

Міністерство освіти і науки України



СИГНАЛИ І КОДИ ТЕЛЕКОМУНІКАЦІЙНИХ ТА ІНФОРМАЦІЙНИХ СИСТЕМ

Конспект лекцій

для здобувачів другого (магістерського) рівня вищої освіти
освітньої програми «Телекомунікації та радіотехніка»
галузі знань 6 Інженерія, виробництво та будівництво
спеціальності 65 Електроніка, електронні комунікації,
приладобудування та радіотехніка
денної та заочної форм навчання

Луцьк 2025

УДК 621.396.93 (07)

С 66

Рекомендовано до видання вченою радою факультету КІТ ЛНТУ,
протокол № _____ від « ____ » _____ 20 25 року.

Голова вченої ради факультету КІТ _____ І. С. Кондіус

Електронна копія друкованого видання передана для внесення в репозитарій ЛНТУ

Директор бібліотеки _____ Н. П. Поліщук

Розглянуто і схвалено на засіданні кафедри електроніки та телекомунікацій
ЛНТУ, протокол № _____ від « ____ » _____ 20 25 року.

Завідувач кафедри ЕіТК _____ В. Ю. Заблоцький

Укладач: _____ Й. Р. Селепина, кандидат технічних наук, доцент
кафедри електроніки та телекомунікацій ЛНТУ

Рецензент: _____ Д. А. Захарчук, кандидат фізико-математичних наук,
завідувач кафедри фізики та вищої математики ЛНТУ

Відповідальний за випуск: _____ В. Ю. Заблоцький, кандидат
технічних наук, доцент, завідувач кафедри електроніки та телекомунікацій ЛНТУ

Сигнали і коди телекомунікаційних та інформаційних систем:
конспект лекцій для здобувачів другого (магістерського) рівня вищої
освіти освітньої програми «Телекомунікації та радіотехніка» галузі знань
С 66 галузі знань G Інженерія, виробництво та будівництво спеціальності G5
Електроніка, електронні комунікації, приладобудування та радіотехніка
денної та заочної форм навчання / уклад. Й. Р. Селепина. Луцьк: ЛНТУ,
2025. 108 с.

Конспект лекцій з дисципліни «**Сигнали і коди телекомунікаційних та
інформаційних систем**»: складений відповідно до діючої програми курсу.

Призначений для здобувачів вищої освіти спеціальності G5 Електроніка,
електронні комунікації, приладобудування та радіотехніка освітньої програми
«Телекомунікації та радіотехніка» .

Й. Р. Селепина 2025

ЗМІСТ

1	МОДУЛЬОВАНІ СИГНАЛИ	4
1.1	Загальні положення	4
1.2	Основні види аналогової амплітудної модуляції	5
1.3	Кутова модуляція	12
1.4	Дискретна модуляція гармонічного переносника	19
1.5	Імпульсна модуляція	22
1.6	Модуляція шумової несучої	26
1.7	Енергетичний спектр модульованих сигналів	30
1.8	Цифрові методи модуляції	32
2	МЕТОДИ ОБРОБКИ СИГНАЛІВ	41
2.1	Загальні відомості про приймання сигналів	41
2.2	Приймання сигналів з накопиченням	42
2.3	Кореляційне приймання	47
2.4	Автокореляційне приймання	48
2.5	Когерентне та некогерентне приймання	50
2.6	Приймання на погоджений фільтр	54
3	КОДУВАННЯ ДЖЕРЕЛ ІНФОРМАЦІЇ ТА КАНАЛІВ ЗВ'ЯЗКУ	62
3.1	Основні поняття та означення	62
3.2	Коректувальні коди	66
3.3	Циклічні коди	73
3.4	Коди Боуза–Чоудхурі–Хоквінгема	76
3.5	Несистематичні коди	79
3.6	Ітеративні коди	80
3.7	Ефективність систематичних коректувальних кодів	81
3.8	Мажоритарне декодування	84
3.9	Узагальнення теорії кодування на недвійкові коди	87
3.10	Системи зі зворотним зв'язком	91
3.11	Поєднання процедур демодуляції і декодування	94
3.12	Згорткові (гратчасті) коди	98
	ПЕРЕЛІК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ	107

1 МОДУЛЬОВАНІ СИГНАЛИ

1.1 Загальні положення

Модуляція – це процес змінення одного чи кількох параметрів переносника відповідно до змінення сигналу, який діє на нього. Як переносник широко використовують гармонічні коливання, періодичну послідовність імпульсів і вузькосмуговий випадковий процес. Сигнал, який діє на керований параметр переносника, називається *модулюючим*.

Параметри переносника, які змінюються в часі під впливом модулюючого сигналу, називаються *інформаційними*, оскільки в їх змінних міститься інформація, що передається. Фізичний процес змінення параметрів переносника і є модуляцією. Таким чином, будь-який модулятор (рис. 1.1) має два входи: один – для переносника, інший – для модулюючого сигналу [1].

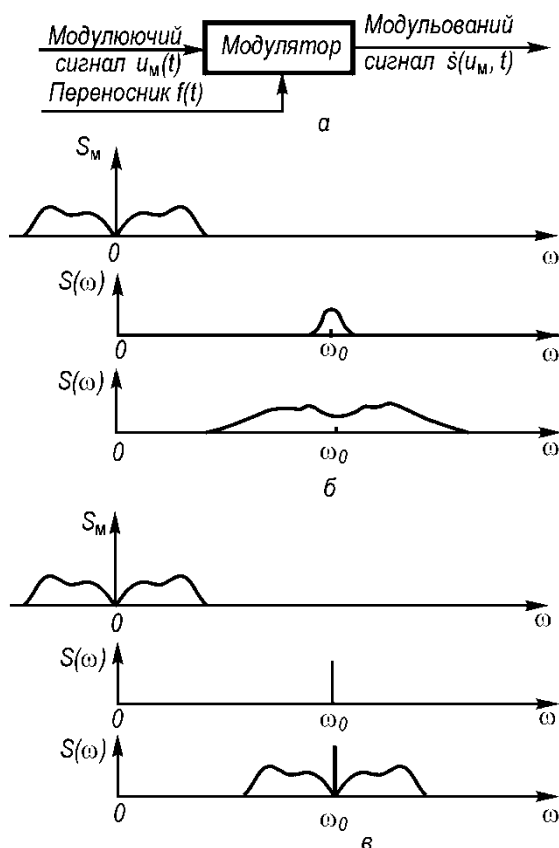


Рисунок 1.1 – Модулятор (а) і зсув спектра при модуляції: загальний випадок аналітичної несучої (б) та гармонічної несучої (в)

Параметри, які залишаються незмінними, є постійними ознаками сигналу. Вони можуть бути використані на прийомі для розрізнення сигналу від завад. У багатьох випадках модулювання сигнал можна подати як добуток двох функцій:

$$s(t) = f(t)M[u_m(t)], \quad (1.1)$$

де $f(t)$ – функція, яка зображує несучу коливання (переносника); $M(t)$ – модуляційна функція, яка виражає дію сигналу, що передається, тобто $u_m(t)$, на несучу $f(t)$.

Якщо для зображення несучої вибирається *аналітичний* сигнал, то для кожної модуляційної функції $M(t)$ існує комплексний модульований сигнал $s(t)$; його дійсна і уявна частини відповідають реально існуючому модульованому сигналу, а модуль визначає обвідну. У випадку, коли несучою є *гармонічне* коливання $A_0 e^{i\omega_0 t}$, модуляційна функція виражає діяння відеосигналу $u_M(t)$ на амплітуду (частоту або фазу) несучої.

Спектр модульованого коливання (1.1) згідно з теоремою про спектр добутку визначається згорткою [1]:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_f(\nu) S_M(\omega - \nu) d\nu, \quad (1.2)$$

де $S_f(\nu)$ – спектральна щільність сигналу $s(t)$; ν – частота модулюючого сигналу $M[u_M(t)]$; S_M – спектральна щільність модульованого сигналу.

Звідси випливає, що процес модуляції призводить до складного перетворення спектра сигналу. Якщо несучою є *вужкосмугове* коливання, то модуляція спричинює розширення спектра і перенесення його в область навколо несучої частоти (рис. 1.1, б). Якщо несуча – чиста *синусоїда*, то має місце просто зсув спектра (рис. 1.1, в).

1.2 Основні види аналогової амплітудної модуляції

Одним із видів аналогової модуляції (тобто первинний сигнал і сигнал-переносник неперервні) є безпосереднє передавання та амплітудна модуляція (АМ), різновиди якої – балансна (БМ) і односмугова (ОМ) модуляції.

Безпосереднє передавання. Найпростішим сигналом для передавання неперервного повідомлення $s(t)$ є сигнал, пропорційний $u_M(t)$:

$$s(t) = A u_M(t), \quad (1.3)$$

де A – якась стала.

Такий сигнал відповідає формулі (1.1), якщо взяти $f(t) = A$ і $M[u_M(t)] = u_M(t)$. Прикладом безпосереднього передавання повідомлень є звичайний телефонний зв'язок по проводах [2].

Амплітудна модуляція (АМ). Для цього виду модуляції

$$s(t) = A_0 e^{i\omega_0 t}, \quad M[u_M(t)] = 1 + m u_M(t), \quad (1.4)$$

де m – коефіцієнт модуляції. Тоді модульований сигнал запишеться так:

$$s(t) = A_0 [1 + m u_M(t)] e^{i\omega_0 t} = A_0 [1 + m u_M(t)] (\cos \omega_0 t + i \sin \omega_0 t). \quad (1.5)$$

Цей вираз дає зображення реального амплітудно-модульованого сигналу (АМ-сигналу):

$$s(t) = \operatorname{Re} s(t) = A_0 [1 + m u_M(t)] \cos \omega_0 t. \quad (1.6)$$

Спектр сигналу в загальному випадку визначається як перетворення Фур'є від $s(t)$:

$$S(i\omega) = \int_{-\infty}^{\infty} s(t)e^{-i\omega t} dt = A_0 \int_{-\infty}^{\infty} \cos \omega_0 t e^{-i\omega t} dt + mA_0 \int_{-\infty}^{\infty} u(t) \cos \omega_0 t e^{-i\omega t} dt.$$

Враховавши, що $\cos \omega t = \frac{1}{2}(e^{i\omega t} + e^{-i\omega t})$, а $\int_{-\infty}^{\infty} e^{i\omega t} dt = 2\pi\delta(\omega)$, отримаємо:

$$S(\omega) = \pi A_0 \delta(\omega - \omega_0) + \frac{mA_0}{2} S_u(\omega_0 + \omega) + \frac{mA_0}{2} S_u(\omega_0 - \omega), \quad (1.7)$$

де $S_u(\omega)$ – спектр повідомлення, яке передається.

Звідси бачимо, що при АМ відбувається перенесення спектра повідомлення на частоту ω_0 (рис. 1.1, в). Ширина спектра сигналу F при АМ удвічі більша за ширину спектра повідомлень F_m , тобто $F = 2F_m$.

При модуляції одним тоном, тобто синусоїдальним сигналом, коли $u(t) = \cos \Omega t$,

$$\begin{aligned} s(t) &= A_0 (1 + m \cos \Omega t) \cos \omega_0 t = \\ &= A_0 \cos \omega_0 t + \frac{mA_0}{2} \cos(\omega_0 + \Omega) + \frac{mA_0}{2} \cos(\omega_0 - \Omega). \end{aligned} \quad (1.8)$$

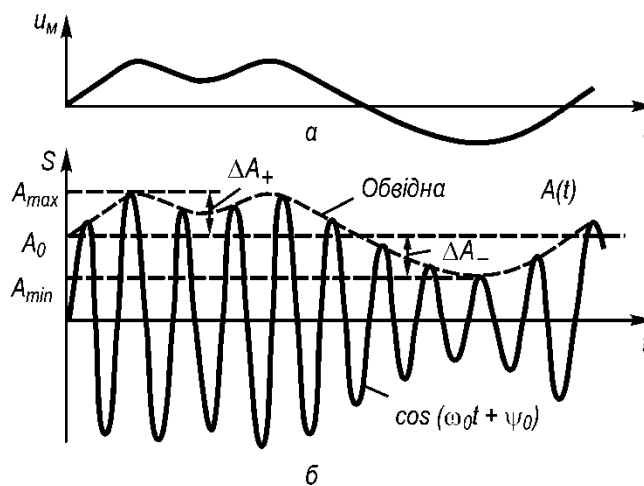


Рисунок 1.2 – Амплітудна модуляція сигналу a – модулюючий сигнал $u_m(t)$; b – АМ-сигнал [2]

При амплітудній модуляції гармонічного переносника приріст амплітуди переносника пропорційний миттєвим значенням модулюючого сигналу $u_m(t)$, тобто:

$$A(t) = A_0 + \Delta A(t) = A_0 + au_m(t) \quad (1.9)$$

де a – коефіцієнт пропорційності; частота і фаза гармонічного переносника залишаються сталими.

Часову діаграму АМ-сигналу наведено на рисунку 1.2, з якого видно, що згідно з миттєвими значеннями $u_m(t)$ амплітуда $A(t)$ то підвищується до A_{\max} , одержуючи при цьому приріст $\Delta A_+ = A_{\max} - A_0 = au_{m\max}$, то зменшується до A_{\min} , одержуючи приріст

$\Delta A_- = A_0 - a u_{m \min}$. Зверніть увагу, що амплітуда повторює форму модулюючого сигналу $u_m(t)$. В АМ-сигналі амплітуда $A(t)$ є обвідною високочастотного заповнення $\cos(\omega_0 + \psi_0)$.

Коефіцієнт модуляції. Для математичного опису АМ-сигналу в формулу (1.9) замість коефіцієнта a , що залежить від конкретної схеми модулятора, вводять коефіцієнт модуляції $m_{AM} = \Delta A_{\text{сеп}} / A_0$, який дорівнює відносному значенню приросту амплітуди. Тут $\Delta A_{\text{сеп}} = (\Delta A_+ + \Delta A_-) / 2$ – середньоарифметичне значення приросту амплітуди АМ-сигналу. Оскільки середнє значення амплітуди АМ-сигналу за час модуляції $A_0 = (A_{\text{max}} + A_{\text{min}}) / 2$, то коефіцієнт модуляції:

$$m_{AM} = \Delta A_{\text{сеп}} / A_0 = (A_{\text{max}} - A_{\text{min}}) / (A_{\text{max}} + A_{\text{min}}). \quad (1.10)$$

Отже, коефіцієнт модуляції – це відношення різниці між максимальним і мінімальним значеннями амплітуд АМ-сигналу до суми цих значень. Досить часто коефіцієнт модуляції визначається у відсотках, але при всіх розрахунках АМ-сигналів користуються коефіцієнтом модуляції m_{AM} у відносних одиницях.

Для симетричного модулюючого сигналу $u_m(t)$ (тобто середнє значення амплітуди якого дорівнює нулю) АМ-сигнал також симетричний: $\Delta A_+ = \Delta A_- = \Delta A$ і

$$m_{AM} = \Delta A / A_0; \quad (1.11)$$

отже, коефіцієнт модуляції дорівнює відношенню максимального приросту амплітуди до амплітуди переносника. Він характеризує глибину амплітудної модуляції і може змінюватись у межах $0 \leq m_{AM} \leq 1$.

Приклад 1.1. Визначити коефіцієнт модуляції АМ-сигналу, числову діаграму якого в лінійному масштабі зображено на рисунку 1.2, б.

Для визначення коефіцієнта модуляції скористаємося формулою (1.10). При розрахунках знати абсолютне значення амплітуд не обов'язково. Обчислення можна зробити в умовних одиницях. За 1 ум. од. візьмемо 1 мм на вертикальній осі амплітуд. Тоді $A_{\text{min}} = 4$ ум. од., $A_{\text{max}} = 12$ ум. од. і $m_{AM} = (A_{\text{max}} - A_{\text{min}}) / (A_{\text{max}} + A_{\text{min}}) = (12 - 4) / (12 + 4) = 0,5 \cdot 100 = 50 \%$.

Амплітудна модуляція гармонічним коливанням. У найпростішому випадку модулюючий сигнал $u_m(t)$ є гармонічним коливанням із частотою Ω і початковою фазою ψ . При цьому:

$$s_{AM}(u_m, t) = A_0 [1 + m_{AM} \cos(\Omega t + \psi) \cos(\omega_0 t + \psi_0)] \quad (1.12)$$

є аналітичним виразом (математичною моделлю) однотонального АМ-сигналу, тобто модульованого одним гармонічним коливанням тональної частоти. На рисунку.1.3, а, б зображено часові діаграми однотонального АМ-сигналу при різних значеннях m_{AM} . На ньому дуже добре видно симетричність модуляції та характерні спотворення при *перемодуляції*, коли форма обвідної вже не повторює форму модулюючого гармонічного коливання (рис. 1.3, в) [2].

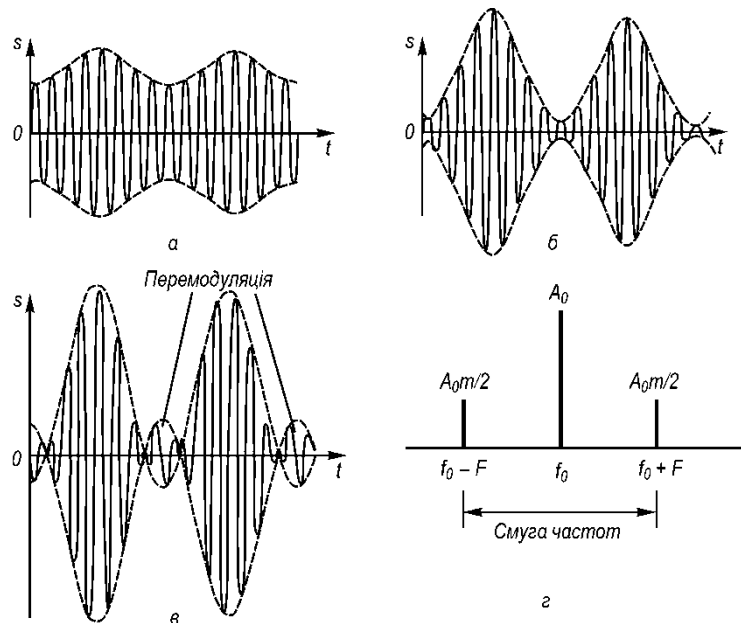


Рисунок 1.3 – Однотональна амплітудна модуляція:

a – з малим коефіцієнтом модуляції; *b* – глибока модуляція; *в* – перемодуляція; *г* – спектральна діаграма АМ-сигналу для однотональної модуляції [3]

Однотональний АМ-сигнал можна подати також у вигляді суми гармонічних складових. Якщо використати тригонометричну формулу добутку косинусів $\cos \alpha \cos \beta = 0,5[\cos(\alpha + \beta) + \cos(\alpha - \beta)]$, із формули (1.12) дістанемо:

$$\begin{aligned}
 s_{\text{AM}}(u_m, t) = & A_0 \cos(\omega_0 t + \psi_0) + \\
 & + 0,5 A_0 m_{\text{AM}} \cos[(\omega_0 + \Omega)t + \psi_0 + \psi] + \\
 & + 0,5 A_0 m_{\text{AM}} \cos[(\omega_0 - \Omega)t + \psi_0 - \psi].
 \end{aligned}
 \tag{1.13}$$

Із формули (1.13) випливає, що однотональний АМ-сигнал має три гармонічні спектральні складові з частотами f_0 (переносника); $f_0 + F$ (верхньою бічною); $f_0 - F$ (нижньою бічною).

Спектральна діаграма однотонального АМ-сигналу, побудована за формулою (1.13), є симетричною відносно частоти переносника f_0 (рис. 1.3, *г*). Амплітуди бічних коливань однакові, і навіть для $m_{\text{AM}} = 1$ не перевищують половини амплітуди переносника A_0 .

Амплітудна модуляція при складному модулюючому сигналі. Гармонічні модулюючі сигнали і відповідно однотональний АМ-сигнал практично зустрічаються рідко. У більшості випадків модулюючі первинні сигнали є складними функціями часу. Аналітичний вираз АМ-сигналу можна подати у вигляді формули:

$$s_{\text{AM}}(u_m, t) = A_0 [1 + m_{\text{AM}} u_m(t)] \cos(\omega_0 + \psi_0).
 \tag{1.14}$$

Спектр АМ-сигналу при складному модулюючому сигналі якісно визначається з таких міркувань. Будь-який складний сигнал $u_m(t)$ можна подати у вигляді скінченної (чи нескінченної) суми гармонічних складових, якщо

скористатись для цього рядом чи інтегральним перетворенням Фур'є. Кожна гармонічна складова сигналу $u_m(t)$ з частотою Ω_i сприяє появі в спектрі АМ-сигналу двох бічних складових з частотами $f_0 \pm F_i$, а множина гармонічних складових модулюючого сигналу $\sum F_i$ – множині бічних складових з частотами $\sum(f_0 \pm F_i)$. Для наочності такі перетворення спектра для АМ наведено на рисунку 1.4.

