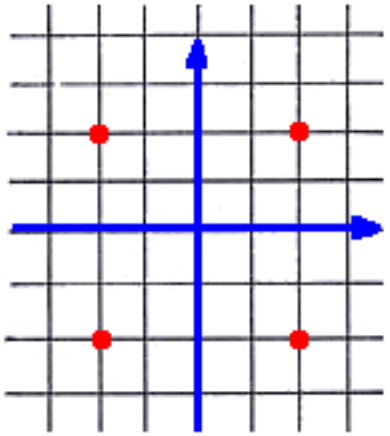
Понятие
о помехоустойчивом кодировании

Теорема Шеннона-Хартли:

$$B = n \log\left(1 + \frac{S}{N}\right)$$

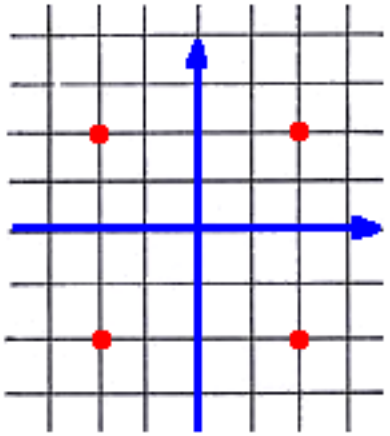
связывает физические параметры канала передачи данных и битовую скорость в канале

Представление сигналов
в пространстве (частота х фаза)

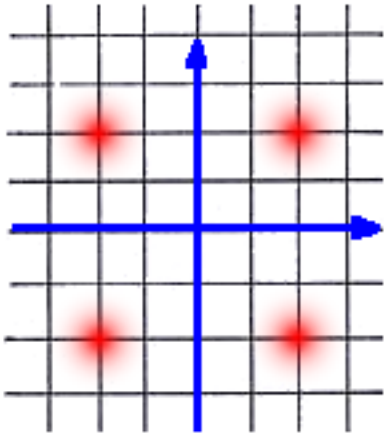


Идеальный
сигнал

Представление сигналов
в пространстве (частота х фаза)

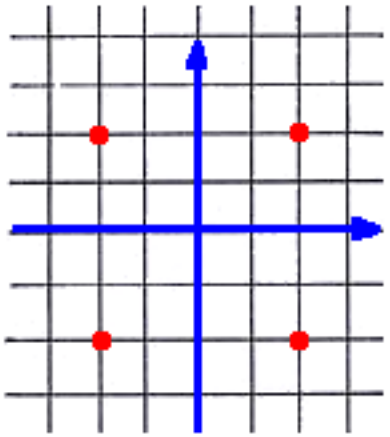


Идеальный
сигнал

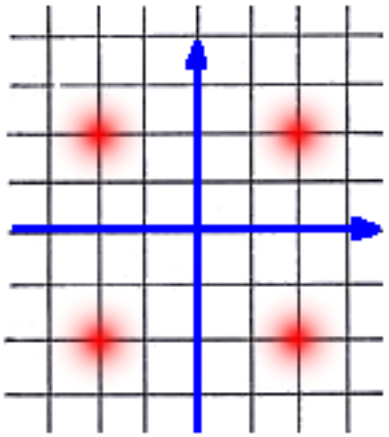


Реальный
сигнал

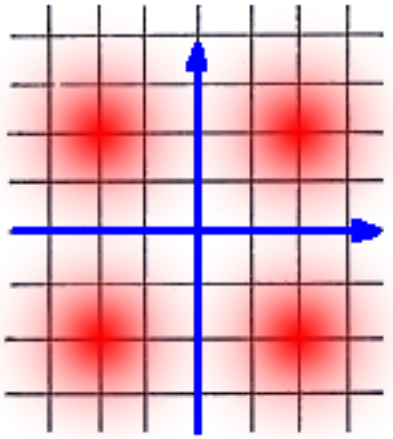
Представление сигналов
в пространстве (частота x фаза)



