаний из n по i . Отсюда следует, что число N разрешенных комбинаций блочного (n,k) -кода будет ограничено значением

$$N \leq \frac{2^n}{\sum_{i=0}^t C_n^i}, \quad (2.35)$$

где C_n^i — количество запрещенных комбинаций длины n , отличающихся от A_i в i позициях. Наиболее часто встречаются блочные (n,k) -коды, у которых число разрешенных комбинаций равно $N = 2^k$. В этом случае выражение (2.35) принимает вид

$$2^k \leq \frac{2^n}{\sum_{i=0}^t C_n^i}$$

или

$$2^{n-k} \geq \sum_{i=0}^t C_n^i. \quad (2.36)$$

Логарифмируя по основанию 2 выражение (2.36), получим, что абсолютная избыточность блочного (n,k) -кода, исправляющего до t ошибок включительно, должна удовлетворять неравенству

$$n-k \geq \log_2 \sum_{i=0}^t C_n^i. \quad (2.37)$$

Выражение (2.37) представляет собой границу Хемминга для избыточного блочного (n,k) -кода при заданных значениях k и t .

Обобщение этой границы для недвоичных кодов (основание кода $p \neq 2$) имеет вид:

$$n-k \geq \log_p \left[\sum_{i=0}^t C_n^i (p-1)^i \right]. \quad (2.38)$$

Из этого следует, что абсолютная избыточность двоичного помехоустойчивого кода, исправляющего однократные ошибки ($t=1$), определяется неравенством $n - k \geq \log_2(1 + n)$.

Эффективность помехоустойчивого кодирования.

Эффективность помехоустойчивого кода, наряду с такими показателями как **скорость** R кода и **достоверность** передачи информации, часто оценивается **энергетическим выигрышем** от применения такого кодирования [1,2]. В качестве энергетического показателя системы связи выбирают отношение сигнал/шум, которое требуется для достижения заданной вероятности ошибки. *Энергетический выигрыш оценивается как разность значений соотношений сигнал/шум, которые обеспечивают заданную вероятность ошибки в системе без помехоустойчивого кодирования и системе с применением помехоустойчивого кода.* При этом энергетический выигрыш измеряется, как правило, в дБ.

Ниже приводится пример сравнения вероятностей ошибок при передаче двоичных сигналов с двоичной фазовой модуляцией BPSK информационными комбинациями длиной $k=11$ элементов каждая. Сравнение проводится для обычной передачи со скоростью $R_1=4800$ бит/с без помехоустойчивого кодирования и передач с блочным помехоустойчивым кодированием $(n,k)=(15,11)$ с исправлением однократных битовых ошибок. Рассматривается синхронная система связи с непрерывной передачей, поэтому скорость передачи двоичных сигналов при использовании помехоустойчивого кода должна быть увеличена в n/k раз, т.е. равной $R_2=4800 \times 15/11=6545$ бит/с.

Вероятность битовой ошибки в канале с АБГШ и с модуляцией BPSK с антиподными (противоположными) сигналами и когерентным приемом оценивается выражением

$$p_b = \int_{\sqrt{2E_b/N_0}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = Q\left(\sqrt{\frac{2E_b}{N_0}}\right),$$

где $Q(x)$ – гауссов интеграл ошибки, E_b – энергия сигнала, приходящаяся на один двоичный символ, N_0 – спектральная плотность мощности гауссова шума.

Тогда вероятность битовой ошибки при передаче без помехоустойчивого кодирования будет оцениваться выражением

$$p_1 = Q\left(\sqrt{\frac{2E_1}{N_0}}\right),$$

где E_1/N_0 – отношение энергии битового сигнала без помехоустойчивого кодирования к спектральной плотности мощности шума, равно $\frac{E_1}{N_0} = \frac{P_s}{N_0} \left(\frac{1}{R_1}\right)$.

Аналогично, вероятность битовой ошибки при передаче с использованием помехоустойчивого кода оценивается выражением

$$p_2 = Q\left(\sqrt{\frac{2E_2}{N_0}}\right),$$

где E_2/N_0 – отношение энергии битового сигнала с помехоустойчивым кодированием к спектральной плотности мощности шума, равное $\frac{E_2}{N_0} = \frac{P_s}{N_0} \left(\frac{1}{R_2} \right)$.