З рисунка 1.4 видно, що в спектрі складного АМ-сигналу, крім складової з частотою переносника f_0 , містяться групи *верхніх* ($f_0 + F$) та *нижніх* ($f_0 - F$) *бічних коливань*, що утворюють відповідно верхню та нижню бічні смуги частот АМ-сигналу. При цьому верхня бічна смуга частот є масштабною копією як дискретного, так і неперервного спектра модулюючого сигналу, зсунутого за частотою на величину f_0 . Нижня бічна смуга частот також повторює спектральну діаграму (спектральну щільність) сигналу $u_m(t)$, але з дзеркальним розташуванням частот (у зворотному порядку) відносно частоти переносника f_0 .

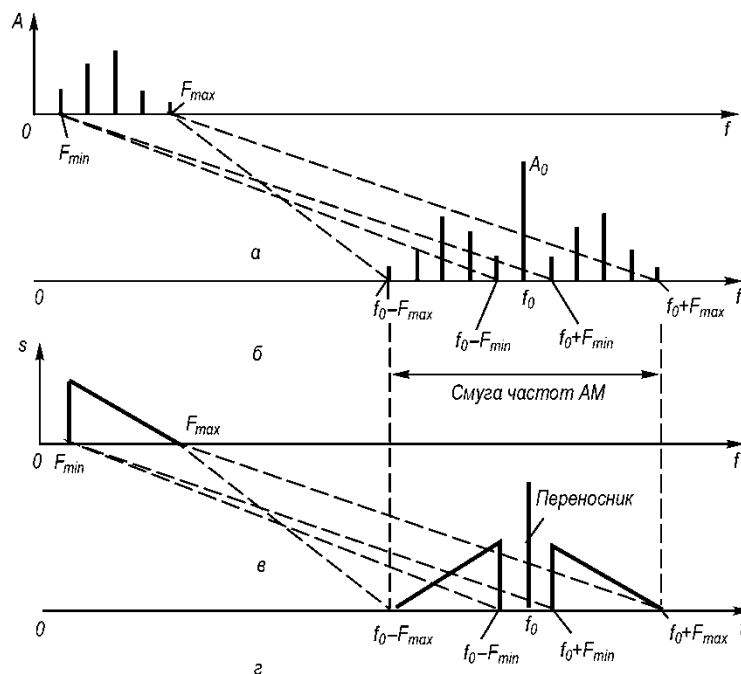


Рисунок 1.4 – Спектральні діаграми:

а – модулюючого сигналу; *б* – АМ-сигналу для багатоканальної модуляції; *в* – модулюючого сигналу з неперервним спектром; *г* – АМ-сигналу для модуляції сигналом із неперервним спектром [3]

Із зазначеного можна дійти важливого висновку: *ширина спектра АМ-сигналу* F_{AM} *дорівнює подвоєному значенню максимальної частоти* F_{max} *спектра модулюючого сигналу, тобто:*

$$F_{AM} = 2F_{max} \quad (1.15)$$

З цього виразу випливає, що амплітуда модульованого сигналу змінюється від $A_{min} = A_0(1 - m)$ до $A_{max} = A_0(1 + m)$, а потужність сигналу – від $P_{min} = P_n(1 - m)^2$ до

$P_{\max} = P_H(1+m)^2$, де $P_H = \frac{A_0^2}{2}$ – потужність несучого коливання. Середня потужність АМ-сигналу дорівнює:

$$P_{\text{ср}} = \frac{A_0^2}{2T} \int_0^T (1 + m \cos \Omega t)^2 dt = P_H \left(1 + \frac{m^2}{2} \right). \quad (1.16)$$

При $m=1$ $P_{\max} = 4P_H$, а $P_{\text{ср}} = 1,5P_H$; відношення середньої потужності до максимальної дорівнює 0,375, що свідчить про суттєвий недолік амплітудної модуляції – погане використання потужності передавача.

Балансна модуляція (БМ). Крім звичайної АМ, застосовується передавання АМ-сигналу без несучої – балансна модуляція. Для цього виду модуляції:

$$f(t) = A_0 e^{i\omega_0 t}, \quad M[u(t)] = u(t); \quad (1.17)$$

Тоді

$$\begin{aligned} s(t) &= A_0 u(t) \cos \omega_0 t + i A_0 u(t) \sin \omega_0 t; \\ s(t) &= \operatorname{Re} \left[A_0 u(t) e^{i\omega_0 t} \right]. \end{aligned} \quad (1.18)$$

Спектр сигналу при БМ:

$$S(\omega) = \frac{1}{2} A_0 [S_u(\omega_0 + \omega) + S_u(\omega_0 - \omega)]. \quad (1.19)$$

Тут знаходяться тільки дві бічні смуги, оскільки несуча відсутня.

При **односмуговій модуляції (ОМ)** передається тільки одна бічна смуга. При передаванні *верхньої бічної смуги*:

$$f(t) = A_0 e^{i\omega_0 t}, \quad M[u_M(t)] = u_M(t) + i u_M^*(t); \quad (1.20)$$

$$\begin{aligned} s(t) &= A_0 [u_M(t) \cos \omega_0 t - u_M^*(t) \sin \omega_0 t] + \\ &+ i A_0 [u_M(t) \sin \omega_0 t + u_M^*(t) \cos \omega_0 t]; \\ s(t) &= \operatorname{Re} \left[A_0 [u_M(t) \cos \omega_0 t - u_M^*(t) \sin \omega_0 t] \right]. \end{aligned} \quad (1.21)$$

Спектр сигналу ОМ:

$$S(\omega) = \frac{A_0 m}{2} S_u(\omega_0 + \omega). \quad (1.22)$$

Дійсно, якщо розкласти функції $u_M(t)$ і $u_M^*(t)$ у ряд Фур'є:

$$\begin{aligned} u_M(t) &= \sum_{k=0}^{\infty} C_k \cos(\Omega_k t + \varphi_k); \\ u_M^*(t) &= \sum_{k=0}^{\infty} C_k \sin(\Omega_k t + \varphi_k) \end{aligned}$$

і врахувати, що $\cos x$ і $\sin x$ є парою перетворення Гільберта, то отримаємо:

$$s(t) = \sum_{k=0}^{\infty} C_k \cos[(\omega_0 + \Omega_k)t + \varphi_k] + i \sum_{k=0}^{\infty} C_k \sin[(\omega_0 + \Omega_k)t + \varphi_k].$$

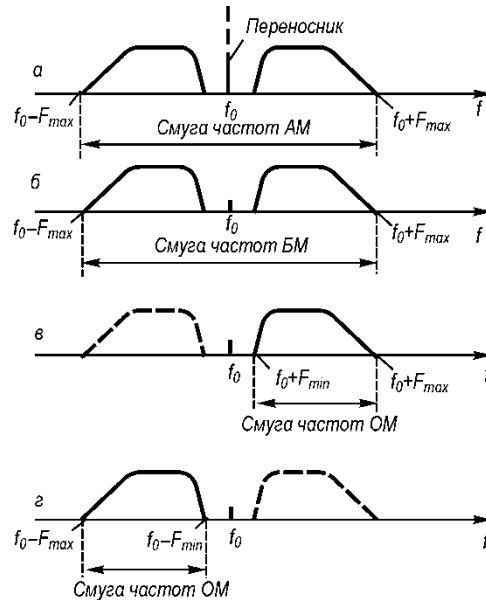


Рисунок 1.5 – Спектральні діаграми:

a – АМ-сигналу; *б* – БМ-сигналу; *в* – ОМ-сигналу (верхня бічна смуга); *г* – ОМ-сигналу (нижня бічна смуга) [4]

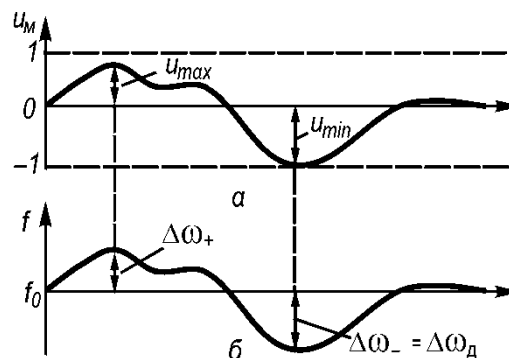


Рисунок 1.6 – Миттєва частота ЧМ-сигналу: *a* – модулюючий сигнал; *б* – зміна миттєвої частоти [4]

Таке зображення є аналітичним для усіх $\omega_0 > 0$. Із заміною модуляційної функції $M[u_m(t)]$ на спряжену $M^*[u_m(t)] = u_m(t) - iu_m^*(t)$ форма сигналу $s(t)$ відповідає нижній бічній смузі.

Спектри БМ- і ОМ-сигналів можна дістати зі спектра АМ-сигналу, якщо з нього вилучити або складову на частоті переносника для БМ-сигналу, або складову на частоті переносника та одну з бічних смуг (верхню чи нижню) для ОМ. Такі перетворення спектра АМ-сигналу показані на рисунку 1.5.

Важливою перевагою БМ- і ОМ-сигналів є підвищення швидкості ефективності використання потужності передавача, що покращує відповідно якість приймання таких сигналів. Крім того, при ОМ у два рази зменшується ширина спектра модульованого сигналу, що дозволяє вдвічі збільшити кількість сигналів у заданій смузі частот. Тому ОМ широко застосовується в багатоканальному зв'язку з частотним розділенням.

1.3 Кутова модуляція

До кутових видів модуляції належать частотна (ЧМ) та фазова (ФМ) модуляції. Для частотної модуляції відхилення частоти модульованого сигналу від частоти переносника ω_0 змінюється пропорційно миттєвим значенням модулюючого сигналу:

$$\omega(t) = \omega_0 + \Delta\omega_d u_M(t), \quad (1.23)$$

де $\Delta\omega_d$ – коефіцієнт пропорційності, який називають *девіацією частоти* (від лат. deviato – відхилення); він дорівнює максимальному відхиленню частоти модульованого сигналу від частоти переносника ω_0 .

Зміну частоти ЧМ-сигналу графічно подано на рис. 1.6, де зазначено девіацію частоти $\Delta\omega_d$, яка відповідає найбільшому відхиленню частоти вниз: $\Delta\omega_d = \Delta\omega_-$, оскільки $\Delta\omega_+ < \Delta\omega_-$. У виразі (1.23) величина $u_M(t)$ нормована, тобто $|u_M(t)| \leq 1$.

Девіація частоти є одним із головних параметрів частоти модуляторів і може набувати значень від одиниць герців до сотень мегагерців у модуляторах різного призначення при обов'язковому виконанні умови $\Delta\omega_d \ll \omega_0$.

Для фазової модуляції відхилення (зсув) фази від лінійної $\cos(\omega_0 t + \psi_0)$ змінюється пропорційно миттєвим значенням модулюючого сигналу:

$$\psi(t) = \omega_0 t + \psi_0 + \Delta\omega_d u_M(t). \quad (1.24)$$

Фізичний зміст цього коефіцієнта пояснює рисунок 1.7, де зображені модулюючий сигнал і повна фаза ФМ-сигналу. Якщо $u_M(t)$ зростає, то і повна фаза зростає, причому швидше, ніж за лінійним законом. Якщо ж сигнал $u_M(t)$ спадає, маємо зменшення швидкості зростання фази $\psi(t)$. Абсолютна величина відхилення (зсуву) фази від лінійного закону найбільша, коли $u_M(t)$ досягає екстремальних значень.

На рисунку 1.7, б зазначені максимальні відхилення фази вгору $\Delta\varphi_+$ та вниз $\Delta\varphi_-$. Найбільше відхилення фази від лінійної і є девіацією фази $\Delta\varphi_d$ для ФМ. У прикладі, що наведений на рисунку 1.7, $\Delta\varphi_d = \Delta\varphi_+$. Вимірюється $\Delta\varphi_d$ в радіанах і може набирати значень від одиниць до десятків тисяч радіан.

Аналітичний вираз (математичну модель) ФМ-сигналу можна записати так:

$$s_{\text{ФМ}}(u_M, t) = A_0 \cos[\omega_0 t + \psi_0 + \Delta\omega_d u_M(t)]. \quad (1.25)$$

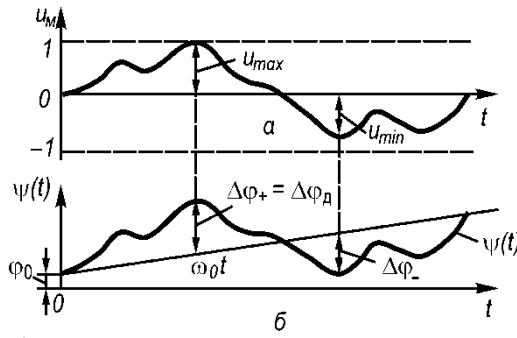


Рисунок 1.7 – Повна фаза ФМ-сигналу: *a* – модулюючий сигнал; *б* – зміна повної фази [4]

У випадку кутової модуляції (ЧМ та ФМ) модульована функція має вигляд:

$$M[u_M(t)] = e^{i\varphi(t)}. \quad (1.26)$$

При синусоїдальній несучій $f(t) = A_0 e^{i\omega_0 t}$ модульований сигнал матиме наступний вираз:

$$s(t) = A_0 e^{i[\omega_0 t + \varphi(t)]} = A_0 \cos[\omega_0 t + \varphi(t)] + iA_0 \sin[\omega_0 t + \varphi(t)]. \quad (1.27)$$

Реальний сигнал:

$$s(t) = \text{Re} \{ A_0 e^{i[\omega_0 t + \varphi(t)]} \} = A_0 \cos \psi(t). \quad (1.28)$$

Це звичайне зображення сигналу з кутовою модуляцією. Згідно з (1.28) повна фаза високочастотного коливання дорівнює:

$$\psi(t) = \omega_0 t + \varphi(t), \quad (1.29)$$

а миттєва частота коливання змінюється за законом похідної від $\psi(t)$, тобто:

$$\omega(t) = \frac{d\psi(t)}{dt} = \omega_0 + \frac{d\varphi(t)}{dt}. \quad (1.30)$$

Навпаки, у разі змінення частоти за законом $\omega(t)$ (1.30) фаза коливання:

$$\psi(t) = \int \omega(t) dt. \quad (1.31)$$

У випадку фазової модуляції $\varphi(t) = \varphi_0 + \Delta\varphi_m u(t)$, де $\Delta\varphi_m$ – девіація фази. Тоді на підставі (1.28) і (1.29):

$$\varphi(t) = \omega_0 t + \Delta\varphi_m u(t) + \varphi_0; \quad (1.32)$$

$$s(t) = A_0 \cos[\omega_0 t + \Delta\varphi_m u(t) + \varphi_0]. \quad (1.33)$$

При частотній модуляції за законом первинного сигналу, який характеризує повідомлення, змінюється частота несучої коливання:

$$\omega = \omega_0 + \Delta\omega_d u(t), \quad (1.34)$$

де $\Delta\omega_d$ – амплітуда частотного відхилення (девіація частоти).

Повна фаза коливання при цьому:

$$\psi(t) = \int \omega(t) dt + \varphi_0 = \omega_0 t + \Delta\omega_d \int u(t) dt + \varphi_0. \quad (1.35)$$

Тоді ЧМ-сигнал:

$$s(t) = A_0 \cos \left[\omega_0 t + \Delta\omega_d \int u(t) dt + \varphi_0 \right]. \quad (1.36)$$

При модуляції одним тоном, коли $u(t) = \cos \Omega t$, сигнали при ФМ і ЧМ за формою мають однаковий вигляд:

$$s(t) = A_0 \cos(\omega_0 t + m \cos \Omega t + \varphi_0), \quad (1.37)$$

де m – індекс модуляції: при фазовій модуляції $m = \Delta\varphi_m$, а при частотній модуляції $m = \frac{\Delta\omega_d}{\Omega}$.

Для визначення спектра сигналу замінимо в (1.37) косинус суми двох кутів за відомими тригонометричними формулами, взявши для спрощення $\varphi_0 = 0$:

$$s(t) = A_0 \cos(m \sin \Omega t) \cos \omega_0 t - A_0 \sin(m \sin \Omega t) \sin \omega_0 t. \quad (1.38)$$

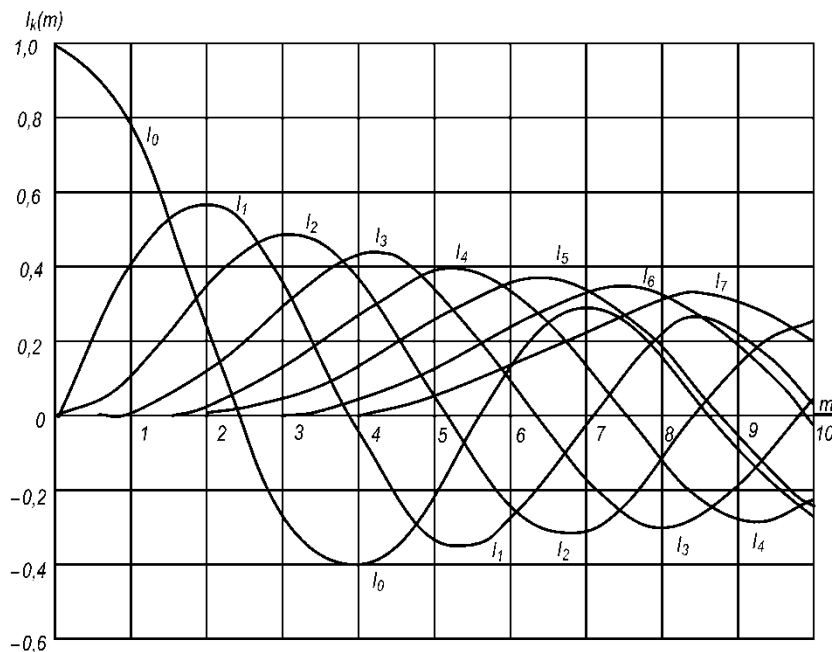


Рисунок 1.8 – Графіки функції Бесселя [5]

З теорії бesselевих функцій відомі наступні співвідношення:

$$\sin(m \sin \Omega t) = 2 \sum_{k=1}^{\infty} I_{2k-1}(m) \sin(2k-1)\Omega t, \quad (1.39)$$

$$\cos(m \sin \Omega t) = I_0(m) + 2 \sum_{k=1}^{\infty} I_{2k}(m) \cos 2k\Omega t, \quad (1.40)$$

де $I_k(m)$ – бesselева функція першого роду k -го порядку від аргументу m (рис. 1.8).

Після підстановки (1.39) і (1.40) в (1.38) отримаємо:

$$s(t) = A_0 I_0(m) \cos \omega_0 t + \sum_{k=1}^{\infty} A_0 I_k(m) \cos(\omega_0 + k\Omega)t + \sum_{k=1}^{\infty} (-1)^k A_0 I_k(m) \cos(\omega_0 - k\Omega)t. \quad (1.41)$$

Таким чином, навіть при синусоїдальних ЧМ і ФМ отримаємо теоретично нескінченний спектр. Він складається з несучої ω_0 і двох бічних смуг $\omega_0 \pm k\Omega$. Амплітуда несучої $A_0 I_0(m)$ при ЧМ і ФМ, на відміну від АМ, залежить від модулюючого коливання. При деяких значеннях m ($m=2, 3, 4, 5$) вона може дорівнювати нулю. Амплітуда бічних частот $A_k = A_0 I_k(m)$.

Однак практично ширина спектра ЧМ- і ФМ-сигналів обмежена. На рисунку 1.9 наведено спектр сигналу з кутовою модуляцією одним тоном при $m=5$. Як бачимо, амплітуди бічних частот швидко зменшуються зі збільшенням номера гармоніки k . При $k > m$ складові спектра малі, тому ними можна знехтувати. Практично ширина спектра сигналу при кутовій модуляції дорівнює $F = 2(m+1)F_m$, де $F_m = \frac{\Omega}{2\pi}$ – частота модулюючого коливання.