Идеальный
сигнал

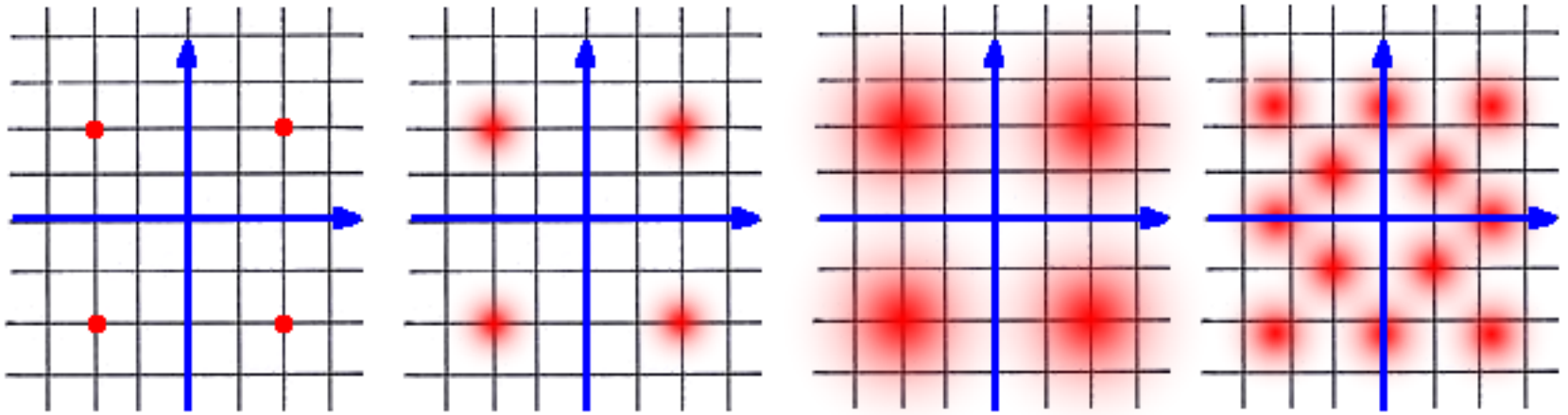


Реальный
сигнал



Реальные сигналы
вблизи
границы Шеннона

Представление сигналов
в пространстве (частота x фаза)



Идеальный сигнал

Реальный сигнал

Реальные сигналы вблизи границы Шеннона

	Ошибкодетектирующие коды	Ошибковосстанавливающие коды
Блок-ориентированные Шины параллельной передачи данных, внутренние шины HDD, SDD, CD-, DVD-ROM, BlueRay, транспортный уровень модели ISO/OSI	стробокоды, коды со счётчиками CRC-12, 16, CCITT, 32 MD4, MD5, BlowFish SHA1, SHA256 Digital Fingerprints Теория: поля неприводимых полиномов над полем Z (поля Галуа)	коды Хемминга БЧХ-коды коды Рида-Соломона Теория: поля неприводимых полиномов над полем Z (поля Галуа), теорема Хемминга
Бит-Ориентированные Шины последовательной передачи данных, SATA, PCI-e, канальный и физический уровни модели ISO/OSI, GSM	Свёрточные коды Решёточные коды (коды Виттерби) Каскадные коды Теория: марковские цепи, теорема максимального правдоподобия (Виттерби)	

Стробовое кодирование

В блок данных (фрейм) добавляются специальные управляющие символы или эскейп-последовательности, определяющие начало и конец блоков данных, количество данных в них.
Могут использоваться для управления потоком.

·DLE-протокол (DLESTX/DLEETX)
·Стартовые фреймы Ethernet (0x55.55.55.55.55.55.55.5d)

Теория Хемминга

Метрика Хемминга (лексикографическое расстояние, расстояние Хемминга, $Hm(A,B)$) – для двух слов одинаковой длины это число различных букв на одинаковых позициях.

Для двоичных слов

$$Hm(\overline{A_N A_{N-1} \dots A_1 A_0}, \overline{B_N B_{N-1} \dots B_1 B_0})$$

равно числу единиц в

$$\overline{A_N A_{N-1} \dots A_1 A_0} \oplus \overline{B_N B_{N-1} \dots B_1 B_0} = \overline{P_n P_{n-1} \dots P_1 P_0}$$

Это число равно $P(1)$, где $P(x) = \sum_{i=0}^{N} P_i x^i$ – называется представляющим полиномом для числа $\overline{P_n P_{n-1} \dots P_1 P_0}$

Теорема Хемминга: Метрика Хемминга эквивалентна метрике в поле многочленов над $\mathbb{Z}_n$ (поля Галуа)

Cyclic Redundant Codes

Для числа $\overline{P_N P_{N_1} \dots P_1 P_0}$ строится представляющий полином $P(x) = \sum_{i=0}^{N} P_i x^i$ тогда

$$R(x) = P(x) x^N \bmod G(x)$$, где

$R(x)$ – полином, представляющий значение CRC

$G(x)$ – т. н. порождающий полином

тогда $CRC(\overline{P_N P_{N_1} \dots P_1 P_0}) = R(1)$

CRC type	N	Порождающий полином
CRC12	12	$x^{12}+x^{11}+x^3+x^2+x+1$
CRC16	16	$x^{16}+x^{15}+x^2+1$
CRC CCITT	16	$x^{16}+x^{12}+x^5+1$
CRC 32	32	$x^{32}+x^{26}+x^{23}+x^{22}+x^{16}+x^{12}+x^{11}+x^{10}+x^8+x^7+x^5+x^4+x^2+x+1$

Ошибковосстанавливающие блочные коды появились при дальнейшем исследовании метрических свойств полей Галуа

Поле Галуа с заданным порождающим полиномом является группой

Свободные делители этой группы соответствуют кодовым словам CRC-кода

Оказалось, что можно построить CRC-коды с минимальным расстоянием между двумя парами не совпадающих кодовых слов

Корректирующая способность кода C это способность восстановить C бит на блок данных