В приведенных формулах P_s/N_0 – это усредненное отношение мощности несущей к спектральной плотности мощности шума, которое будет одинаковым для обоих случаев передачи и равным [1]

$$\frac{P_s}{N_0} = \frac{E_1 R_1}{N_0} = \frac{E_2 R_2}{N_0}$$

На рис. 2.13 представлены битовые вероятности ошибки $p_1 = G(z1)$ (без помехоустойчивого кодирования, сплошная кривая) и вероятности ошибки $p_2 = Q(z1)$ (с использованием помехоустойчивого кодирования, пунктирная кривая) в зависимости от соотношения сигнал/шум в децибелах $u_1(z1) = 10 \log(P_s/N_0)$. Из графиков видно, что уменьшение длительности элементарного сигнала при использовании помехоустойчивого кодирования приводит к увеличению битовой вероятности ошибки.

Сравним теперь вероятности выдачи потребителю после декодирования комбинации, содержащей $k=11$ информационных элементов, с ошибками.

В случае передачи простым кодом без помехоустойчивого кодирования вероятность ошибочного приема и комбинации будет равна

$$P_1(k=11) = 1 - (1 - p_1)^k.$$

Если же использовать помехоустойчивый блочный код $(n,k)=(15,11)$ с исправлением однократных битовых ошибок, то вероятность получения ошибочного сообщения будет равна

$$P_2(n,k) = P_2(15,11) = \sum_{j=2}^{n=15} C_n^j (p_2)^j (1 - p_2)^{n-j}.$$

Зависимости этих вероятностей от соотношения сигнал/шум в децибелах $u_1(z1) = 10 \log(P_s/N_0)$ показаны на рис. 2.14, где $P1(z1) = P_1(k=11)$ (без помехоустойчивого кодирования, сплошная кривая); $P2(z1) = P_2(n,k) = P_2(15,11)$ (с использованием помехоустойчивого кодирования, пунктирная кривая). Сравнивая кривые, становится очевидным, что применение помехоустойчивого кодирования обеспечивает уменьшение вероятности ошибочного получения сообщения по сравнению с передачей простым кодом. При этом степень повышения достоверности зависит от соотношения сигнал/шум в канале.

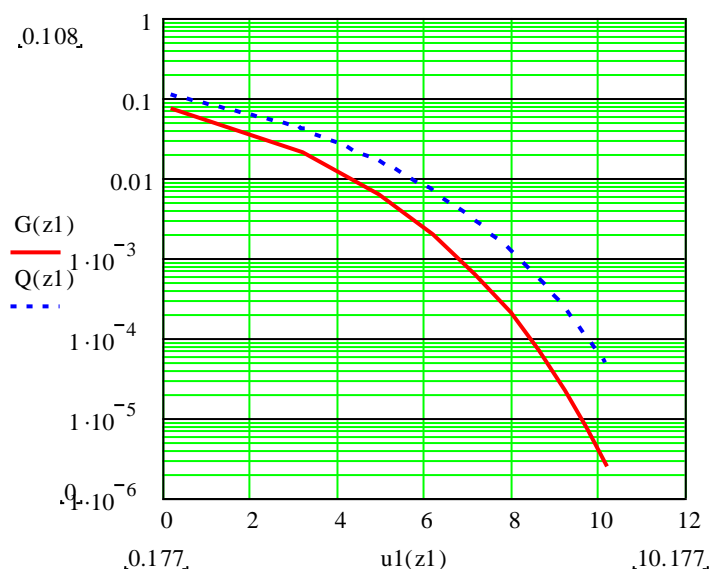


Рис. 2.13. Зависимости вероятностей битовой ошибки при передаче простым кодом длины $k=11$ (сплошная кривая) и избыточным кодом $(15,11)$ с исправлением одиночных ошибок (пунктирная кривая)