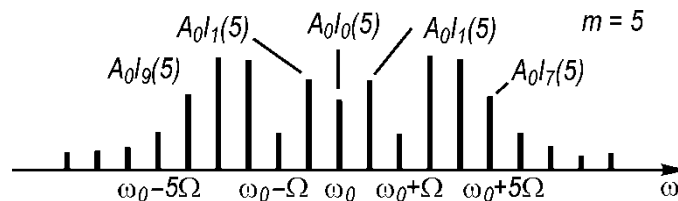


Рисунок 1.9 – Спектр сигналу з кутовою модуляцією [5]

Різницю між ФМ і ЧМ можна краще усвідомити, розглядаючи ФМ- і ЧМ-сигнали при модуляції повідомлення $u(t) = \cos \Omega t$. У цьому випадку:

$$s_{\text{ФМ}}(t) = A_0 \cos(\omega_0 t + \Delta \varphi_m \cos \Omega t + \varphi_0); \quad (1.42)$$

$$\begin{aligned} s_{\text{ЧМ}}(t) &= A_0 \cos\left(\omega_0 t + \Delta \varphi_m \int \cos \Omega t dt + \varphi_0\right) = \\ &= A_0 \cos\left(\omega_0 t + \frac{\Delta \omega_m}{\Omega} \sin \Omega t + \varphi_0\right). \end{aligned} \quad (1.43)$$

Порівнюючи (1.42) і (1.43), визначимо еквівалентне значення зміни фази при ЧМ:

$$\Delta\varphi_{\text{екв ЧМ}} = \frac{\Delta\omega_m}{\Omega}. \quad (1.44)$$

З отриманого виразу випливає, що при ЧМ зміна фази сигналу обернено пропорційна частоті модулюючого повідомлення, тоді як при ФМ $\Delta\varphi$ є величина, яка залежить не від частоти модулюючого повідомлення, а від амплітуди:

$$\Delta\omega_{\text{екв ФМ}} = \Delta\varphi_m \Omega. \quad (1.45)$$

Отже, при ЧМ девіація частоти не залежить від модулюючого повідомлення, а при ФМ девіація частоти прямо пропорційна частоті модулюючого сигналу. Залежності (1.44) і (1.45) проілюстровані на рисунку 1.10 [4].

Різниця між ЧМ і ФМ полягає і в тому, що ширина спектра ФМ-коливання залежить від ширини спектра модулюючого повідомлення, тоді як ширина спектра при ЧМ практично не залежить від ширини спектра модулюючого повідомлення, якщо $\Delta\omega_m \ll \Omega$, і дорівнює:

$$\Delta f_{\text{ЧМ}} \approx \frac{2\Delta\omega_m}{2\pi}.$$

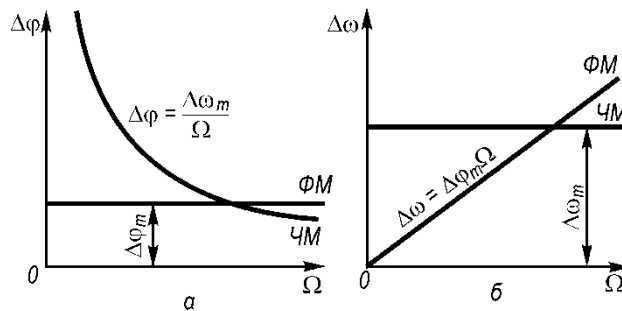


Рисунок 1.10 – Залежність фази (а) і частоти (б) модульованих коливань при ФМ і ЧМ [4]

Функції Бесселя мають цікаву закономірність: чим вищий порядок k функції Бесселя, тим при більших значеннях аргументу m спостерігається її максимум, але якщо $k > m$, значення функції Бесселя є малою величиною. З цього випливає, що малими будуть і складові спектра, а тому ними можна нехтувати.

Ширина спектра кутових модуляцій залежить від того, з якою амплітудою ми відкидаємо складові спектра. Практично вважають, що можна нехтувати тими спектральними складовими, номери яких $k > m + 1$ (значення амплітуди їх менше за 5 % амплітуди переносника). З цього випливає, що ширина спектра сигналів з кутовими модуляціями дорівнює:

$$F_{\text{ЧМ}} = F_{\text{ФМ}} \approx 2(m + 1)F_m, \quad (1.46)$$

де F_m – частота гармонічного модулюючого сигналу.

Інколи вважають, що для забезпечення передавання модулюючого сигналу з більш високою точністю потрібно передавати спектральні складові із значеннями амплітуди до 1 % амплітуди переносника. Тоді ширина кутових модуляцій:

$$F_{\text{ЧМ}} = F_{\text{ФМ}} \approx 2(m + \sqrt{m} + 1)F_M.$$

Якщо $m < 0,6$, то кутова модуляція вважається *вузькосмуговою*; її ширина спектра порівнянна з шириною спектра амплітудної модуляції.

Якщо $m \gg 1$, то кутова модуляція є *широкосмуговою*; з формули (1.46) випливає, що ширина смуги частот приблизно дорівнює подвоєній девіації частоти. Кутові модуляції, особливо широкосмугові, мають більшу завадостійкість, ніж амплітудна модуляція, тому вони знаходять широке застосування в системах зв'язку для якісного передавання повідомлень. Однак слід пам'ятати, що при цьому значно розширюється смуга частот модульованого сигналу.

Якщо модулюючий сигнал є складним, тобто містить у собі ряд гармонічних складових, то спектр ЧМ і ФМ можна знайти описаним раніше способом, використавши вираз для косинуса кількох аргументів. Спектр модульованого сигналу при цьому буде дуже складним і міститиме різні комбінаційні частоти. Але і в цьому разі загальна смуга частот модульованого сигналу наближено обчислюється за формулою (1.46), куди замість F_M підставляється максимальна частота спектра модулюючого сигналу; індекс модуляції обчислюється на цій максимальній частоті.

У випадку малого індексу модуляції спектр ЧМ- і ФМ-сигналів, як і для АМ, має тільки три складові:

$$s(t) = A_0 \cos \omega_0 t + \frac{mA_0}{2} \cos(\omega_0 + \Omega)t - \frac{mA_0}{2} \cos(\omega_0 - \Omega)t. \quad (1.47)$$

Це безпосередньо випливає з (1.41), якщо врахувати, що при $m \ll 1$ $\sin(m \sin \Omega t) \approx m \sin \Omega t$, а $\cos(m \sin \Omega t) \approx 1$.

Порівняння (1.8) і (1.47) показує, що розбіжність спектрів сигналів при амплітудній і кутовій модуляції полягає тільки в зсуві фаз коливання нижньої бічної частоти на 180° відносно його положення при АМ. Ця розбіжність є суттєвою та ілюструється векторними діаграмами, зображеними на рисунку 1.11.

Досить часто використовується *односмугова* кутова модуляція. Якщо функція $\varphi(t)$ аналітична, тобто:

$$\varphi(t) = \varphi(t) + i\varphi^*(t),$$

то сигнал:

$$\begin{aligned} s(t) &= A_0 e^{i(\omega_0 t + \varphi(t))} M[u_m(t)] = A_0 e^{i[\varphi(t) + i\varphi^*(t)]} = \\ &= A_0 e^{-\varphi^*(t)} \sin[\omega_0 t + \varphi(t)] \end{aligned} \quad (1.48)$$

також є аналітичною функцією при $\omega_0 > 0$. Він не містить від'ємних частот, хоча і має нескінченний спектр в області додатних частот:

$$s(t) = \text{Re} \left\{ A_0 e^{i(\omega_0 t + \varphi(t))} \right\} = A_0 e^{-\varphi^*(t)} \cos[\omega_0 t + \varphi(t)]. \quad (1.49)$$

Вираз (1.47) визначає новий модулюючий сигнал, який є варіантом сигналу односмугової кутової модуляції.

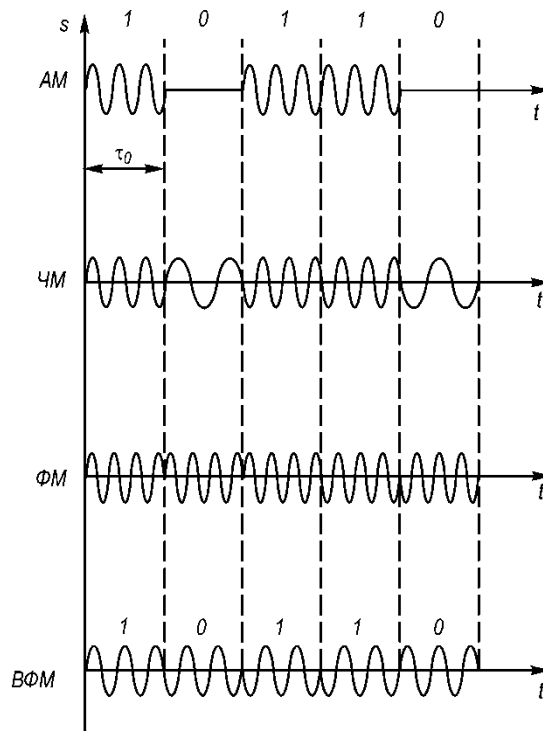


Рисунок 1.12 – Сигнали при різних видах дискретної модуляції [6]

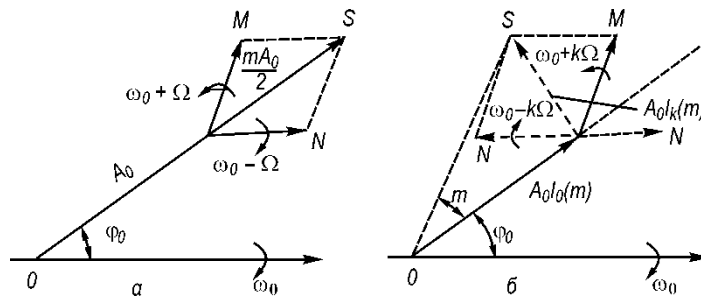


Рисунок 1.11 – Векторні діаграми АМ-сигналу (а) та ЧМ-сигналу ($m < 1$) (б)

Щоб довести це, розглянемо частотну модуляцію одним тоном $u(t) = \sin \Omega t$. Для цього випадку функція $\varphi(t)$ і її перетворення Гільберта набувають вигляду:

$$\varphi(t) = \Delta\omega_d \int u_m(t) dt = m \sin \Omega t, \quad \varphi^*(t) = -m \cos \Omega t,$$

де $m = \frac{\Delta\omega_d}{\Omega}$ – індекс модуляції. Модулююча функція в такому разі:

$$M[u_m(t)] = e^{i[\varphi(t) + i\varphi^*(t)]} = \sum_{k=0}^{\infty} \frac{m^k}{k!} e^{ik\Omega t},$$

а модулюючий сигнал:

$$s(t), \quad M[u_m(t)] e^{i\omega_0 t} = \sum_{k=0}^{\infty} \frac{m^k}{k!} e^{i(\omega_0 + k\Omega)t}$$

Звідси бачимо, що спектр модульованого сигналу складається з одної бічної смуги частот. Односмуговий ЧМ-сигнал можна отримати зі звичайного ФМ-сигналу шляхом перетворення Гільберта (наприклад, фазовим зсувом на $\pi/2$) і модуляції амплітуди за експоненціальним законом. Тоді обмеження такого сигналу в приймачі відновить нижню бічну смугу частот і дозволить застосувати для детектування звичайний дискримінатор.

1.4 Дискретна модуляція гармонічного переносника

Дискретна модуляція (маніпуляція) є окремим випадком модуляції гармонічного переносника, коли модулюючий сигнал $u_m(t)$ – дискретний. При дискретній модуляції закодоване повідомлення $u(t)$ послідовність кодових символів $\{a_i\}$ – перетворюється в послідовність елементів сигналу $\{s_i\}$. Останні відрізняються від кодових символів тільки електричним представленням. Дискретна модуляція може також полягати у взаємодії кодових символів $\{a_i\}$ з переносником $f(t)$. Така дискретна модуляція аналогічна неперервній.

У процесі модуляції один з параметрів переносника змінюється за законом, який визначається кодом. У разі безпосереднього передавання переносником може бути постійний струм, змінними параметрами якого є величина і напрямок. Звичайно за переносник, як і при неперервній модуляції, використовується змінний струм (гармонічне коливання). В цьому випадку можна отримати амплітудну (АМ), частотну (ЧМ) і фазову (ФМ) модуляції. Дискретну модуляцію часто називають *маніпуляцією*, а пристрій, що здійснює дискретну модуляцію (дискретний модулятор), називають *маніпулятором* або *генератором сигналів*.

На рисунку 1.12 наведені графіки сигналів при різних видах маніпуляції. При АМ символу 1 відповідає передавання несучого коливання впродовж часу τ_0 (посилка), а символу 0 – відсутність коливання (пауза). При ЧМ передавання несучого коливання з частотою f_1 відповідає символу 1, а передавання коливання з частотою f_0 відповідає 0. При ФМ змінюється фаза несучої на 180° на кожному переході від 1 до 0 і від 0 до 1.

Крім цих видів маніпуляції, застосовується *відносна фазова модуляція* (ВФМ). На відміну від ФМ, у системі ВФМ фаза несучого коливання змінюється на 180° при передаванні символів 1 і залишається незмінною при передаванні символів 0. Маніпуляція кожної даної послілки здійснюється відносно попередньої. Очевидно, таким чином можна змінювати будь-який параметр несучого коливання: змінюючи частоту, отримуємо *відносну частотну маніпуляцію* (ВЧМ), а змінюючи амплітуду, – *відносну амплітудну маніпуляцію* (ВАМ) [7].

Розглянемо спектри сигналів при деяких видах дискретної модуляції. Вважатимемо, що модуляція виконується двійковим повідомленням $u(t)$, що є періодичною послідовністю імпульсів з періодом $T = 2\tau_0$.

Амплітудна маніпуляція. АМ-сигнал можна записати у вигляді:

$$s(t) = A_0 u(t) \sin(\omega_0 t + \varphi_0), \quad (1.50)$$

де періодична функція $u(t)$ на інтервалі $T = 2\tau_0$ дорівнює:

$$u(t) = \begin{cases} 1 & \text{при } 0 < t < \tau_0, \\ 0 & \text{при } -\tau_0 < t < 0. \end{cases} \quad (1.51)$$

Подамо $u(t)$ через ряд Фур'є:

$$u(t) = \frac{1}{2} + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega t. \quad (1.52)$$

Тоді АМ-сигнал запишеться у вигляді

$$\begin{aligned} s(t) &= A_0 \left(\frac{1}{2} + \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega t \right) \sin(\omega_0 t + \varphi_0) = \\ &= \frac{1}{2} A_0 \sin(\omega_0 t + \varphi_0) + \frac{A_0}{2\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \cos(\omega_0 t - k\Omega t + \varphi_0) - \\ &\quad - \frac{A_0}{2\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \cos(\omega_0 t + k\Omega t + \varphi_0). \end{aligned} \quad (1.53)$$

Спектр АМ-сигналу, побудований за формулою (1.53), показано на рисунку 1.13, а. Він складається з несучого коливання з амплітудою $\frac{1}{2}A_0$ і двох бічних смуг, спектральні складові яких мають амплітуди:

$$A_k = \frac{A_0}{2\pi} \frac{1 - \cos k\pi}{k} = \frac{A_0}{k\pi} \sin \frac{k\pi}{2}. \quad (1.54)$$

Обвідна спектра дискретного АМ-сигналу виражається формулою:

$$A(\omega) = \frac{A_0 \tau_0}{2} \frac{\sin \frac{1}{2}(\omega - \omega_0)\tau_0}{\frac{1}{2}(\omega - \omega_0)\tau_0}, \quad (1.55)$$

тобто спектр є зсунутим на частоту ω_0 спектром одиночного імпульсного сигналу $u(t)$.

Фазова маніпуляція. ФМ-сигнал можна записати у вигляді:

$$\begin{aligned} s(t) &= A_0 \sin[\omega_0 t + \Delta\varphi u(t) + \varphi_0] = \\ &= A_0 \cos[\Delta\varphi u(t)] \sin(\omega_0 t + \varphi_0) + A_0 \sin[\Delta\varphi u(t)] \cos(\omega_0 t + \varphi_0). \end{aligned} \quad (1.56)$$

Періодична функція, яка визначає закон змінення фази на інтервалі $T = 2\tau_0$, виражається формулою:

$$u(t) = \begin{cases} 1 & \text{при } 0 < t < \tau_0, \\ -1 & \text{при } -\tau_0 < t < 0. \end{cases} \quad (1.57)$$

Підстановка (1.57) у вираз (1.56) дає:

$$s(t) = A_0 \cos \Delta\varphi \sin(\omega_0 t + \varphi_0) + u(t) \sin \Delta\varphi \cos(\omega_0 t + \varphi_0).$$

Подано $u(t)$ через ряд Фур'є:

$$u(t) = \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega t.$$

Тоді ФМ-сигнал запишеться у вигляді:

$$\begin{aligned}
s(t) &= A_0 \cos \Delta\varphi \sin(\omega_0 t + \varphi_0) + \frac{2A_0}{\pi} \sin \Delta\varphi \cos(\omega_0 t + \varphi_0) \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin k\Omega t = \\
&= A_0 \cos \Delta\varphi \sin(\omega_0 t + \varphi_0) - \frac{A_0}{\pi} \sin \Delta\varphi \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin(\omega_0 t - k\Omega t + \varphi_0) + \\
&\quad + \frac{A_0}{2} \sin \Delta\varphi \sum_{k=1}^{\infty} \frac{1 - \cos k\pi}{k} \sin(\omega_0 t - k\Omega t + \varphi_0).
\end{aligned} \tag{1.58}$$

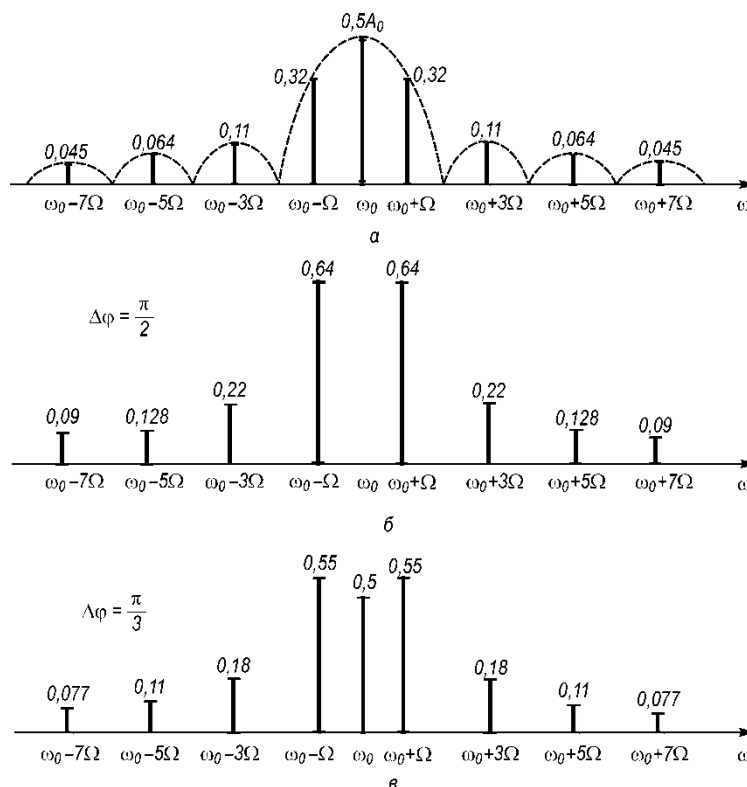


Рисунок 1.13 – Спектр сигналу при амплітудній (а) і фазовій (б, в) маніпуляціях [7]

Спектр ФМ-сигналу для різноманітних значень девіацій фази $\Delta\varphi$, побудований за формулою (1.58), наведений на рисунку 1.13, б, в. Він складається з несучого коливання і двох бічних смуг. Амплітуда несучого коливання залежить від $\Delta\varphi$: $A(\omega_0) = A_0 \cos \Delta\varphi$ і при $\Delta\varphi = \pm \frac{\pi}{2}$ дорівнює нулю. Амплітуди спектральних складових у бічних смугах також залежать від $\Delta\varphi$. Із збільшенням $\Delta\varphi$ від 0 до $\pi/2$ (див. рисунок 1.13, б, в) амплітуда несучого коливання зменшується до нуля, а амплітуди бічних частот збільшуються.