Вторая теорема Хемминга:

$$C = \lfloor \frac{d_{min} - 1}{2} \rfloor$$

Для ошибковосстанавливающих кодов принято обозначение (m,n), где m – число сигнальных (переданных битов), n – число информационных (значащих) битов.
(m – n)/n = избыточность кода на бит

Код Хемминга – простейший ошибковосстанавливающий код имеющий C=1 ($d_{min}=3$)

i_n – информационные биты, r_n – сигнальные корректирующие биты

$$r_1 = i_1 \oplus i_2 \oplus i_3$$
$$r_2 = i_2 \oplus i_3 \oplus i_4$$
$$r_3 = i_1 \oplus i_2 \oplus i_4$$

На вход декодера приходят биты $i'_1, i'_2, i'_3, i'_4, r'_1, r'_2, r'_3$

Вычислим вектор (S_1,S_2,S_3), называемый синдромом

$$S_1 = r'_1 \oplus i'_1 \oplus i'_2 \oplus i'_3$$
$$S_2 = r'_2 \oplus i'_2 \oplus i'_3 \oplus i'_4$$
$$S_3 = r'_3 \oplus i'_1 \oplus i'_2 \oplus i'_4$$

В случае любой одиночной ошибки значение синдрома позволяет определить сбойный бит

Синдром	001	010	011	100	101	110	111
Битовая маска ошибок	0000 001	0000 010	0001 000	0000 100	1000 000	0010 000	0100 000
Сбойный символ	r1	r2	i4	r3	i1	i3	i2

Код Боуза — Чоудхури — Окенгема (БЧХ) позволяет строить системы кодирования с заданной корректирующей способностью

Одним из простейших 16-ричных кодов БЧХ является код Рида-Соломона (44,60) с избыточностью 0,36, который позволяет гарантированно исправлять любую одиночную ошибку на ниббл (полубайт) блока и с некоторой (довольно низкой) вероятностью двойные и тройные ошибки

Свёрточные коды

СССР — Л.М. Финк (1955)

США – Д. Хегельбергер (1959)

Бит-ориентированные (потоковые) коды, допускающие программирование с использованием сдвиговых регистров

Простейший код – код Финка (i_n – сигнальные биты, r_n – корректирующие биты) $i_1 r_1 i_2 r_2 \ldots i_k r_k i_{k+1} r_{k+1}$, где

$$r_k = i_{k-s} \oplus i_{k+s+1}$$

называется условием свёртки.
Число s – называется шагом кодирования
Код (2,3), избыточность 0,5

При ошибочном приёме корректирующего бита r'_n условие свёртки не будет выполнено при $k=n$, а при ошибочном приёме информационного бита условие свёртки нарушается дважды, при $k=n-s-1$ и $k=n+s$. В первом случае бит r'_n можно просто отбросить. Если же условие свёртки нарушено дважды при при k_1 и k_2, таких что $k_2-k_1=2s+1$, то бит r'_{k_1+s+1} исправляется на противоположный. Одновременную ошибку в битах, отстоящих друг от друга на $2s+1$ код не детектирует.

Достоинства

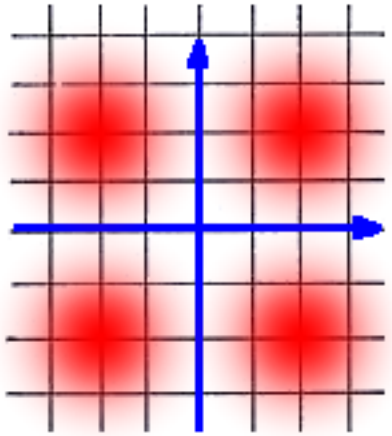
· Простота алгоритма кодирования
· Возможность кодирования в реальном времени
· Для кодирования достаточно сумматора и сдвиговых регистров

Недостатки

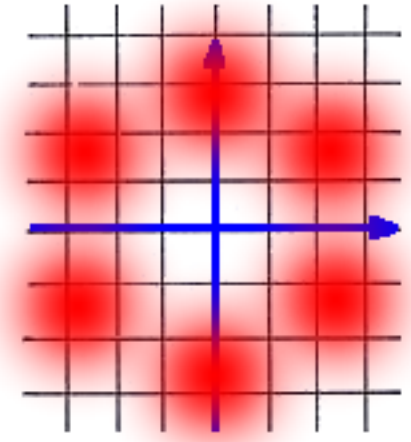
· Декодирование происходит с задержкой на длину шага кодирования
· Сложность алгоритма декодирования растёт полиномиально с увеличением избыточности

· Алгоритм декодирования носит названия алгоритма Виттерби или декодирования на решётке
· Свёрточный код (n,n+1) носит название треллис-кодирования. Корректирующий бит в этом случае носит название треллис-бита

Представление сигналов
в пространстве (частота x фаза)



Исходный сигнал



Сигнал после
применения
кодирования Финка с s=1

Спасибо за внимание!