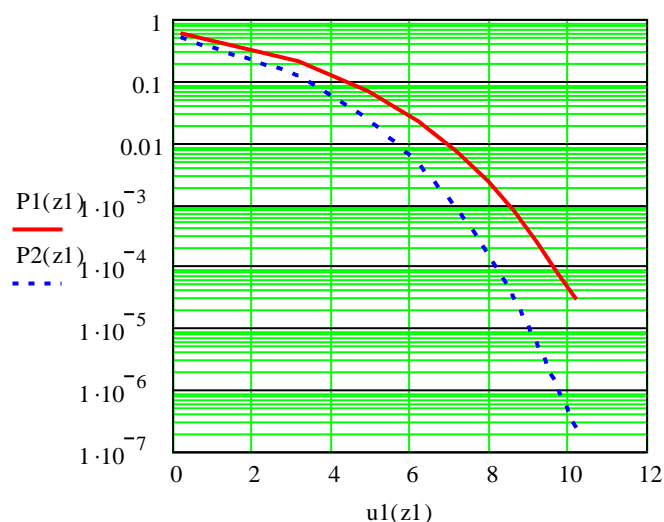


Рис. 2.14. Зависимости вероятностей ошибочного получения комбинации при передаче простым кодом длины $k=11$ (сплошная кривая) и избыточным кодом $(15,11)$ с исправлением одиночных ошибок (пунктирная кривая)

Графики на рис. 2.14 также демонстрируют наличие энергетического выигрыша от применения помехоустойчивого кодирования. Так, например, вероятность ошибочного приема комбинации не хуже 10^{-4} без использования помехоустойчивого кода достигается при соотношении сигнал/шум не ниже 9,6 дБ, тогда как применение помехоустойчивого кода $(15,11)$ с исправлением одиночных ошибок обеспечивает такую же вероятность ошибочного получения сообщения при соотношении сигнал/шум около 8,1 дБ, т.е. энергетический выигрыш составляет около 1,5 дБ.

Ряд авторов оценивают эффективность помехоустойчивого кодирования эквивалентной вероятностью ошибки p_3 . Эта вероятность определяется как вероятность ошибочного приема кодового символа, при которой в некотором гипотетическом (эквивалентном) канале с безизбыточным кодом

обеспечивается такая же вероятность правильного приема комбинации из k информационных символов как и в канале с использованием избыточного помехоустойчивого кода.

Эквивалентная вероятность ошибки по Л. М. Финку определяется следующим образом.

Пусть вероятность правильного приема кодовой комбинации (n,k) -кода в реальном канале с помехоустойчивым кодом будет равна $P_{\text{пп}}$. Те же k информационных элементов в эквивалентном канале с применением кода без избыточности будут приняты без ошибок с вероятностью $(q_3)^k$, где q_3 – вероятность правильного приема одного кодового символа. Приравняв обе вероятности, получим $(q_3)^k = P_{\text{пп}} = 1 - P_{\text{оп}}$, где $P_{\text{оп}}$ – вероятность ошибочного декодирования комбинации избыточного (n,k) -кода. Полученное равенство можно записать как $q_3 = 1 - p_3 = (1 - P_{\text{оп}})^{\frac{1}{k}}$. Тогда, при $P_{\text{оп}} \ll 1$, можно считать, что эквивалентная вероятность ошибочного приёма одного символа будет равна

$$p_3 \approx \frac{P_{\text{оп}}}{k}.$$

Вторая теорема Шеннона.

Вторая теорема Шеннона о возможности передачи информации с произвольно малой частотой ошибок в дискретном канале.

Теорема звучит следующим образом:

«Если скорость производства информации источником сообщений меньше пропускной способности канала, то существует такая система кодирования, что сообщения источника могут быть переданы по каналу с произвольно малой частотой ошибок (или со сколь угодно малой ненадежностью)».

Докажем эту теорему на частном примере двоичного симметричного канала.

Допустим существует источник на выходе которого может появиться одно из N сообщений A_i подлежащих передаче. Обозначим эти сообщения через $(A_1, A_2, A_3, \dots, A_N)$.

Каждое сообщение появляется со своей ему присущей вероятностью P_i , т.е. $(P_1, P_2, P_3, \dots, P_N)$

Энтропия источника H , т.е. количество двоичных единиц информации приходящихся в среднем на одно сообщение, равна:

$$H = - \sum_{i=1}^N p_i \log_2 p_i .$$

Для передачи сообщений по каналу связи сообщения кодируются двоичным кодом. Предположим теперь, что кодируется не каждое отдельное сообщение A_i , а S сообщений вместе (S -цепочки). Если брать S достаточно большим, то каждое сообщение A_i в S -цепочке появится в среднем (с большой вероятностью) $P_i \cdot S$ раз. Таким образом сообщение A_1 появится в какой-либо S -цепочке $P_1 S$ раз, сообщение A_2 – $P_2 S$ раз и т.д.