Коли $\Delta\varphi = \frac{\pi}{2}$, вся енергія ФМ-сигналу зосереджується тільки в бічних смугах. Як і при АМ, обвідна дискретного спектра бічних частот є зсунутим на частоту ω_0 спектром одиночного імпульсного сигналу $u(t)$, помноженим на $\sin \Delta\varphi$:

$$A(\omega) = A_0 \tau_0 \sin \Delta\varphi \frac{\sin \frac{1}{2}(\omega - \omega_0) \tau_0}{\frac{1}{2}(\omega - \omega_0) \tau_0}. \tag{1.59}$$

Аналогічно визначається спектр сигналу при частотній маніпуляції.

1.5 Імпульсна модуляція

У системах зв'язку з імпульсною модуляцією переносником інформації служить періодична послідовність імпульсів однакової форми (рис. 1.14, а):

$$f(t) = \sum_{k=-\infty}^{\infty} A_0 u(t - t_k), \quad (1.60)$$

де $u(t)$ – нормована функція, яка характеризує форму імпульсу; A_0 – амплітуда імпульсу; $t_k = kT_i + t_0$ – початок переднього фронту k -го імпульсу; T_i – період прямування імпульсів; t_0 – початок відліків послідовності.

При модуляції один з параметрів послідовності змінюється відповідно до модульованого сигналу. Так, при *амплітудно-імпульсній модуляції* (АІМ) змінюється амплітуда імпульсу (рис. 1.14, в):

$$A(t) = A_0 + \Delta A u(t). \quad (1.61)$$

При *широтно-імпульсній модуляції* (ШІМ) змінюється тривалість імпульсу (рис. 1.14, з):

$$\tau(t) = \tau_0 + 2\Delta\tau_m u(t), \quad (1.62)$$

де $\Delta\tau_m$ – максимальне відхилення фронту імпульсів в один бік.

При *фазово-імпульсній модуляції* (ФІМ) змінюється зсув імпульсів відносно тактових точок kT_i (рис. 1.14, е):

$$t_k = \theta(t) = kT_i + \Delta\tau_m u(t). \quad (1.63)$$

При *частотно-імпульсній модуляції* (ЧІМ) відповідно до повідомлення, що передається, змінюється час-тота прямування імпульсів (рис. 1.14, д).

Як при ФІМ, імпульси зміщуються відносно тактових точок, але за іншою закономірністю. Різниця ФІМ і ЧІМ аналогічна різниці між ФМ і ЧМ синусоїдального переносника.

Періодичну послідовність прямокутних імпульсів можна записати у вигляді:

$$f(t) = \sum_{k=-\infty}^{\infty} A_0 u(t - t_k) = \begin{cases} A_0 & \text{при } \theta_1 + kT_i < t < \theta_2 + kT_i, \\ 0 & \text{при } \theta_2 + kT_i < t < \theta_1 + (k+1)T_i. \end{cases} \quad (1.64)$$

Її можна подати рядом Фур'є:

$$f(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \dot{\cdot} \quad ,$$

$$\text{де } \int_{-T/2}^{T/2} f(t) e^{-i\omega_1 t} dt, \quad \omega_1 = \frac{2\pi}{T}.$$

У нашому випадку:

$$\int_{-T_i/2}^{T_i/2} e^{-ik\omega_i t} dt = \frac{2A_0\tau_0 \left(e^{-i\frac{k\omega_i\tau_0}{2}} - e^{i\frac{k\omega_i\tau_0}{2}} \right)}{ik\omega_i T_i} = \frac{2A_0\tau_0}{T_i} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}}; \quad (1.65)$$

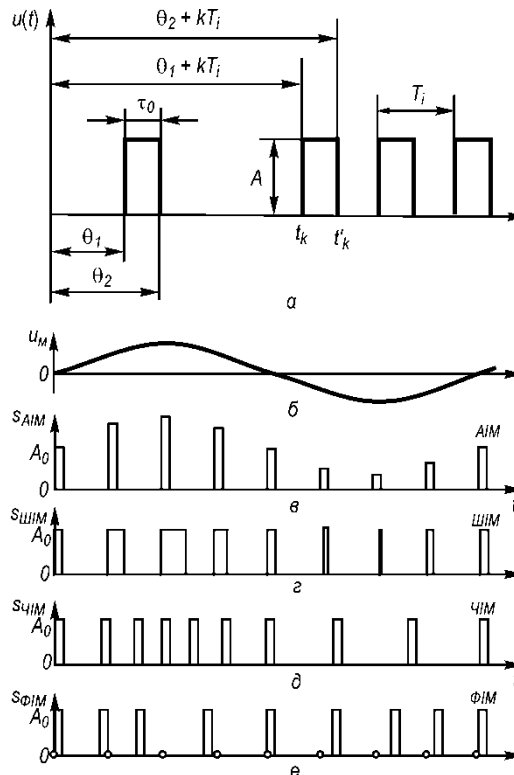


Рисунок 1.14 – Параметри періодичної послідовності прямокутних імпульсів (а), модулюючий сигнал (б) та модульовані сигнали при різних видах імпульсної модуляції (в–е) [6]

Тоді:

$$f(t) = \frac{A_0\tau_0}{T_i} \sum_{k=-\infty}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} e^{ik\omega_i(t-t_0)} = \frac{A_0\tau_0}{T_i} \left[1 + 2 \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i\tau_0}{2}}{\frac{k\omega_i\tau_0}{2}} \cos k\omega_i(t-t_0) \right], \quad (1.66)$$

$$\text{де } \tau_0 = \theta_2 - \theta_1; \quad t_0 = \frac{\theta_1 + \theta_2}{2}; \quad \omega_i = \frac{2\pi}{T_i}.$$

Спектр амплітуд періодичної послідовності прямокутних імпульсів наведений на рисунку 1.15. Амплітуди спектральних компонентів A_k визначаються модулями спектральної щільності $|A_k|$ (1.65) на гармоніках частоти повторення $\omega_i = \frac{2\pi}{T_i}$. Форма обвідної частотного спектра періодичної послідовності залежить від форми

окремого імпульсу. Зі збільшенням періоду повторення інтервал частот між сусідніми спектральними компонентами скорочується, їх число збільшується, а амплітуда кожної компоненти зменшується із збереженням сталого співвідношення між ними. У разі необмеженого збільшення T_i періодична послідовність вироджується в одиничний імпульс, а лінійний спектр стає суцільним.

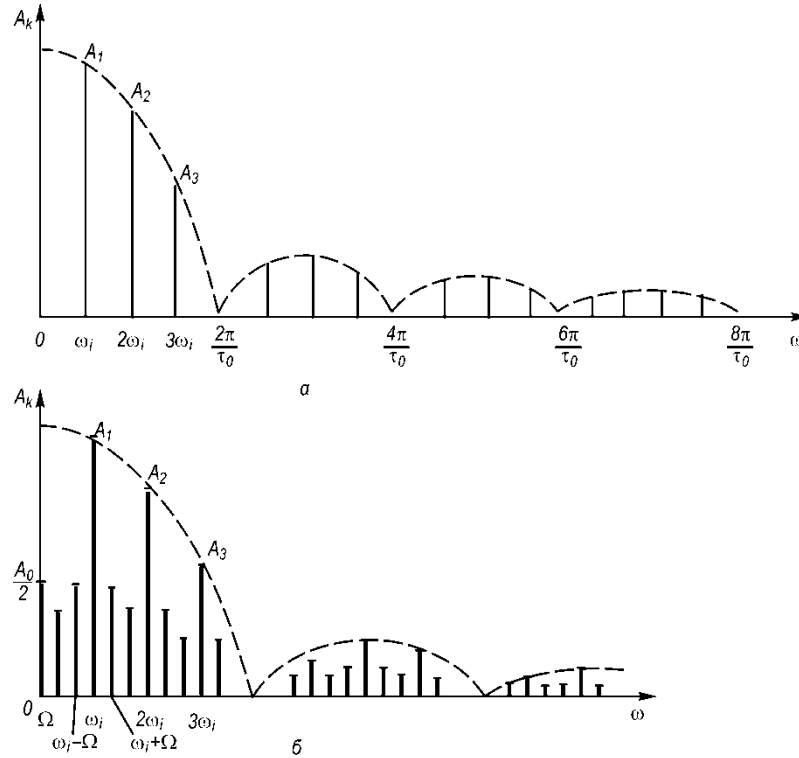


Рисунок 1.15 – Спектр періодичної послідовності прямокутних імпульсів (а) та спектр АІМ-сигналу (б) [8]

Спектр періодичної послідовності радіоімпульсів отримується зі спектра послідовності відеоімпульсів перенесенням шкали частот на несучу частоту ω_0 і доповненням отриманого спектра його дзеркальним відображенням.

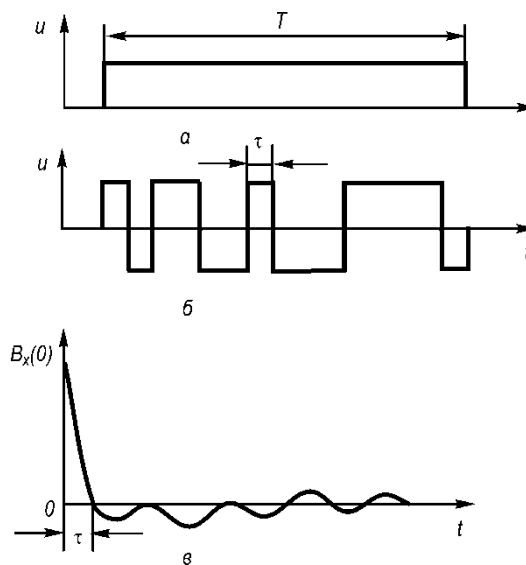


Рисунок 1.16 – Принцип побудови складного широкосмугового сигналу [8]

При модуляції параметри, які входять у вирази (1.64) і (1.66), є функціями часу: $A_0 = A(t)$; $\theta_1 = \theta_1(t)$; $\theta_2 = \theta_2(t)$. Модульована послідовність буде тепер уже неперіодичною функцією, яка деформована відносно початкової:

$$s(t) = \sum_{k=-\infty}^{\infty} A(t)u(t-t_k) = \begin{cases} u(t) & \text{при } \theta_1(t) + kT_i < t < \theta_2(t) + kT_i, \\ 0 & \text{при } \theta_2(t) + kT_i < t < \theta_1(t) + (k+1)T_i, \end{cases} \quad (1.67)$$

Або

$$s(t) = u(t) \frac{\theta_2(t) - \theta_1(t)}{T_i} = 1 + 4 \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i}{2} [\theta_2(t) - \theta_1(t)]}{k\omega_i [\theta_2(t) - \theta_1(t)]} \cos k\omega_i \left[t - \frac{\theta_1(t) + \theta_2(t)}{2} \right]. \quad (1.68)$$

Остання формула визначає частотний спектр деформованої послідовності імпульсів. Для отримання спектрів каналів при різних видах модуляції в формулу (1.68) необхідно підставити відповідний вираз для модульованого сигналу.

Для прикладу знайдемо спектр при АІМ. У разі модуляції одним тоном $u(t) = \sin \Omega t$ і $A = A_0(1 + m \sin \Omega t)$; інші параметри послідовності незмінні: $\theta_2 - \theta_1 = \tau_0$, $\theta_1 = 0$ і $\theta_2 = \tau_0$. Після підстановки цих значень для частотного спектра АІМ-сигналу дістанемо:

$$s(t) = \frac{A_0 \tau_0}{T_i} \left[1 + 2 \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i \tau_0}{2}}{\frac{k\omega_i \tau_0}{2}} \right] \cos k\omega_i \left(t - \frac{\tau_0}{2} \right) + \frac{mA_0 \tau_0}{T_i} \sin \Omega t + \frac{mA_0 \tau_0}{T_i} \times \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i \tau_0}{2}}{\frac{k\omega_i \tau_0}{2}} \sin \left[(k\omega_i + \Omega)t - \frac{k\omega_i \tau_0}{2} \right] + \frac{mA_0 \tau_0}{T_i} \sum_{k=1}^{\infty} \frac{\sin \frac{k\omega_i \tau_0}{2}}{\frac{k\omega_i \tau_0}{2}} \sin \left[(k\omega_i + \Omega)t - \frac{k\omega_i \tau_0}{2} \right]. \quad (1.69)$$

На рисунку 1.16 наведено графік спектра АІМ-сигналу. Порівняння його з рисунком 1.15 показує, що при АІМ модулюється за амплітудою кожна складова спектра немодульованої послідовності імпульсів, як ізольована несуча. В спектрі є низькочастотне модулююче повідомлення $u(t)$ з частотою Ω ; отже, демодуляція при АІМ може бути здійснена за допомогою фільтра нижніх частот, який пропускає низькочастотні коливання $u(t)$.

Аналогічно визначається спектр й інших видів імпульсної модуляції. Для обчислення спектра при ФІМ в (1.68) слід підставити вираз (1.63), яким визначається зміна положення імпульсу відповідно до повідомлення, що передається, а при ШІМ – вираз (1.62), за яким визначається зміна тривалості імпульсу.

При імпульсно-кодовій модуляції (ІКМ) передавання окремих значень сигналу зводиться до передавання окремих груп імпульсів. Ці групи передаються одна за одною через відносно великі проміжки часу порівняно з тривалістю окремих

імпульсів. Кожна кодова група імпульсів є регулярним неперіодичним сигналом, спектр якого може бути обчислений на підставі перетворень Фур'є.

Ширина спектра послідовності імпульсів практично не залежить від частоти повторення ω_i і визначається, головним чином, шириною спектра одного імпульсу. У разі наявності модуляції будь-якого виду спектр розширюється незначно внаслідок бічних частот крайніх складових спектра немодульованих імпульсів. Тому робоча смуга частот, яка зайнята імпульсними сигналами, практично залежить не від виду модуляції, а від тривалості та форми імпульсу.

1.6 Модуляція шумової несучої

За переносники можна використовувати не тільки періодичні коливання, але і вузькосмуговий випадковий процес. Такі переносники знаходять практичне застосування, наприклад, в оптичних системах зв'язку з некогерентним випромінюванням, де сигналом є вузькосмуговий гауссів шум.

Вузькосмуговий випадковий процес можна зобразити як квазігармонічне коливання:

$$f(t) = u(t) \cos[\omega_0 t + \varphi(t)] = u(t) \cos \psi(t), \quad (1.70)$$

обвідна якого повільно змінюється $u(t)$ і фазою $\psi(t)$. При амплітудній модуляції відповідно до повідомлення, що передається, змінюється обвідна $u(t)$, при фазовій модуляції – фаза $\varphi(t)$, при частотній – миттєва частота $\omega(t) = \frac{d\psi}{dt}$.

Розглянемо амплітудну модуляцію шумової несучої. Вираз для модульованої несучої можна записати у вигляді:

$$y(t) = [1 + m u(t)] f(t) \quad (1.71)$$

де $f(t)$ – переносник; $u(t)$ – модулююча функція (відеосигнал); m – коефіцієнт модуляції.

Припускають, що модулюючий процес $u(t)$ є стаціонарним нормальним процесом з середнім значенням, яке дорівнює нулю: $u(t) = 0$. Процеси $f(t)$ і $u(t)$ незалежні. За цих умов функція кореляції $B_y(\tau)$ модульованої за амплітудою шумової несучої буде:

$$\begin{aligned} B_y(\tau) &= \overline{y(t)y(1+\tau)} = \overline{[1 + m u(t)] f(t) [1 + m u(1+\tau)] f(1+\tau)} = \\ &= \overline{[1 + m u(t) + m u(1+\tau) + m^2 u(t)u(1+\tau)] f(t) f(1+\tau)} = \\ &= [1 + m^2 B_u(\tau)] B_f(\tau), \end{aligned} \quad (1.72)$$

де $B_u(\tau)$ – функція кореляції, що характеризує модулюючу функцію; B_f – функція кореляції, що характеризує сигнал-переносник.

Тепер знаходимо енергетичний спектр:

$$G_y(\omega) = 2 \int_0^{\infty} B_y(\tau) \cos \omega \tau d\tau = 2 \int_0^{\infty} B_f(\tau) \cos \omega \tau d\tau + 2m^2 \int_0^{\infty} B_u(\tau) \cos \omega \tau d\tau. \quad (1.73)$$

Перший інтеграл дає енергетичний спектр шумової несучої $G_j(\omega)$. Для другого інтеграла на основі теореми про спектр добутків маємо:

$$\int_0^{\infty} B_u(\tau) B_f(\tau) \cos \omega \tau d\tau = \frac{1}{2} \int_{-\infty}^{\infty} G_f(\nu) G_u(\omega - \nu) d\nu.$$

Остаточно спектр модульованої несучої дорівнює:

$$G_y(\omega) = G_j(\omega) + m^2 \int_{-\infty}^{\infty} G_f(\tau - \omega) G_u(\omega - \nu) d\nu.$$

Таким чином, спектр, модульований за амплітудою шумової несучої, отримують суперпозицією спектра несучої і згорткою цього спектра зі спектром повідомлення, яке передається, зсунутого в область високих частот на величину ω_0 . Аналогічно визначаються функція кореляції та енергетичний спектр при ФМ і ЧМ.

Застосування «шумових» сигналів дозволяє послабити вплив завмирання в каналах з багатопробним розповсюдженням радіохвиль. Пояснимо це на простому прикладі. Нехай на вхід приймача надходять сигнали двох променів $\xi(t)$ і $\xi(t - \tau)$ зі зсувом на час τ . Потужність результуючого сигналу, яка визначається за достатньо великий час T ,

$$P = \overline{[\xi(t) + \xi(t - \tau)]^2} = 2[P_0 + B_\xi(\tau)],$$

де $B_\xi(\tau)$ – функція кореляції сигналу; P_0 – його середня потужність. Функція кореляції шуму зменшується зі збільшенням τ і тим швидше, чим ширший його спектр. Відповідно при достатньо великій ширині спектра можна вважати, що $B_\xi(\tau) \approx 0$ і $P \approx 2P_0$, тобто середня потужність прийнятого сигналу, незважаючи на завмирання, залишається приблизно постійною.

Застосування в якості переносника реалізацій *реального шуму* пов'язано зі складностями, що виникають при формуванні і прийманні таких сигналів. Тому на практиці знайшли застосування шумоподібні сигнали, які не є випадковими; вони формуються за певним алгоритмом. Однак їх статистичні властивості близькі до властивостей шуму: енергетичний спектр майже рівномірний, а функція кореляції має вузький основний пік і невеликі бічні викиди [7].

У теперішній час відомі методи формування шумоподібних сигналів, які у разі великої бази ($2TF$) дозволяють незалежно відтворювати їх на приймальному і передавальному кінцях і відповідають вимогам синхронізації цих сигналів.

Широко застосовуються *дискретні сигнали*, що будуються на розбитті інформаційної послідовності тривалістю T на N бінарних елементів тривалістю $\tau = \frac{T}{N}$.

Таке розбиття дозволяє отримати сигнал тривалістю T зі смугою $F \approx \frac{1}{\tau}$ і значенням бази $2TF \square$.

Послідовності бінарних елементів утворюють коди, які вибирають так, щоб забезпечити задані властивості сигналу. За допомогою модуляції або гетеродинування формується високочастотний сигнал, який передається по каналу. Часто при цьому використовується модуляція фази на два положення: 0 і π .

Функція кореляції дискретних сигналів при достатньо великому числі елементів N має головний максимум, зосереджений в області $-\tau \leq \Delta\tau \leq \tau$, і бічні пелюстки, які мають порівняно малий рівень (рис. 1.16). Ця функція дуже нагадує функцію автокореляції відрізка шуму зі смугою F .

У системах зв'язку, де використовуються шумоподібні (складові) сигнали, кожний елемент повідомлення передається не одним, а кількома елементами сигналу, що несуть (повторюють) одну й ту саму інформацію. Число N може сягати кількох сотень і навіть тисяч. Як буде показано в подальшому, це дає змогу реалізувати накопичення сигналу і забезпечити високу завадостійкість навіть тоді, коли рівень сигналу нижчий за рівень завад [8].

Деякі дискретні сигнали будуються на основі *лінійних рекурентних послідовностей*. Ці сигнали мають добрі кореляційні властивості та порівняно нескладну практичну реалізацію. Структура сигналів носить випадковий характер, хоча спосіб їх формування цілком регулярний. Неперервні ФМ-сигнали, побудовані на рекурентних послідовностях, можуть мати ідеальну автокореляційну функцію.

Серед лінійних рекурентних послідовностей особливе місце займають *псевдовипадкові M-послідовності Хаффмена*. Вони являють собою сукупність N періодично повторюваних символів d_i , кожний з яких може набувати одного з двох значень: +1 або -1, що визначається взятим з протилежним знаком добутком двох або більшого числа (але завжди парного) попередніх сигналів:

$$d_i = -d_{i-n}d_{i-m} \dots d_{i-e}d_{i-k}, \text{ причому } n > m > \dots > e > k \geq 1; i = (n+1) \div N. \quad (1.74)$$

Якщо вибрати початкову послідовність d_1, d_2, \dots, d_n , то на підставі (1.74) можна утворити неповторну елементарну послідовність $\{d_i\}$ із N символів, тут:

$$N = 2^n - 1. \quad (1.75)$$

Вона вміщуватиме всі комбінації n символів з двох елементів: +1 і -1, окрім комбінацій, які складаються з одних від'ємних одиниць. Внаслідок цього кожна послідовність $\{d_i\}$ вміщує 2^{n-1} додатних одиниць і $2^{n-1} - 1$ від'ємних одиниць. Тому:

$$\sum_{i=1}^N d_i = 1. \quad (1.76)$$

При $i > N$ символи повторюються в тому самому порядку, тобто при будь-якому цілому:

$$d_{i+pN} = d_i. \quad (1.77)$$

Розглянемо найпростіший приклад: $n=2$. У цьому випадку $d_i = d_{i-2}d_{i-1}$. Виберемо початкову послідовність $d_1 = -1$, $d_2 = +1$. Тоді шукана послідовність буде такою: ...-1, +1; +1, -1; +1, +1; -1, +1; ... Вона складає всі можливі комбінації з двох символів: -1, +1; +1, +1; +1, -1, окрім забороненої комбінації -1, -1. Отримана послідовність -1, +1, +1 повторюється через $N=2^2-1=3$ символи. При $n=3$ можна утворити дві послідовності з $N=2^3-1=7$ символів за правилами $d_i = -d_{i-3}d_{i-2}$ і $d_i = -d_{i-3}d_{i-1}$. Майже кожному цілому числу n відповідає кілька чисел k , при яких за правилом (1.74) утворюється послідовність.

З виразу (1.77) випливає, що число N є максимальним періодом нескінченної послідовності Хаффмена. Можуть утворитися також послідовності меншого періоду. Максимальне число різних послідовностей максимального періоду для будь-якого n дорівнює:

$$M = \frac{1}{2} \varphi(2^n - 1), \quad (1.78)$$

де $\varphi(x)$ – функція Ейлера [функція Ейлера $\varphi(x)$ дорівнює кількості цілих чисел, включаючи одиницю, менших за число x і взаємно простих з x . Наприклад, якщо x дорівнює простому числу, то $\varphi(q) = q - 1$].

Бінарні псевдовипадкові послідовності Хаффмена мають ряд чудових властивостей. Нормована функція автокореляції в неперервному режимі роботи має головний максимум, який дорівнює одиниці, і однакові за величиною бічні пелюстки, які дорівнюють $-1/N$. Функція взаємної кореляції для різних послідовностей дорівнює $-1/M$. В імпульсному режимі роботи рівень бічних пелюсток не перевищує $1/\sqrt{N}$. Будь-які послідовності при заданому n відрізняються як порядком чергування символів +1 і -1, так і максимальним значенням бічних пелюсток. При цьому можна вказати послідовність, максимальний рівень бічних пелюсток якої буде найменшим серед можливих послідовностей для заданого n . Генерування псевдовипадкових послідовностей Хаффмена порівняно просто здійснюється за допомогою регістрів зсуву.

Крім сигналів Хаффмена, практичне застосування знаходять й інші види дискретних сигналів: сигнали Пелі–Плоткіна, послідовність символів Лежандра, коди Баркера, багатофазні коди Френка. Можливі, нарешті, різні варіанти складних сигналів.

У радіолокації широко застосовуються сигнали з *лінійними змінами частоти усередині імпульсу* (ЛЧМ). Пояснюється це тим, що ЛЧМ-сигнали мають добрі кореляційні властивості і прийняти їх можна за допомогою погоджених фільтрів.

Шумоподібний сигнал може піддаватися модуляції всіма відомими способами. При амплітудній модуляції змінюються амплітуди всіх елементів. У разі частотної модуляції варіанти сигналу відрізняються середньою частотою, а фазової – різницею фаз між елементами двох посилок.

Специфічним видом модуляції, яка властива тільки широкосмуговим системам зв'язку, є *структурна модуляція* або *модуляція за формою* сигналу. В цьому випадку за варіанти сигналу використовуються коливання, побудовані на однакових елементах, але по-різному розташованих. Наприклад, двійкове передавання можна здійснити за допомогою сигналів виду:

$$s_1(t) = f_1(t) \cos \omega_0 t, \quad s_2(t) = f_2(t) \cos \omega_0 t, \quad (1.79)$$

де $f_1(t)$ і $f_2(t)$ – функції, які набувають значень ± 1 відповідно до заданої кодової послідовності і задовольняють умову:

$$\int_0^T f_1(t) f_2(t) dt \approx 0. \quad (1.80)$$

У разі виконання останньої умови маємо систему майже ортогональних сигналів. Як функції $f_1(t)$ і $f_2(t)$, що задовольняють умову (1.80), можна використовувати розглянуту вище двійкову псевдовипадкову послідовність Хаффмена $f_1(t) = f(t)$ і цю ж послідовність зі зсувом у часі: $f_2(t) = f(t + \Delta t)$.

Аналогічно будуються *багатопозиційні* широкосмугові системи зі структурною модуляцією. В цьому випадку застосовується ансамбль шумоподібних сигналів $f_i(t)$ ($i = 1, 2, \dots, m$). При цьому, обов'язково, різниця між сигналами має бути достатньою для розділення на прийомі. З цієї точки зору великий інтерес становлять протилежні й ортогональні сигнали.

1.7 Енергетичний спектр модульованих сигналів

До цих пір ми розглядали модуляцію несучого коливання детермінованим процесом $u(t)$, який відображає певне повідомлення або окрему його реалізацію. Сукупність можливих повідомлень є деяким випадковим процесом. Так, при передаванні мови або музики статистичні властивості повідомлень, що передаються, дуже близькі до властивостей нормального випадкового процесу. Важливими характеристиками коливання, модульованого випадковим процесом, є функція кореляції та енергетичний спектр.

Слід підкреслити, що модульований сигнал є нестационарним випадковим процесом навіть тоді, коли модулюючі процеси (повідомлення) стаціонарні. Енергетичний спектр нестационарного випадкового процесу визначається за допомогою двократного усереднення – за множиною і за часом. Спочатку визначається усереднена за часом кореляційна функція, а потім оберненим перетворенням Фур'є – шуканий енергетичний спектр [9].

Розглянемо випадок, коли передаване повідомлення $u(t)$ являє собою стаціонарний процес з середнім значенням $\overline{u(t)}$, а переносник – гармонічні коливання $A_0 \cos \omega_0 t$.

При амплітудній модуляції:

$$s(t) = A_0 [1 + m u(t)] \cos \omega_0 t,$$

де m – середньоквадратичне значення коефіцієнта модуляції.

Функція кореляції модульованого сигналу:

$$\begin{aligned} B(t, \tau) &= \overline{s(t) s(t + \tau)} = A_0^2 \overline{[1 + m u(t)] [1 + m u(t + \tau)] \cos \omega_0 t \cos(\omega_0 t + \tau)} = \\ &= 0,5 A_0^2 [1 + m^2 B_u(\tau)] [\cos(2\omega_0 t + \tau) + \cos \omega_0 \tau], \end{aligned} \quad (1.81)$$

де $B_u(t)$ – функція кореляції повідомлення $u(t)$, що передається.

Як бачимо, функція $B(t, \tau)$ залежить від часу, що вказує на нестационарність модульованого сигналу. Після усереднення за часом отримаємо:

$$B(\tau) = \overline{B(t, \tau)} = \frac{1}{2} A_0^2 [1 + m^2 B_u(\tau)] \cos \omega_0 \tau.$$

Застосовуючи до $B(\tau)$ перетворення Фур'є, знаходимо енергетичний спектр сигналу при АМ:

$$\begin{aligned} G(\omega) &= 2 \int_0^{\infty} B(\tau) \cos \omega \tau d\tau = A_0^2 \int_0^{\infty} [1 + m^2 B_u(\tau)] \cos \omega_0 \tau \cos \omega \tau d\tau = \\ &= \frac{\pi A_0^2}{2} \delta(\omega - \omega_0) + \frac{m^2 A_0^2}{4} G_u(\omega - \omega_0). \end{aligned}$$

Таким чином, спектр гармонічного коливання, модульованого за амплітудою випадковим процесом, складається з несучого коливання з частотою ω_0 і зсунутого на ω_0 спектра повідомлення $u(t)$, яке передається.

Сигнал при кутовій модуляції (ЧМ і ФМ) можна записати в загальному вигляді:

$$s(t) = A_0 \cos[\omega_0 t + \varphi(t)].$$

При ФМ $\varphi(t) = \Delta\varphi u(t)$, а при ЧМ $\varphi(t) = \Delta\omega \int u(t) dt$. Тут $\Delta\varphi$ і $\Delta\omega$ – середньоквадратичні значення девіації відповідно фази і частоти.

Функція кореляції модульованого сигналу:

$$\begin{aligned} B(t, \tau) &= \overline{s(t)s(t+\tau)} = A_0^2 \overline{\cos[\omega_0 t + \varphi(t)] \cos[\omega_0(t+\tau) + \varphi(t+\tau)]} = \\ &= \frac{A_0^2}{2} \left\{ \overline{\cos[2\omega_0 t + \omega_0 \tau + \varphi(t+\tau)]} + \overline{\cos[\omega_0 \tau + \varphi(t+\tau) - \varphi(t)]} \right\}. \end{aligned}$$

При усередненні за часом перший доданок обертається в нуль. Другий доданок не залежить від часу t , тому:

$$B(\tau) = \overline{B(t, \tau)} = \frac{A_0^2}{2} \overline{\cos[\omega_0 \tau + \varphi(t+\tau) - \varphi(t)]}.$$

Позначимо різницю $\varphi(t+\tau) - \varphi(t) = \theta(\tau)$ і за відомою формулою подамо косинус суми двох кутів у вигляді:

$$\overline{\cos(\omega_0 \tau + \theta)} = \overline{\cos \omega_0 \tau \cos \theta} - \overline{\sin \omega_0 \tau \sin \theta}.$$

Середні за множиною значення косинуса і синуса θ можна знайти, якщо відомий закон розподілу ймовірностей повідомлень $u(t)$. Якщо $u(t)$ підпорядковується нормальному закону, то $\theta(\tau)$, яке є лінійним перетворенням від $u(t)$, також матиме нормальний розподіл з нульовим середнім значенням і дисперсією σ_θ^2 . Легко переконатися, що в цьому випадку:

$$\overline{\cos \theta} = e^{-\frac{\sigma_\theta^2}{2}}; \quad \overline{\sin \theta} = 0.$$

Отже, усереднена за часом функція кореляції сигналу при кутовій модуляції:

$$B(\tau) = \frac{A_0^2}{2} \cos \omega_0 \tau e^{-\frac{\sigma_\theta^2}{2}}. \quad (1.82)$$

Дисперсію процесу $\theta(\tau)$ можна виразити через функцію кореляції або енергетичний спектр повідомлення $u(t)$. Дійсно,

$$\sigma_{\varphi}^2 = \overline{[\varphi(t+\tau) - \varphi(t)]^2} = 2[B_{\varphi}(0) - B_{\varphi}(t)],$$

де $B_{\varphi}(t)$ – функція кореляції процесу $\varphi(t)$. При ФМ $\varphi(t) = \Delta\varphi u(t)$, тому $B_{\varphi}(\tau) = \Delta\varphi^2 B_u(\tau)$ при ЧМ $\varphi(t) = \Delta\omega\psi$, де $\psi = \int u(t)dt$, тому $B_{\varphi}(\tau) = \Delta\omega^2 B_{\psi}(\tau)$. Потім можна визначити енергетичний спектр модульованого сигналу шляхом перетворення Фур'є від функції $B(\tau)$ (2.82).

1.8 Цифрові методи модуляції

Для передавання цифрового первинного сигналу по каналу використовують різні несучі $f(t)$. Нижче ми розглянемо тільки гармонічну несучу.

Цифрова амплітудна модуляція (ЦАМ). Канальний сигнал при ЦАМ (лінійній модуляції) можна записати як [8]:

$$u_{\text{ЦАМ}}(t) = \left(U_0 + K_{\text{АМ}} \sum_{n=0}^{\infty} b_n^{(k)} v(t-nT) \right) \cos(\omega_0 t + \varphi_0), \quad (1.83)$$

де $K_{\text{АМ}}$ – девіація амплітуди; U_0 – стала складова.

Спектр цього сигналу містить несучу і дві бічні смуги, кожна з яких повторює спектр первинного сигналу $\hat{b}_n(t)$. Формування сигналу та його детектування можна здійснити параметричними і нелінійними схемами. Проте з отриманої оцінки первинного сигналу $\hat{b}_n(t)$ потрібно вилучити на окремих тактових інтервалах T оцінки кодових символів \hat{b}_n . Для цього треба мати розв'язувальний пристрій, який знаходить ці оцінки. На рисунку 1.17 для окремих блоків введені позначення: АД – амплітудний детектор (когерентний або некогерентний), РП – розв'язувальний пристрій, який з тактовим інтервалом T приймає рішення про кодові символи $b_n^{(k)}$. Ці рішення фіксуються в запам'ятовуючому пристрої ЗП.

Зазначимо, що для приймання дискретних повідомлень слід забезпечити не тільки синхронізацію за частотою несучої ω_0 і за фазою сигналу (при когерентному детектуванні), але і тактову, а також циклічну (з урахуванням методів передавання дискретних повідомлень по реальних каналах) синхронізацію. Через наявність адитивного шуму в каналі приймання повідомлень за схемою рисунку 1.17 не є найкращим.

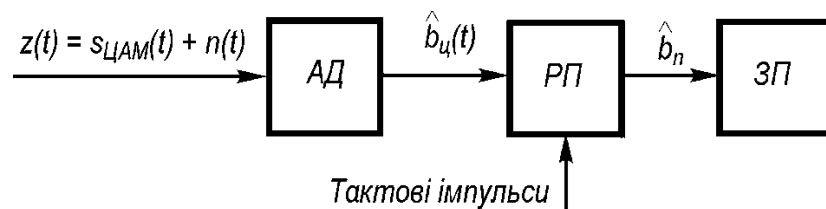


Рисунок 1.17 – Схема детектування сигналу ЦАМ і прийняття рішення про кодовий символ [9]

Якщо в (1.83) $U_0 = 0$, то маємо сигнал ЦАМ без несучої (ЦБАМ). В цьому випадку можливе або синхронне (когерентне) детектування, або некогерентне детектування з відновленням несучої в місці приймання. Методи ЦАМ дозволяють

за квадратурною схемою (рис. 1.18) організувати передавання і приймання двох незалежних дискретних повідомлень.

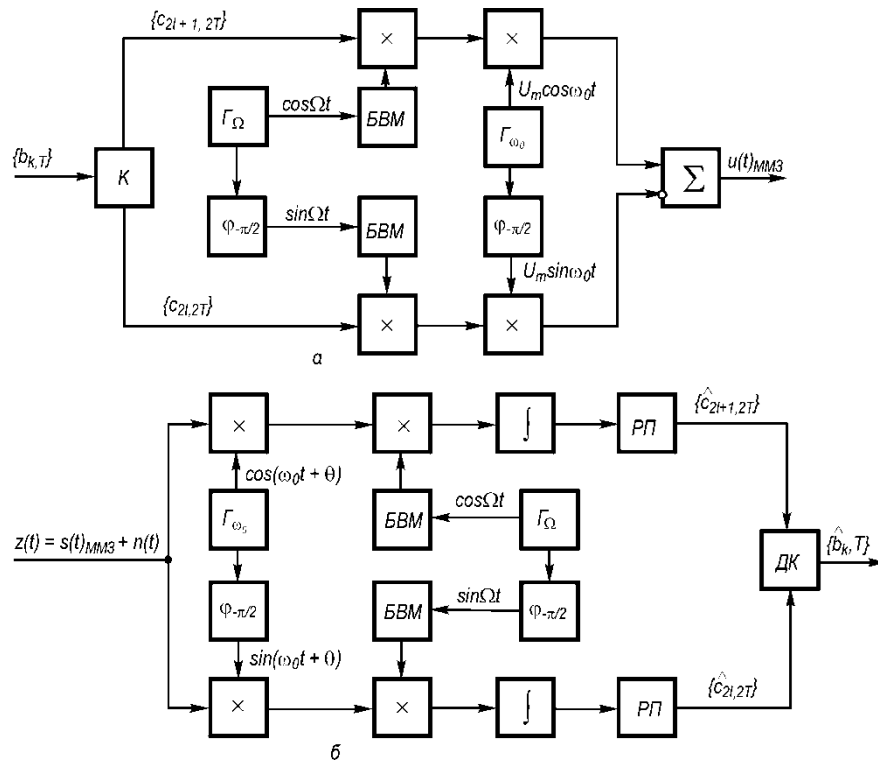


Рисунок 1.18 – Квадратурні схеми:

a – двійкового сигналу ММЗ; *б* – детектування сигналів ММЗ [10]

Сигнал, який передається в цифровій квадратурній амплітудній модуляції, тобто коли вектори сигналів розташовані в квадратній решітці (ЦКАМ),

$$u_{\text{ЦКАМ}}(t) = b_{11}(t) \cos(\omega_0 t + \varphi_0) + b_{12}(t) \sin(\omega_0 t + \varphi_0). \quad (1.84)$$

Приймання в цьому випадку може бути тільки когерентним. Якщо покласти в (1.84) $b_{12}(t) = \hat{b}_{11}(t)$, то можна за квадратурною схемою організувати і односмугову ЦАМ.

Цифрова фазова модуляція (ЦФМ). Канальний сигнал при ЦФМ можна записати у вигляді:

$$\begin{aligned} u_{\text{ЦФМ}}(t) &= U_m \cos \left(\omega_0 t + K_{\text{ФМ}} \sum_{n=0}^{\infty} b_n^{(k)} v(t - nT) \right) = \\ &= U_m \cos \left(K_{\text{ФМ}} \sum_{n=0}^{\infty} b_n^{(k)} v(t - nT) \right) \cos \omega_0 t - U_m \sin \left(K_{\text{ФМ}} \sum_{n=0}^{\infty} b_n^{(k)} v(t - nT) \right) \sin \omega_0 t, \end{aligned} \quad (1.85)$$

де U_m – амплітуда каналного сигналу; $K_{\text{ФМ}}$ – девіація фази.

Вид цього сигналу, реалізація модулятора і детектора суттєво спрощуються при використанні прямокутних імпульсів $v(t)$ з одиничною амплітудою і

тривалістю T . В цьому випадку відсутня міжсимвольна інтерференція (МСІ), і замість (1.85) при $b_n = \pm 1$ маємо:

$$u_{\text{ЦФМ}}(t) = U_m \cos\left(\frac{\Delta\theta}{2}\right) \cos \omega_0 t - U_m \sum_{n=0}^{\infty} b_n^{(k)} v(t-nT) \sin\left(\frac{\Delta\theta}{2}\right) \sin \omega_0 t, \quad (1.86)$$

де $\Delta\theta = 2K_{\text{ФМ}}$ – різниця фаз для двох позицій коду.

Спектральний склад сигналу (1.86) за змістом не відрізняється від сигналу (1.83). Відповідно, ЦФМ і когерентне детектування в розглянутому випадку реалізуються так само, як для ЦАМ, зокрема за квадратурною схемою (ЦКАМ). Зауважимо, що коли різниця фаз при двопозиційному коді $\Delta\theta = \pi$ (використовуються протилежні сигнали), то несуча в спектрі сигналу ЦФМ зникає.