Если сообщения являются взаимно независимыми, то можно найти вероятность P какой либо реализации блока из S сообщений как

$$P = \prod_{i=1}^N (P_i)^{P_i \cdot S}$$

Таким образом S -цепочки можно считать равновероятными и, следовательно, их количество можно приближенно определить как:

$$M_{\text{вер}} = \frac{1}{P}$$

Учитывая, что все S -цепочки равновероятны, их можно закодировать в соответствии с первой теоремой Шеннона равномерным k -элементным двоичным кодом, где

$$k = \log_2 M_{\text{вер}} = -S \sum_{i=1}^N P_i \log_2 P_i = SH.$$

Допустим, что задан канал с помехами, для которого вероятность ошибки на один кодовый элемент равна P_0 . Для защиты информации от ошибок необходимо ввести избыточное кодирование. Каждый блок из S сообщений (S -цепочки) необходимо закодировать n -элементным помехоустойчивым кодом. Естественнo, что n должно быть больше k .

Для нахождения n определим наиболее вероятное количество ошибок r в одной кодовой комбинации длины n при условии, что S , а следовательно и n , являются достаточно большими. Тогда можно считать $r = P_0 \cdot n$.

При этом чем больше n , тем меньше реальное количество искаженных элементов в кодовой комбинации отличается от r .

Число сочетаний C_n^r дает максимально возможное количество кодовых комбинаций длины n , в которых искажены равно r элементов и которые соответствуют одной S -цепочке.

Таким образом, если на приеме каждому блоку из S исходных сообщений (разрешенной n -элементной комбинации) поставить в соответствие C_n^r кодовых комбинаций, составленных определенным образом, то большая часть ошибок с большой вероятностью будет обнаруживаться и исправляться. Тогда для однозначного правильного декодирования всех S -цепочек понадобится всего $M_{\text{вер}} \cdot C_n^r$ кодовых n -элементных комбинаций. Общее количество комбинаций n -элементного двоичного кода равно 2^n . Следовательно, чтобы закодировать все $M_{\text{вер}}$ сообщений (S -цепочек) различными n -элементными кодовыми комбинациями нужно, чтобы выполнялось условие

$$2^n \geq M_{\text{вер}} \cdot C_n^r$$

Логарифмируя это неравенство, находим

$$n \geq \log_2 M_{\text{вер}} + \log C_n^r = SH + \log \frac{n!}{(P_0 \cdot n)! (1 - P_0 \cdot n)!}$$

где $\vartheta = 1 - P_3$.

Существует ассиметрическая формула для больших n :

$$\log_2 n! \rightarrow n \log_2 n,$$

тогда, с учетом этого, получим

$$\begin{aligned} n &\geq SH + n \log_2 n - P_3 * n \log_2 P_3 * n - n \vartheta \log_2 n \vartheta \\ &= SH + n \log_2 n - P_3 * n \log_2 P_3 - P_3 * n \log_2 n - n * \vartheta \log_2 n - n * \vartheta \log_2 \vartheta \\ &= SH - n(P_3 \log_2 P_3 - \vartheta \log_2 \vartheta) \end{aligned}$$

Поделив обе части неравенства на n , после перестановки получим:

$$\frac{SH}{n} \leq 1 + P_3 \log_2 P_3 + \vartheta \log_2 \vartheta$$

Рассмотрим обе части полученного неравенства. Произведение SH представляет собой количество k двоичных единиц информации переносимое одним блоком из S сообщений, т.е. $SH=k$.

Очевидно, что $\frac{SH}{n}$ – это количество информации переносимое одним элементом избыточного кода, т.е это относительная скорость передачи информации, равная

$$R = \frac{SH}{n} = \frac{k}{n}.$$

Учитывая, что правая часть неравенства представляет собой пропускную способность двоичного симметричного канала C , получаем неравенство $R \leq C$, что и требовалось доказать.

Литература.

1. Скляр, Д. Цифровая связь. Теоретические основы и практическое применение / Д. Скляр; пер. с англ. – М. :Издательский дом «Вильямс», 2003. – 1104 с.
2. Морелос-Сарагоса, Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение / Р. Морелос-Сарагоса; пер. с англ. – М. : Техносфера, 2006. – 319
3. Когновицкий О.С., Охорзин В. М. Теория помехоустойчивого кодирования. Часть 1. Циклические коды.: Учебное пособие/ СПбГУТ. – СПб., 2013. – 94 с.