Розглядають також методи відносної (різницевої) ЦФМ. У цьому випадку можна крім когерентного використовувати і некогерентний демодулятор. На практиці широко розповсюджені багатопозиційні системи ЦФМ, коли початкова фаза несучої набуває не двох, а m значень. У техніці передавання даних широко застосовуються як багатопозиційні ($m > 2$) системи ЦАМ (багаторівневі системи з лінійною модуляцією) та ЦФМ (з нелінійною модуляцією), так і їх змішані варіанти (ЦАФМ). Амплітудно-фазові діаграми деяких систем сигналів наведені на рисунку 1.19 (знаком «+» позначається початок координат). Для підвищення якості передавання (мінімізації середньої ймовірності помилкового приймання в каналі з шумом) намагаються підібрати таку сигнально-кодову конструкцію, щоб сигнальні точки дозволених кодових комбінацій знаходились одна від одної на максимально можливій відстані. Часто конструкція рисунку 1.19, б (чотиріпозиційна ЦФМ або система ФМ-4), яка має широке застосування в техніці зв'язку, реалізується двійковою ЦКАМ (з фізичної точки зору ЦКАМ є фазовою модуляцією).

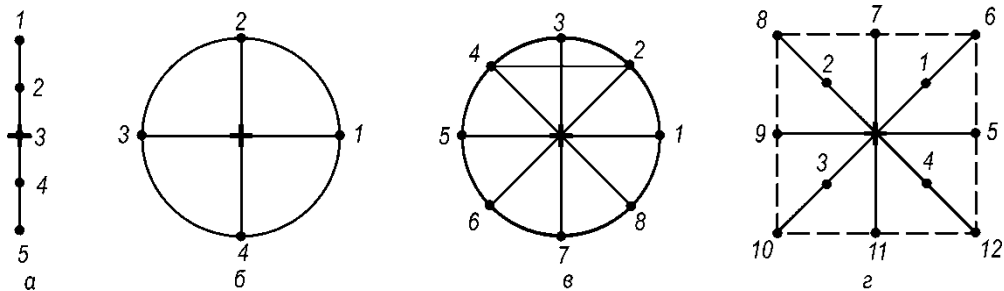


Рисунок 1.19 – Розташування сигнальних точок у 5-позиційній ЦАМ (а), у 4-позиційній ЦФМ (ФМ-4) (б), у 8-позиційній ЦФМ (в), у 12-позиційній ЦАФМ (г)

Це можна зробити наступним перекодуванням: непарні тактові послілки тривалості T двійкової вхідної послідовності подають на один вхід квадратурного модулятора; на інший вхід квадратурного модулятора подають парні послілки вхідної двійкової послідовності; послілки передають по каналу тривалістю інтервалу $T_k = 2T$. Зберігши швидкість передавання інформації $R_i = \frac{\log_2 4}{T_k} = \frac{2}{2T} = \frac{1}{T}$, можна підвищити завадостійкість системи, використовуючи деякі методи кодування.

Цифрова частотна модуляція (ЦЧМ). Якщо частотна модуляція реалізується вибором одного з m незалежних гармонічних сигналів, то в загальному випадку при кожному перемиканні (з i -тої на j -ту позицію; $i, j \in \overline{1, m-1}$) виникає розрив фази каналного сигналу. Дійсно, якщо i -тий гармонічний сигнал:

$$u_i(t) = U_m \cos(\omega_i t + \varphi_i), \quad (1.87)$$

то в момент комутації $t_k = kT$ маємо:

$$u_i(t_k) = U_m \cos(\omega_i kT + \varphi_i), \quad \psi_i(t_k) = \omega_i kT + \varphi_i$$

При комутації j -го генератора

$$u_j(t_k) = U_m \cos(\omega_j kT + \varphi_j), \quad \psi_j(t_k) = \omega_j kT + \varphi_j$$

На практиці смуговий фільтр, який обмежує ширину спектра сигналу, як правило, ввімкнений, і стрибки фази призводять до перехідного процесу у фільтрі. Внаслідок цього виникає паразитна амплітудна модуляція сигналу, і пік-фактор сигналу (відношення його пікової і середньої потужностей) збільшується. Крім того, при використанні m незалежних генераторів для забезпечення ортогональності системи сигналів необхідний рознос частот $\Delta f = i/T$ ($i=1, 2, \dots$), тобто мінімальний рознос $\Delta f_{\min} = 1/T$.

З метою звуження спектра і збереження мінімального пік-фактора сигналу потрібно забезпечити неперервність зміни миттєвої фази сигналу при менших значеннях Δf_{\min} .

Частотну модуляцію з неперервною фазою скорочено позначають ЧМНФ. У системах ЧМНФ миттєва частота сигналу змінюється за законом:

$$\omega(t) = \omega_0 + K_{\text{ЧМ}} \sum_{n=0}^{\infty} b_n^{(k)} v_{\text{ч}}(t - nT), \quad (1.88)$$

а каналний сигнал :

$$u_{\text{ЧМНФ}}(t) = U_m \left(\cos \omega_0 t + K_{\text{ЧМ}} \sum_{n=0}^{\infty} b_n^{(k)} v_{\text{ч}}(t - nT + \varphi_0) \right), \quad (1.89)$$

де $v_{\text{ч}}(t)$ – частотний імпульс; $v(t) = \int v_x(t_1) dt_1$ – фазовий імпульс; φ_0 – початкова фаза; $K_{\text{ЧМ}}$ – девіація частоти.

Якщо $v_{\text{ч}}(t)$ – прямокутний імпульс одиничної висоти, тоді фазовий імпульс:

$$v(t) = \int_0^1 dt_1 = t, \quad t \in [0, T].$$

Для забезпечення більш «гладкого» змінення фази і частоти (відповідно звуження спектра сигналу) на практиці (наприклад, у цифрових системах мобільного стільникового зв'язку за загальноєвропейським стандартом GSM) використовують гауссову форму частотного імпульсу та інтегральну форму фазового:

$$v_{\text{ч}}(t) = \frac{1}{\sqrt{2\pi\tau_i^2}} \exp\left[-\frac{\left(t - \frac{T}{2}\right)^2}{2\tau_i^2}\right], \quad v(t) = F\left(\frac{t - \frac{T}{2}}{\tau_i}\right),$$

де τ_i – величина, пропорційна ефективній тривалості частотного імпульсу.

У подальшому для спрощення аналізу вважатимемо, що частотний імпульс $v_{\text{ч}}(t)$ є прямокутним з одиничною амплітудою і тривалістю T .

Запишемо сигнал (1.88) на відрізку $[0, T]$ при передаванні i -тої позиції символу:

$$u_i(t) = U_0 \cos(\omega_i t + \varphi_{0,k}), \quad k = 1, 2, 3, \dots, \quad (1.90)$$

де $\omega_i = \omega_0 + \frac{2\pi}{T}i$; $\varphi_{0,k}$ – початкова фаза до даного (k -го) тактового інтервалу.

Здійснюючи ЧМНФ, можна забезпечити (при коефіцієнті модуляції $m = 0,5$) ортогональність сигналів (1.90) при частотному зсуві:

$$\frac{\Delta\omega_{\min}}{2\pi} = \Delta f_{\min} = \frac{1}{2T}. \quad (1.91)$$

Дійсно, частотний зсув:

$$(\mathbf{u}_i, \mathbf{u}_j) = U_0^2 \int_0^T \cos(\omega_i t + \varphi_{0,i}) \cos(\omega_j t + \varphi_{0,j}) dt,$$

або

$$(\mathbf{u}_i, \mathbf{u}_j) = \frac{U_0^2 T}{2} \cdot \frac{\sin(\omega_i - \omega_j)T}{(\omega_i - \omega_j)T}. \quad (1.92)$$

При отриманні цього результату враховано, що доданок, який відповідає сумарній частоті $\omega_i + \omega_j$, малий. Мінімальний рознос частот, при якому права частина (1.92) обертається в нуль, визначається зі співвідношення $2\Delta f \pi T = \pi$, звідки випливає (1.91).

Цифрову ЧМ з неперервною фазою і параметром (1.91), тобто $\Delta f_{\min} = \frac{1}{2T}$, називають *модуляцією з мінімальним (частотним) зсувом* – ММЗ (minimum shift keying – MSK). Покажемо, що коефіцієнт модуляції в цій системі $m = 0,5$. Визначимо коефіцієнт частотної модуляції m як відношення девіації частоти $K_{\text{ЧМ}} = \Delta\omega/2$ (максимальне відхилення від середньої частоти) до частоти модулюючого сигналу типу «точка» $\Omega = \pi/T [F = 1/(2T)]$:

$$M = K_{\text{ЧМ}} / \Omega = \frac{\Delta\omega T}{2\pi} = \Delta f T. \quad (1.93)$$

З урахуванням (1.91) і (1.93) отримаємо, що для ММЗ коефіцієнт модуляції дійсно дорівнює 0,5. При використанні для ЧМ сигналів (1.87), ортогональних у підсилювальному значенні, мінімальний індекс $m = 1$. Відсутність стрибків фази в

системах ЧМНФ сприятливо позначається на формі амплітудного спектра сигналу. При $m=0,5$ амплітудний спектр сигналу ЧМНФ дуже вузький і зосереджений поблизу частоти несучої. При значеннях $M > 1$ амплітудний спектр сигналів ЧМНФ стає широким.

Цифрову частотну модуляцію можна реалізувати різними способами, наприклад керуванням частоти генератора гармонічних сигналів за законом (1.88). При цьому початкова фаза опорного сигналу на n -му тактовому інтервалі:

$$\varphi_{\text{оп}} = \pi M \sum_{l=0}^{n-1} b_l + \varphi_0. \quad (1.94)$$

Для систем ММЗ з коефіцієнтом $m=0,5$ широко застосовується квадратурний метод модуляції зі зсувом модулюючих функцій. Для обґрунтування такої модуляції зобразимо сигнал (1.89) при ММЗ $[K_{\text{чМ}} = \pi/(2T)]$ на k -му і $(k+1)$ -му тактових інтервалах як:

$$\begin{aligned} u(t)_{\text{ММЗ}_{k,k+1}} &= U_m \cos \left[\omega_0 t + \frac{\pi}{2T} b_k v(t-kT) + \frac{\pi}{2T} b_{k+1} v[t-(k+1)T] + \varphi_{0k} \right] = \\ &= U_c(t) \cos \omega_0 t - U_s(t) \sin \omega_0 t, \quad t \in [kT, (k+2)T], \end{aligned} \quad (1.95)$$

де U_c , U_s – косинусна і синусна складові сигналу;

$$\varphi_{0k} = \frac{\pi}{2} \sum_{l=0}^{k-1} b_l, \quad (1.96)$$

а квадратурні компоненти:

$$\begin{aligned} U_c(t) &= U_m \cos \left\{ \omega_{0k} + \frac{\pi}{2T} b_k v \left[(t-kT) + \frac{\pi}{2T} b_{k+1} v [t-(k+1)T] \right] \right\}, \\ U_s(t) &= U_m \sin \left\{ \omega_{0k} + \frac{\pi}{2T} b_k v \left[(t-kT) + \frac{\pi}{2T} b_{k+1} v [t-(k+1)T] \right] \right\}. \end{aligned} \quad (1.97)$$

Розіб'ємо інформаційний потік $\{b_k\}$ на два потоки, які відповідають парним і непарним індексам k : $\{b_{2l}\}$ і $\{b_{2l+1}\}$. З урахуванням (1.96) при $k=2l$ маємо $\varphi_{0,2l} = \pi \frac{1}{2} 2m$ (m – ціле число). Тоді, користуючись формулами косинуса і синуса суми кутів, можна записати:

$$U_{s,l}(t) = U_m \cos \pi m \sin \left\{ \frac{\pi}{2T} b_{2l} v(t-2lT) + \frac{\pi}{2T} b_{2l+1} v[t-(2l+1)T] \right\}. \quad (1.98)$$

На відріжку $[2lT; (2l+1)T]$ формула (5.98) набуває вигляду:

$$U_{s,l}(t) = U_m \cos \pi m b_{2l} \sin \left[\frac{\pi}{2T} (t-2lT) \right].$$

На відрізку $[(2l+2)T]$, маючи на увазі, що перший доданок під знаком синуса дорівнює $\frac{\pi}{2}b_{2l}$, і використовуючи формулу $\sin(\pm\pi/2 + \psi) = +\cos\psi$, отримаємо:

$$U_{s,l}(t) = \cos\pi m b_{2l} \cos\left\{\frac{\pi}{2T}[t - (2l+1)T]\right\}.$$

Позначимо $c_{2l} = \cos\pi m b_{2l}$, що можна інтерпретувати як деяке перекодування. Таким чином, на відрізку $[2lT; (2l+2)T]$ для $U_{s,l}(t)$ маємо єдиний вираз:

$$U_{s,l}(t) = U_m c_{2l} \left\{ \sin\left[\frac{\pi}{2T}(t - 2lT)\right] \lambda(t - 2lT) + \cos\left[\frac{\pi}{2T}[t - (2l+1)T]\right] \lambda[t - (2l+1)T] \right\}, \quad (1.99)$$

де $\lambda(t) = \begin{cases} 1, & t \in [0, T] \\ 0, & t \notin [0, T] \end{cases}$ – дискретна функція.

Формулу (1.99) можна подати більш компактно:

$$U_{s,l}(t) = U_m c_{2l} \left| \sin\frac{\pi}{2T}(t - 2lT) \right| \lambda\left(\frac{t - 2lT}{2}\right). \quad (1.100)$$

Розглянемо випадок $k = 2l + 1$. Тоді з урахуванням (1.96):

$$\varphi_{0,2l+1} = \pi m + \frac{\pi}{2} b_{2l}.$$

Підставивши в $U_{c,l}(t)$ значення $k = 2l + 1$, отримаємо:

$$U_{c,l}(t) = U_m \sin\left(\pi m + \frac{\pi}{2} b_{2l}\right) \sin\left\{\frac{\pi}{2T} b_{2l+1} \nu[t - (2l+1)T] + \frac{\pi}{2T} b_{2l+2} \nu[t - (2l+2)T]\right\}.$$

На відрізку $[(2l+1)T; (2l+2)T]$ останній вираз набирає вигляду:

$$U_{c,l}(t) = -U_m \cos\pi m \cdot b_{2l} b_{2l+1} \sin\left\{\frac{\pi}{2T}[t - (2l+1)T]\right\}.$$

На відрізку $[(2l+2)T; (2l+3)T]$ отримаємо:

$$U_{c,l}(t) = -U_m \cos\pi m \cdot b_{2l} b_{2l+1} \cos\left\{\frac{\pi}{2T}[t - (2l+2)T]\right\}.$$

Позначимо:

$$-\cos\pi m \cdot b_{2l} b_{2l+1} = -c_{2l} b_{2l+1} = c_{2l+1},$$

що можна інтерпретувати як деяке перекодування. Тоді на відрізку $[(2l+2)T; (2l+3)T]$ для $U_{c,l}(t)$ отримаємо єдиний вираз :

$$U_{c,l}(t) = -U_m c_{2l+1} \left| \sin\frac{\pi}{2T}[t - (2l+1)T] \right| \lambda\frac{t - (2l+1)T}{2}. \quad (1.101)$$

З урахуванням (1.100) і (1.101) весь сигнал ММЗ можна записати так:

$$\begin{aligned}
u_{\text{ММЗ}}(t) = & -U_m \sum_{l=0}^{\infty} c_{2l} \left| \sin \frac{\pi}{2T} (t - 2lT) \right| \left| \lambda \frac{t - 2lT}{2} \right| \sin \omega_0 t + \\
& + U_m \sum_{l=0}^{\infty} c_{2l+1} \left| \sin \frac{\pi}{2T} [t - (2l+1)T] \right| \left| \lambda \frac{t - (2l+1)T}{2} \right| \cos \omega_0 t.
\end{aligned} \quad (1.102)$$

На рисунку 1.18, *a* дана квадратурна схема, яка формує сигнал ММЗ згідно з (1.102). Позначення для окремих блоків такі: K – кодер, який перетворює інформаційний потік $\{b_{k,T}\}$ двійкових символів, які прямують з тактовим інтервалом T ($b_k \in [-1; 1]$), в два інформаційних потоки двійкових символів $\{c_{2l,2T}\}$ і $\{c_{2l+1,2T}\}$, що прямують з тактовим інтервалом $2T$ ($c_l \in [-1; 1]$); Γ_{Ω} – генератор гармонічного сигналу з частотою $F = 1/(4T)$; $\varphi_{-\pi/2}$ – фазообертач на $-\pi/2$; BVM – блок визначення модуля; Γ_{ω_0} – генератор гармонічного сигналу (несучої) з частотою ω_0 . При реалізації схеми враховано, що:

$$\sin \left\{ \frac{\pi}{2T} [t - (2l+1)T] \right\} = \cos \left\{ \frac{\pi}{2T} (t - 2lT) \right\}.$$

Існують різноманітні способи демодуляції сигналу ММЗ.

1. Частотне детектування без урахування неперервності фази. При такому способі відбувається оцінювання частоти сигналу протягом одного тактового інтервалу T . Оскільки початкова фаза $\varphi_{0,n}$ при цьому не визначається, можливе тільки некогерентне приймання.

2. Когерентне детектування з відслідкуванням фази сигналу. Якщо попередні символи демодульовані без помилок, можливе передбачення початкової фази у відповідності до (1.94). При цьому можливе когерентне приймання, і ймовірність похибки відповідає прийманню двійкових ортогональних сигналів з поправкою на точність оцінки початкової фази.

3. Квадратурна когерентна демодуляція протягом двох тактових інтервалів $2T$ (рис. 1.18, *б*). Можливість такої демодуляції впливає з зображення (1.10). Демодуляція елемента сигналу з номером $n = 2l$ провадиться на інтервалі $[2lT; 2(l+1)T]$ по синусній гілці, а демодуляція елемента сигналу з номером $n = 2l+1$ – на інтервалі $[(2l+1)T; (2l+3)T]$ по косинусній гілці.

На рисунку 1.18, *б* введені такі позначення: $РП$ – розв’язувальний пристрій; $ДК$ – декодер, який здійснює об’єднання символів $\{\hat{c}_{2l,2T}\}$ і $\{\hat{c}_{2l+1,2T}\}$ в єдиний потік, їх декодування і видачу інформаційного потоку $\{\hat{b}_{k,T}\}$; \int – блок інтегрування. Сусідні елементи інформаційної послідовності виділяються в двох гілках обробки, при цьому інтегрування в обох гілках здійснюється зі зсувом на T . Інтегратор верхньої гілки здійснює інтегрування в межах $[kT; (k+2)T]$, інтегратор нижньої гілки – в межах $[(k+1)T; (k+3)T]$. Після виявлення знаку напруги в блоках $РП$ інтегратори зводяться до нульових початкових умов. Вхідні корисні сигнали в кожній з гілок обробки:

$$\begin{aligned}
s_{\text{ЦЧМ}}(t)_1 &= \pm \gamma U_m \left| \sin \Omega t \right| \cos(\omega_0 t + \theta_{\text{кан}}), \\
s_{\text{ЦЧМ}}(t)_2 &= \pm \gamma U_m \left| \cos \Omega t \right| \sin(\omega_0 t + \theta_{\text{кан}}),
\end{aligned}$$

де γ і $\theta_{\text{кан}}$ – коефіцієнт передачі і фазовий зсув в каналі відповідно.

Отже, в кожній з гілок на інтервалі $2T$ аналізується (з метою прийняття рішення про інформаційний символ c_{2l} або c_{2l+1}) система протилежних сигналів. Енергія сигналу по кожній гілці обробки:

$$E = \gamma^2 U_m^2 \int_0^{2T} (\sin \Omega t \cos \omega_0 t)^2 dt = \gamma^2 U_m^2 \int_0^{2T} (\cos \Omega t \sin \omega_0 t)^2 dt = \frac{\gamma^2 U_m^2}{2} T,$$

тобто збігається з енергією гармонічного сигналу амплітуди γU_m на інтервалі $(0; T)$. Тому ймовірність похибки виходить такою ж самою, як при оптимальному когерентному прийманні двійкових протилежних сигналів з тактовим інтервалом T (ФМ2 – двопозиційна фазова модуляція). Але для системи ФМ2 через стрибки фази на початку тактових інтервалів необхідна більша смуга частот, ніж для ММЗ.

Контрольні запитання для самооцінки рівня знань

1. У чому полягає суть модуляції сигналів?
2. Наведіть векторні діаграми АМ- і ЧМ-сигналів.
3. Які переносники при модуляції Ви знаєте?
4. Як визначити потужність АМ-сигналу?
5. Запишіть аналітичний вираз для АМ-сигналу при модуляції несучої одним тоном.
6. При якому виді модуляції ширина спектра модульованого сигналу мінімальна?
7. Чому дорівнює ширина спектра АМ-сигналу?
8. Чому дорівнює ширина спектра ЧМ-сигналу?
9. Назвіть основні види дискретної модуляції.
10. Назвіть основні види імпульсної модуляції.
11. Що визначає ширину спектра сигналу при імпульсній модуляції?
12. Поясніть принципи модуляції шумової несучої.
13. Поясніть принципи побудови дискретних шумоподібних сигналів.
14. Чи є дискретна псевдовипадкова послідовність випадковим процесом? В чому її подібність з шумом?
15. Як здійснюється модуляція шумоподібних сигналів?
16. Як здійснюється цифрова амплітудна модуляція (ЦАМ)?
17. Які особливості цифрової фазової модуляції (ЦФМ)?
18. Нарисуйте квадратурні схеми формування двійкового сигналу ММЗ та його детектування.

2 МЕТОДИ ОБРОБКИ СИГНАЛІВ

2.1 Загальні відомості про приймання сигналів

Приймальна сторона звичайно має деякі попередні (априорні) відомості про передані сигнали, наприклад, про частоту несучої або вид модуляції. Сигнал, про який заздалегідь усе відомо, не несе інформації, а абсолютно невідомий сигнал не можна було б прийняти.

Відомі параметри сигналу використовуються в приймачі для кращого відділення сигналів від завад. Чим більше ми знаємо про сигнал, тим досконалішими можуть бути методи його приймання.

Параметри, в змінах яких закладена переносна сигналом інформація, називаються *інформаційними*. Зміни цих параметрів на приймальній стороні заздалегідь невідомі.

За видом і призначенням передавання в процесі приймання сигналів виникають такі *основні задачі*, як виявлення сигналів, їх розрізнення та відновлення.

Задача виявлення сигналів зводиться до одержання відповіді на питання, що саме знаходиться на вході приймача: сигнал, а точніше сигнал плюс шум, або тільки шум? З такою задачею звичайно зустрічаються в радіолокації, вона також має місце й у деяких системах зв'язку. Якщо ми в змозі знайти сигнал, то передати інформацію можна за допомогою двійкового коду. Наявність сигналу (посилка) відповідатиме символу 1, відсутність сигналу (пауза) – символу 0. Цей принцип використовується в системах з *пасивною паузою*.

Якщо передаються одночасно два сигнали (s_1 і s_2), виникає задача не виявлення, а розрізнення сигналів. У такому разі необхідно визначити, який на вході приймача сигнал – s_1 чи s_2 ? Відповідь на це питання залежить вже не від властивостей кожного сигналу окремо, а від їхньої розбіжності. Сигнали можуть відрізнятися один від одного своїми параметрами, і очевидно, чим більше вони відрізнятимуться, тим краще, за умови стійкості сигналів до впливу завад. Виявлення сигналів може розглядатися як вироджений випадок розрізнення двох сигналів, коли один з них тотожно дорівнює нулю [10].

Передавання двійковим кодом, коли кожному символу (1 і 0) відповідає певний сигнал (s_1 і s_2), який не дорівнює нулю, називається передаванням з *активною паузою*. Випадок розрізнення багатьох сигналів принципово мало відрізняється від випадку розрізнення двох сигналів.

Задача відновлення повідомлення істотно відрізняється від задач виявлення і розрізнення сигналів. Вона полягає в тому, щоб одержати такий вихідний відеосигнал $v(t)$, який найменше відрізнявся би від переданого повідомлення $u(t)$. При цьому про повідомлення $u(t)$ відомо лише, що воно належить до деякої множини або є реалізацією деякого випадкового процесу.

На рисунку 2.1 наведена функціональна схема *обробки* дискретних сигналів. Прийнятий сигнал, спотворений завадою, піддається в приймачі певній обробці, детектується і надходить для впізнання на розв'язувальний пристрій. Очевидно, що ймовірність правильного впізнання сигналів істотно залежить від відношення потужності сигналу до потужності завади на вході розв'язувального пристрою. Тому основною метою обробки сигналів у приймачі є збільшення відношення

сигнал–шум. Обробка аналів, як правило, зводиться до фільтрації тими чи іншими методами.

На етапах обробки сигнали підсилюються до величин, при яких можуть нормально працювати детектор і розв'язувальний пристрій.

У звичайному приймачі неперервних сигналів *додетекторна обробка сигналів* здійснюється за допомогою резонансних підсилювачів, які забезпечують необхідну частотну селекцію. Функції *післядетекторної обробки* при цьому виконуються відеопідсилювачем (чи підсилювачем низької частоти). Розв'язувальний пристрій у таких приймачах відсутній. Замість нього на виході є реєструвальний пристрій, що відтворює чи записує прийняте повідомлення.

У системах передавання дискретних повідомлень основними видами обробки сигналів у приймачі є фільтрація зі стробуванням, інтегрування і кореляційна обробка.

Стробування сигналу є найпростішим видом обробки, давно відомим і широко застосовуваним у практиці. Часто його називають ще прийманням за методом укорочених контактів. При стробуванні даного елемента сигналу його поточне значення (напруга або струм) відліковується у визначений момент часу. Такий момент вибирається в тій частині елемента, яка спотворюється найменше. Оскільки статистичні характеристики завад мало залежать від вибору моменту часу, то момент стробування має відповідати максимальному значенню корисного сигналу. Внаслідок перехідних процесів у фільтрах границі елементів сигналу в значній мірі спотворені; тому максимальне значення сигналу, як правило, відповідає середній частині елемента. Стробування виконується за допомогою спеціальних сигналів, що надходять від системи синхронізації.

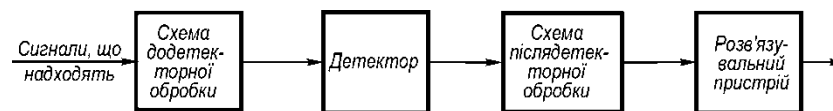


Рисунок 2.1 – Функціональна схема обробки дискретних сигналів [8]

Фільтрація прийнятих сигналів може виконуватися як до, так і після детектора. У разі використання синхронного детектора фільтрація до і після нього принципово може дати той самий результат. Практично ж звичайно сигнал фільтрують двічі: як до, так і після детектора.

Операція *інтегрування* може розглядатися або як процес накопичення, або як визначення середнього значення сигналу. Будь-який фільтр тою чи іншою мірою виконує інтегрування коливання, що надходить на його вхід. За певних умов операція інтегрування еквівалентна фільтрації. Інтегрування, як і фільтрація, може здійснюватися або до, або після детектора.

Методи приймання можна класифікувати за видами застосовуваних детекторів, способами додетекторної і післядетекторної обробки. Розрізняють наступні методи приймання сигналів: кореляційний; автокореляційний; когерентний; некогерентний, які будуть розглянуті далі.

2.2 Приймання сигналів з накопиченням

Одним з найвідоміших методів приймання сигналів за наявності завад є метод накопичення, який полягає в багаторазовому передаванні сигналу (дублюванні) і є

окремим випадком кореляційного методу. Кожний з переданих сигналів по-різному спотворюється завадами, тому що переданий сигнал і завада взаємно незалежні. На приймальному кінці каналу зв'язку кілька зразків прийнятого сигналу порівнюються між собою, внаслідок чого вдається більш точно відновити справжню форму переданого сигналу і завдяки цьому підвищити ймовірність його правильного приймання. Практично порівняння різних зразків прийнятого сигналу здійснюється їхнім підсумовуванням, тому цей метод і одержав назву **метод накопичення** [6].

Прийнятий за методом накопичення сигнал одержує нову властивість – *властивість періодичного повторення*: енергія корисного сигналу збільшується при l -кратному повторенні сигналу в l^2 разів, а енергія завади, яка є випадковим сигналом, – в l разів. Отже, відношення $\frac{P_c}{P_z}$ на виході накопичувача зростає в l разів внаслідок збільшення часу передавання сигналу в l разів, тобто за рахунок зниження ефективності системи зв'язку.

Метод накопичення в найпростішій формі часто застосовується при телефонній розмові в умовах поганої чутності, коли перепитують і повторюють те саме слово кілька разів.

У телеграфному зв'язку кілька разів передається кожна кодова комбінація, що складається з елементів 0 і 1. Якщо ймовірність перебою символів 1 і 0 однакова, то рішення приймається за «більшістю», а саме: відтворюється символ 1 на даній позиції, коли число цих символів на ній більше за число символів 0, і, навпаки, відтворюється 0, коли число нулів більше за число одиниць. Наприклад:

Передана комбінація	01001
Перша прийнята комбінація	00001
Друга прийнята комбінація	11010
Третя прийнята комбінація	01101
Відтворена комбінація	01001

Зазначимо, що можна було б одержати n зразків сигналу не їхнім повторенням у часі, а передаванням по незалежних каналах, розділених за частотою, або яким-небудь іншим способом.

Існують й інші різновиди методу накопичення. До них, зокрема, належить *метод синхронного накопичення*, що застосовується в радіолокації, коли протягом посилення сигналу береться не один відлік, а кілька, які підсумовуються в накопичувачі.

Розглянемо як приклад метод синхронного накопичення для *виявлення слабких періодичних сигналів*. За цим методом прийнятий сигнал розбивається на відрізки, тривалість яких дорівнює $\frac{T}{m}$, де T – період проходження сигналу, а m – ціле число. Відрізки сигналу тривалістю $\frac{T}{m}$ через інтервал T та із зсувом відносно початку відліку $\frac{k}{m}T$ подаються на накопичувач і підсумовуються (рисунку 2.2, а).

Визначимо відношення сигнал–завада на k -му накопичувачі. Позначимо миттєві значення завади, що надходить на накопичувач, через:

$$n_i = n \left(t + k \frac{T}{m} + iT \right), \quad (2.1)$$

де t – поточний час; k – номер накопичувача; $\frac{T}{m}$ – час дії накопичувача; T – період проходження сигналу; i – цикл накопичення.

За l повторень сигналу на накопичувачі середнє значення завади дорівнюватиме:

$$\bar{n}_i = \frac{1}{n} \sum_{i=1}^l n_i. \quad (2.2)$$

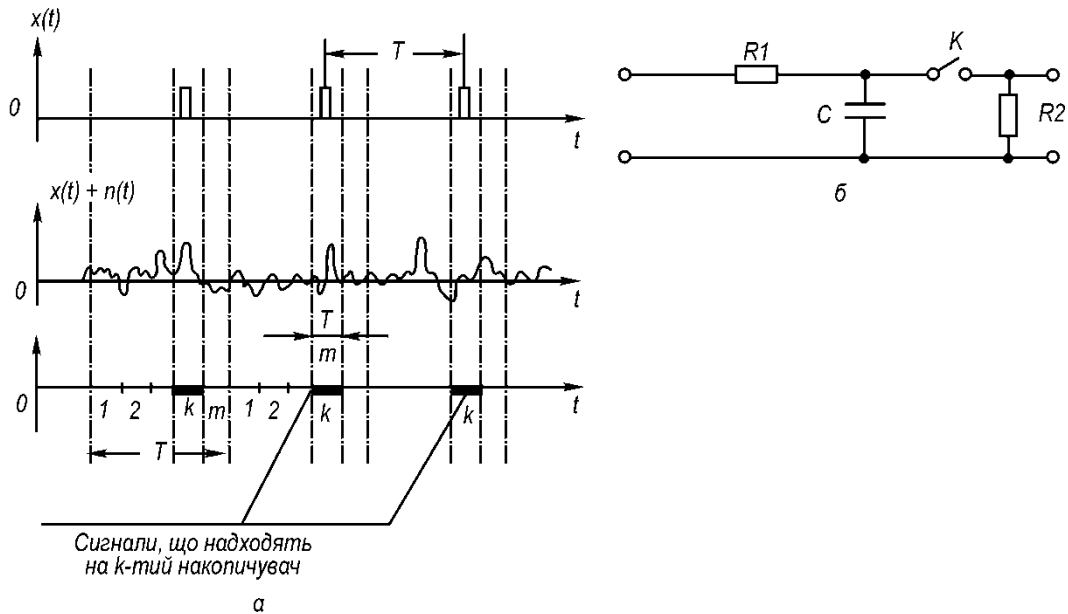


Рисунок 2.2 – Ілюстрація до методу накопичення (а) і найпростіша схема накопичувача (б) [5]

Оскільки n_i – випадкова величина, то при малому числі l значення \bar{n}_i сильно флюктує біля свого математичного сподівання. Зі збільшенням l імовірність великих флюктуацій зменшується.

Корисний сигнал $x(t)$, що надходить на накопичувач за l циклів накопичення, збільшується в l разів, тобто стає рівним $lx(t)$. Помилка при прийманні можлива до тих пір, поки флюктуації перевищуватимуть корисний сигнал за абсолютною величиною.

Визначимо ймовірність того, що відхилення завади \bar{n}_i від математичного сподівання не перевищить величини a – середнього значення сигналу за l повторень. Ця ймовірність визначається нерівністю Чебишова:

$$P\left\{\left|\bar{n}_i - \bar{n}\right| < a\right\} > 1 - \frac{D(n)}{lx^2}, \quad (2.3)$$

де $D(n)$ – дисперсія $n(t)$, яка дорівнює $P_3^2 = \sigma^2$, $x^2 = P_c$.

Розв'язуючи нерівність (6.3) відносно l , одержимо:

$$l < \frac{1}{1-P} \frac{P_3}{P_c} = \frac{1}{1-P} \left(\frac{\sigma}{x} \right)^2. \quad (3.4)$$

Отже, верхня границя частоти повторення l при заданих P , P_c і P_3 може бути визначена за (2.4).

Імовірність $P\{|\bar{n}_i - \bar{n}| < x\}$ може бути знайдена за наближеною формулою:

$$P\{|\bar{n}_i - \bar{n}| < x\} \approx \frac{2}{\sqrt{\pi}} \int_0^{\frac{x}{\sigma} \sqrt{\frac{l}{2}}} e^{-z^2} dz = \Phi\left(\frac{x}{\sigma} \sqrt{\frac{l}{2}}\right). \quad (3.5)$$

Імовірність помилки при прийманні сигналу методом синхронного накопичення дорівнює :

$$P_{\text{пом}} = 1 - P\{|\bar{n}_i - \bar{n}| < x\} \approx 1 - \Phi\left(\frac{x}{\sigma} \sqrt{\frac{l}{2}}\right).$$

З цієї рівності випливає, що завадостійкість системи з синхронним накопиченням тим більша, чим більше відношення сигнал–завада на вході накопичувача і чим більше число циклів накопичення l .

Крім розглянутого методу приймання слабких сигналів, існують й інші. Наведемо деякі з них.

Нехай окремі відліки прийнятого сигналу:

$$x_1 = s + \omega_1, \quad x_2 = s + \omega_2, \dots \quad + \omega_n; \quad (2.6)$$

тоді сума відліків:

$$z = \sum_{k=1}^n x_k = \sum_{k=1}^n (s + \omega_k) = ns + \sum_{k=1}^n \omega_k = b + \xi. \quad (2.7)$$

Величина $b = ns$ – це корисний сигнал на виході приймача, а випадкова величина $\xi = \sum_{k=1}^n \omega_k$ – завада. Відношення сигнал–завада на виході приймача дорівнює:

$$q_{\text{вих}} = \frac{b^2}{D\xi} = \frac{n^2 s^2}{D(\sum \omega_k)} = \frac{n^2 s^2}{\sum D\omega_k} = \frac{n^2 s^2}{nD\omega} = \frac{n^2 s^2}{\sigma_3^2} = nq_{\text{вх}}. \quad (3.8)$$

Тут ми вважали, що значення ω_k некорельовані і мають однаковий розподіл;

$q_{\text{вх}} = \frac{s^2}{\sigma_3^2} = \frac{P_c}{P_3}$ – відношення сигнал–завада на вході приймача.

Отже, за описаних умов накопичення відліків сигналу дозволяє збільшити відношення сигнал–завада на виході приймача рівно в n разів, оскільки потужність сигналу при підсумовуванні зростає пропорційно n^2 (додаються напруги), а потужність завади – пропорційно n (підсумовуються потужності), якщо завади незалежні. За наявності кореляції між значеннями потужності завади цей вигравш буде меншим.

Метод накопичення можна здійснити, взявши не суму відліків x_k , а інтеграл неперервно змінної функції $x(t) = s + \omega(t)$ за час T , який дорівнює тривалості сигналу:

$$\int_0^T x(t) dt = sT + \int_0^T \omega(t) dt = b + \xi.$$

Тут b – стала, яка визначає сигнал на виході накопичувача (інтегратора), а ξ – випадкова величина, що визначає заваду на виході інтегратора.

Дисперсія випадкової величини ξ :

$$\begin{aligned} D\xi &= \left[\int_0^T \omega(t) dt \right]^2 = \left[\int_0^T \int_0^T \omega(t) \omega(t_1) dt dt_1 \right] = \\ &= \int_0^T dt \int_0^T \omega(t) \omega(t_1) dt_1 = \int_0^T dt \int_0^T B_\omega(t-t_1) dt_1, \end{aligned} \quad (2.9)$$

де $B_\omega(t-t_1)$ – функція кореляції завади $\omega(t)$.

Якщо спектр завади рівномірний у досить широкій смузі частот F , тобто інтервал кореляції завади $\Delta\tau \ll T$, то скінченні границі інтегрування можуть бути замінені на нескінченні:

$$\int_0^T B_\omega(t-t_1) dt_1 \approx \int_0^\infty B_\omega(\tau) d\tau = 2 \int_0^\infty B_\omega(\tau) d\tau, \quad (2.10)$$

де $t-t_1 = \tau$.

За означенням інтервал кореляції:

$$\Delta\tau = \frac{1}{B(0)} \int_{-\infty}^\infty B(\tau) d\tau = \frac{G(0)}{B(0)}. \quad (2.11)$$

У розглянутому випадку $B(0) = G(0) \cdot 2F$ і $\Delta\tau = \frac{1}{2F}$. Отже, $D\xi = B(0) \Delta\tau T = B(0) \frac{T}{2F}$. Тоді відношення сигнал–завада на виході інтегратора:

$$q_{\text{вих}} = \frac{b^2}{D\xi} = \frac{s^2 T^2}{B(0) \Delta\tau T} = \frac{T}{\Delta\tau} = 2TF q_{\text{вх}}. \quad (2.12)$$

Як бачимо, виграш, одержуваний при інтегруванні, тим більший, чим більше відношення $T/\Delta\tau$. Описаний метод приймання називається *інтегральним*.

Зазначимо, що $\frac{T}{\Delta\tau} = n$ – це число незалежних відліків завади на інтервалі T .

Це означає, що формули (3.8) і (3.12) збігаються, тобто заміна підсумовування незалежних значень неперервним інтегруванням додаткового виграшу не дає. Однак практично метод інтегрування реалізується простіше, ніж підсумовування дискретних значень. Так, для приймання телеграфних сигналів за інтегратор широко використовується ланцюжок RC, що синхронно розряджається по

закінченні кожної елементарної послідовності (рис. 2.2, б). Наприкінці кожної послідовності заряд на ємності приблизно пропорційний інтегралу вхідного сигналу. Додетекторне інтегрування можна здійснити за допомогою резонатора великої добротності. У найпростішому випадку таким резонатором є коливальний контур.

2.3 Кореляційне приймання

Структурна схема кореляційного приймача наведена на рисунку 2.3, а. Вона складається з перемножника *MPL*, генератора опорного коливання *Г* та інтегратора *I*.

При кореляційному прийманні в деякий момент часу *T* вимірюються значення функції взаємної кореляції *y(T)* прийнятого сигналу $x(t) = s(t) + \omega(t)$ й опорного коливання $s_0(t)$. Якщо опорне коливання тотожне переданому сигналу, тобто $s_0(t) \equiv s(t)$, то цю функцію можна записати в такому вигляді:

$$y(T) = \frac{1}{T} \int_0^T x(t)s(t) dt + i \frac{2}{T} \int_0^T x(t)\hat{s}(t) dt, \quad (2.13)$$

де $s(t)$ і $\hat{s}(t)$ – аналітичні сигнали, що відповідають $x(t)$ і $s(t)$; $\hat{s}(t)$ – функція, комплексно-спряжена з сигналом $s(t)$.

У залежності від методу реєстрації сигналу на виході кореляційне приймання може бути когерентним і некогерентним. При *когерентному* прийманні в момент часу *T* відраховується дійсне значення функції *y(T)*, тобто:

$$\text{Re}[y(T)] = \frac{2}{T} \int_0^T x(t)s(t) dt. \quad (2.14)$$

При *некогерентному* прийманні відраховується модуль цієї функції:

$$\left| \frac{1}{T} \int_0^T x(t)s(t) dt \right| = \frac{2}{T} \sqrt{\left[\int_0^T x(t)s(t) dt \right]^2 + \left[\int_0^T x(t)\hat{s}(t) dt \right]^2}. \quad (2.15)$$

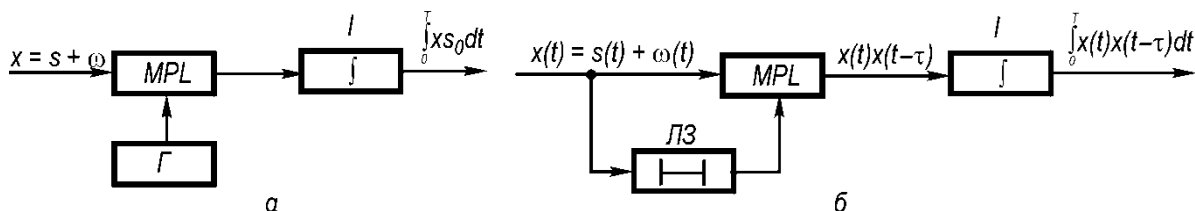


Рисунок 2.3 – Структурні схеми кореляційного (а) і автокореляційного (б) приймачів [1]

Дійсне значення сигналу на виході кореляційного приймача можна зобразити у вигляді суми двох складових:

$$y = \frac{1}{T} \int_0^T x(t) s(t) dt = b + \xi,$$

де $b = \frac{1}{T} \int_0^T s^2(t) dt = P_c$ – корисний сигнал; $\xi = \frac{1}{T} \int_0^T s(t) \omega(t) dt$ – завада на виході приймача.

Дисперсія завади:

$$\begin{aligned} D\xi &= \overline{\left[\frac{1}{T} \int_0^T s(t) \omega(t) dt \right]^2} = \frac{1}{T^2} \overline{\int_0^T \int_0^T s(t) s(t_1) \omega(t) \omega(t_1) dt dt_1} = \\ &= \frac{1}{T^2} \int_0^T s(t) dt \int_0^T s(t_1) \overline{\omega(t) \omega(t_1)} dt_1 = \frac{1}{T^2} \int_0^T s(t) \int_0^T B_\omega(t-t_1) dt_1 dt. \end{aligned} \quad (2.16)$$

Припустимо, що інтервал кореляції завади $\Delta\tau$ настільки малий, що сигнал протягом цього інтервалу практично не змінюється (ширина спектра сигналу мала в порівнянні з шириною спектра завади). Тоді з урахуванням співвідношень (2.10) і (2.11):

$$D\xi \approx \frac{1}{T^2} \int_0^T s^2(t) dt \int_0^T B_\omega(t-t_1) dt_1 \approx \frac{\Delta\tau}{T} P_c P_3, \quad (2.17)$$

де $P_3 = B_\omega(0)$ – потужність завади на вході приймача.

Відношення сигнал–завада на виході кореляційного приймача при когерентному прийманні дорівнюватиме:

$$q_{\text{кг}} = \frac{b^2}{D\xi} = \frac{T}{\Delta\tau} q_{\text{вх}} \approx 2TF q_{\text{вх}}. \quad (2.18)$$

При некогерентному прийманні це відношення вдвічі менше:

$$q_{\text{нкг}} = \frac{b^2}{2D\xi} = \frac{1}{2} \frac{T}{\Delta\tau} q_{\text{вх}} \approx TF q_{\text{вх}}. \quad (2.19)$$

Із порівняння формул (2.12) і (2.18) випливає, що кореляційний метод приймання можна розглядати як узагальнення методу накопичення на сигнали довільної форми.

2.4 Автокореляційне приймання

У схемі автокореляційного приймача немає спеціального генератора опорних коливань (рис. 2.3, б), але є лінія затримки ЛЗ, що здійснює затримку прийнятого сигналу на час τ . Затримане коливання $x(t-\tau)$ використовується як опорне.

У випадку автокореляційного приймання виконується інтегральна операція:

$$y = \frac{1}{T} \int_0^T x(t) x(t-\tau) dt = \frac{1}{T} \int_0^T [s(t) + \omega(t)][s(t-\tau) + \omega(t-\tau)] dt. \quad (2.20)$$

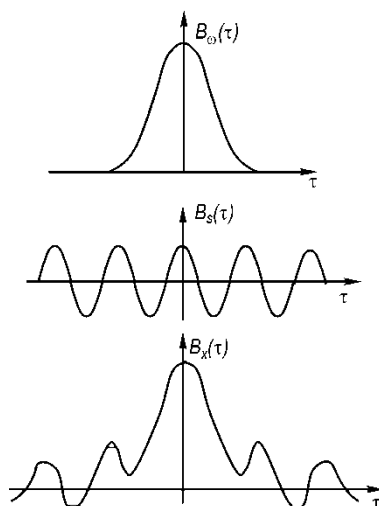


Рисунок 2.4 – Кореляційні функції сигналу та завади

Розкривши дужки і взявши $\tau \square$ дістанемо:

$$y = \frac{1}{T} \int_0^T s(t)s(t-\tau)dt + \frac{1}{T} \int_0^T s(t)\omega(t-\tau) + \frac{1}{T} \int_0^T \omega(t)s(t-\tau)dt + \frac{1}{T} \int_0^T \omega(t)\omega(t-\tau)dt = B_{ss}(\tau) + B_{s\omega}(\tau) + B_{\omega s}(\tau) + B_{\omega\omega}(\tau) = b + \xi, \quad (2.21)$$

де $b = B_{ss}(\tau) = P_c \rho_s(\tau)$; $\xi = B_{s\omega}(\tau) + B_{\omega s}(\tau) + B_{\omega\omega}(\tau)$.

Далі можна показати, що відношення сигнал–завада на виході автокореляційного приймача при малому рівні завад ($q_{\text{вх}} \square$), приблизно таке ж саме, як і на виході кореляційного приймача. Якщо рівень завад великий ($q_{\text{вх}} \square$), то автокореляційний приймач за завадостійкістю наближається до приймача з квадратичним детектором. Менша завадостійкість автокореляційного приймача в порівнянні з кореляційним при будь-яких значеннях $q_{\text{вх}}$ обумовлена наявністю завади в тракці опорної напруги. Автокореляційне приймання можливе і тоді, коли відсутні відомості не тільки про фазу сигналу, але й про частоту, оскільки опорне коливання породжується самим сигналом, що приймається, а не створюється спеціальним генератором у місці приймання.

Для прикладу розглянемо автокореляційне приймання періодичного сигналу за наявності адитивної флуктуаційної завади. Нехай $s(t)$ – періодичний сигнал, а $\omega(t)$ – завада, яку вважатимемо ергодичним процесом. Сумарне коливання сигналу та завади на вході приймача $x(t) = s(t) + \omega(t)$. Визначимо функцію кореляції цього коливання:

$$B_x(\tau) = M \{ [s(t) + \omega(t)][s(t+\tau) + \omega(t+\tau)] \} = B_s(\tau) + B_{s\omega}(\tau) + B_{\omega s}(\tau) + B_{\omega\omega}(\tau).$$

Оскільки сигнал і заваду можна вважати незалежними, то їх взаємкореляційні функції $B_{s\omega}$ і $B_{\omega s}$ перетворюються в нуль; тому

$$B_x(\tau) = B_s(\tau) + B_{\omega\omega}(\tau).$$

Функція автокореляції завади $B_{\omega\omega}$ як ергодичного процесу спадає зі зростанням аргументу τ (рис. 2.4, а), а функція автокореляції періодичного

сигналу – це періодична функція з періодом сигналу (рис. 2.4, б). Функція кореляції прийнятого сигналу є сумою цих двох функцій (рис. 2.4, в). Отже, при досить великому τ можна одержати періодичну криву, що відповідає переданому сигналу $s(t)$, тобто виділити сигнал із шуму. Величина τ , яка дозволяє знехтувати $B_\omega(\tau)$, залежить від інтервалу кореляції завади та відношення сигнал–завада. Очевидно, чим більшим є рівень завад, тим більшим має бути час затримки.

2.5 Когерентне та некогерентне приймання

Функціональна схема когерентного приймання сигналу наведена на рисунку 2.5. Вона складається з перемножника *MPL*, генератора опорної напруги Γ і фільтра нижніх частот *ФНЧ*. Опорне коливання $s_0(t)$ при когерентному прийманні є точною копією переданого сигналу $s(t)$.

Якщо сигналом $s(t)$ є коливання з відомими частотою і фазою, то в приймачі використовується *синхронний детектор*, у якому опорні коливання синхронні з коливанням несучої частоти. Фільтр нижніх частот виконує роль інтегратора, він виділяє на виході напругу, яка практично збігається з обвідною вхідного (високочастотного) сигналу.

При некогерентному прийманні апріорні відомості про початкову фазу прийнятого сигналу не враховуються, тому в схемі приймача можна застосувати не синхронний, а *амплітудний детектор*.

Нехай на приймач надходять регулярний сигнал:

$$s(t) = A_0 \cos \omega_0 t, \quad 0 \leq t \leq T \quad (2.22)$$

і флуктаційна завада $\omega(t)$. Заваду з симетричним щодо частоти сигналу спектром можна записати у вигляді квазігармонічного коливання:

$$\begin{aligned} \omega(t) &= U_1 \cos \omega_0 t + U_2 \sin \omega_0 t = \\ &= U \cos[\omega_0 t + \varphi(t)], \end{aligned} \quad (3.23)$$

де U_1 і U_2 – випадкові незалежні величини, що мають нормальний розподіл імовірностей із середнім значенням, рівним нулю, і дисперсією $\sigma_3^2 = \overline{U_1^2(t)} = \overline{U_2^2(t)} = N_0 F$, де N_0 – спектральна щільність завади, F – ефективна ширина спектра завади.

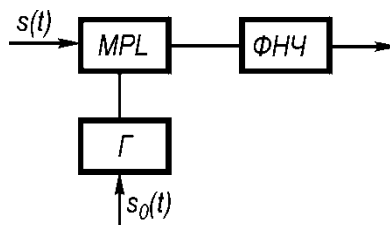


Рисунок 2.5 – Структурна схема когерентного приймача [2]

Сумарне коливання сигналу і завади:

$$x(t) = s(t) + \omega(t) = (A_0 + U_1) \cos \omega_0 t + U_2 \sin \omega_0 t. \quad (2.24)$$

Відповідно до виразу (2.24) завада має дві складові – синусну і косинусну, що знаходяться в квадратурі.

При когерентному прийманні детектор виділяє сигнал і синфазну складову завади U_1 . Квадратурна складова завади U_2 при цьому цілком виключається. Помилки когерентного приймання обумовлені флуктуаціями амплітуди завади U_1 , що має нормальний розподіл імовірностей:

$$P(\omega) = \frac{1}{\sqrt{2\pi} \cdot \sigma_3} e^{-\omega^2/(2\sigma_3^2)}. \quad (2.25)$$

При некогерентному прийманні детектор реагує на обидві складові завади (U_1 та U_2), виділяючи обвідну сумарного коливання сигналу і завади x . Помилки в цьому випадку обумовлені флуктуаціями обвідної U сумарного коливання $x(t)$, що розподілена за узагальненим законом Релея:

$$P(U) = \frac{U}{\sigma_3^2} I_0 \left(\frac{AU}{\sigma_3^2} \right) e^{-\frac{U^2+A^2}{2\sigma_3^2}}, \quad (2.26)$$

де $I_0(x)$ – модифікована функція Бесселя нульового порядку; σ_3^2 – дисперсія завади.

Розподіл (2.26) називається *узагальненим розподілом Релея* (або *розподілом Райса*). При $A=0$ розподіл (2.26) перетворюється в звичайний розподіл Релея.

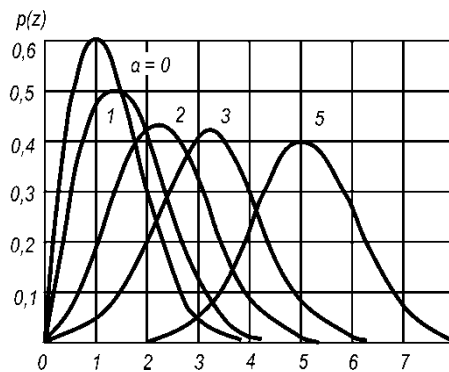


Рисунок 2.6 – Узагальнений розподіл Релея [4]

Взявши $z = U/\sigma_3$ і $a = A/\sigma_3$, розподіл Релея можна записати в такому вигляді:

$$p(z) = z e^{-\frac{a^2+z^2}{2}} I_0(az). \quad (2.27)$$

На рисунку 2.6 зображені графіки цього розподілу для різних значень $a = A/\sigma_3 = \sqrt{2P_c/P_3}$, де $P_c = A^2/2$ – потужність сигналу; $P_3 = \sigma_3^2$ – потужність завади.

Щільність імовірності фази сумарного коливання
:

$$P(\varphi) = \frac{1}{2\pi} e^{-\frac{A^2}{2\sigma_3^2}} + \frac{1}{2} \frac{A \cos \varphi}{\sqrt{2\pi\sigma_3^2}} \left[1 + \Phi \left(\frac{A \cos \varphi}{2\sigma_3^2} \right) \right] e^{-\frac{A^2 \sin^2 \varphi}{2\sigma_3^2}}, \quad (2.28)$$

де $\Phi(x)$ – функція Крампа.

Для некогерентного приймання використовуються різні типи детекторів, які класифікують за видом їхніх характеристик: лінійні, квадратичні тощо.

Розглянемо вплив синусоїдального сигналу й вузькосмугового нормального шуму на амплітудний детектор і визначимо відношення сигнал–завада на його виході. Ця задача розв’язується за допомогою добре розробленого кореляційного методу, який є досить загальним і зводиться до наступного. За заданою характеристикою детектора $y = f(x)$ визначається середнє значення y (струму або напруги):

$$\overline{y(t)} = \int_{-\infty}^{\infty} f(s + \omega) p(\omega) d\omega. \quad (2.29)$$

Для одержання сталої складової вхідного коливання необхідно $\overline{y(t)}$ усереднити ще за часом:

$$y_0 = \overline{\overline{y(t)}} = \overline{\int_{-\infty}^{\infty} f(s + \omega) p(\omega) d\omega}. \quad (2.30)$$

Флуктуаційна складова на виході детектора визначається різницею $y - \overline{y} = \xi$, а регулярна змінна складова $b = \overline{y} - y_0$. Під сигналом звичайно розуміють або низькочастотну складову b , або приріст постійної складової при подаванні на вхід сигналу. Таким чином, коливання на виході детектора можна зобразити у вигляді суми

$$y = y_0 + b + \xi. \quad (2.31)$$

Перший доданок є сталою складовою, другий – регулярною складовою (сигнал) і третій – завадою на виході детектора. Наскільки потужність сигналу більша за потужність завади, можна визначити відношенням b^2 до дисперсії завади $D\xi$:

$$q_{\text{вих}} = \frac{b^2}{D\xi}. \quad (2.32)$$

Нехай характеристика детектора є квадратичною: $y = x^2$. На детектор впливають сигнал (2.22) і завада (2.23). Тоді коливання на виході детектора

$$\begin{aligned} y(t) &= [s(t) + \omega(t)]^2 = \{A_0 \cos \omega_0 t + U(t) \cos[\omega_0 t + \varphi(t)]\}^2 = \\ &= \frac{A_0^2}{2} + \frac{U^2(t)}{2} + A_0 U(t) \cos \varphi(t) + y_{\text{вч}}(t) = y_{\text{нч}}(t) + y_{\text{вч}}(t). \end{aligned}$$

Високочастотні складові $y_{\text{вч}}$ відфільтровуються, а тому можуть не братися до уваги. На виході фільтра нижніх частот

$$y_{\text{нч}}(t) = \frac{A_0^2}{2} + \frac{U^2(t)}{2} + U(t) A_0 \cos \varphi(t) = b + \xi. \quad (2.33)$$

Перший доданок $b = \frac{A_0^2}{2}$ визначає сигнал, другий і третій $\xi = \frac{U^2(t)}{2} + U(t) A_0 \cos \varphi$ – заваду на виході. Визначимо дисперсію шуму на виході детектора:

$$D\xi = (\xi - \bar{\xi})^2 = \sigma_3^4 + A_0^2 \sigma_3^2. \quad (2.34)$$

Тут враховано, що $\overline{U^2(t)} = 2\sigma_3^2$ і $\overline{U(t)\cos\varphi(t)} = 0$. Відношення сигнал–завада на виході квадратичного детектора:

$$q_{\text{вих}} = \frac{b^2}{D\xi} = \frac{A_0^4}{4(\sigma_3^4 + A_0^2 \sigma_3^2)} = \frac{q_{\text{вх}}^2}{1 - 2q_{\text{вх}}}, \quad (2.35)$$

де $q_{\text{вх}} = \frac{A_0^2}{2\sigma_3^2}$ – відношення сигнал–завада на вході детектора.

При $q_{\text{вх}} = \frac{A_0^2}{2\sigma_3^2}$ менше та більше одиниці маємо:

$$q_{\text{вих}} \approx \frac{1}{2} q_{\text{вх}}, \quad (2.36)$$

$$q_{\text{вих}} \approx q_{\text{вх}}^2. \quad (2.37)$$

Оскільки остання залежність є квадратичною, то зі зменшенням $q_{\text{вх}}$ значення $q_{\text{вих}}$ різко зменшується, отже, слабкий сигнал придушується завадою. Так, при $q_{\text{вх}} = 0,1$ маємо $q_{\text{вих}} \approx 0,01$, тобто придушення в десять разів.

Аналогічно розв'язується задача для лінійного й інших типів детекторів. Не зупиняючись на їхньому аналізі, відзначимо, що при слабких сигналах будь-який амплітудний детектор (крім синхронного) є квадратичним.

При когерентному (синхронному) детектуванні опорне коливання $s_0(t)$ вибираємо подібним до очікуваного сигналу:

$$s_0(t) = B_0 \cos \omega_0 t. \quad (2.38)$$

На виході детектора одержуємо добуток двох коливань:

$$y(t) = x(t)s_0(t) = [s(t) + \omega(t)]s_0(t).$$

Відповідно до (2.22), (2.23) і (2.38) маємо:

$$\begin{aligned} y(t) &= A_0 \cos \omega_0 t + U \cos(\omega_0 t + \varphi) B_0 \cos \omega_0 t = \\ &= A_0 B_0 \cos^2 \omega_0 t + A_0 U \cos \omega_0 t \cos(\omega_0 t + \varphi) = \\ &= \frac{A_0 B_0}{2} + \frac{1}{2} A_0 B_0 \cos 2\omega_0 t + \frac{B_0 U}{2} \cos(2\omega_0 t + \varphi) + \frac{1}{2} B_0 U \cos \varphi. \end{aligned}$$

Звідси:

$$y_{\text{нч}}(t) = \frac{A_0 B_0}{2} + \frac{1}{2} B_0 U \cos \varphi = b + \xi, \quad (2.39)$$

де $b = \frac{A_0 B_0}{2}$; $\xi = \frac{1}{2} B_0 U \cos \varphi$.

Дисперсія шуму на виході:

$$D\xi = \frac{1}{4} \overline{B_0^2 U^2 \cos^2 \varphi} = \frac{1}{8} B_0^2 \overline{U^2} = \frac{1}{4} B_0^2 \sigma_3^2. \quad (2.40)$$

Перевищення сигналу над завадою на виході синхронного детектора:

$$q_{\text{вих}} = \frac{b^2}{D\xi} = \frac{A_0^2}{\sigma_3^2} = 2q_{\text{вх}}. \quad (2.41)$$

Отримане співвідношення показує, що при когерентному детектуванні відношення сигнал–завада на виході лінійно залежить від відношення сигнал–завада на вході. Явище придушення сигналу завадою відсутнє. Лінійні властивості когерентного приймача визначають його високі показники завадостійкості. Переваги когерентного детектора в порівнянні з амплітудним особливо помітні при малих значеннях $q_{\text{вх}}$.

Так, якщо $q_{\text{вх}} = 0,1$, то при квадратичному детектуванні $q_{\text{вих}} = 0,01$, а при синхронному $q_{\text{вих}} = 0,2$.

Хоча наведене вище стосується гармонічного сигналу, отримані висновки можуть бути цілком поширені і на модульовані (маніпульовані) сигнали.

2.6 Приймання на погоджений фільтр

Існує великий клас задач, у яких потрібно знайти сигнал за його відомою формою. До їхнього числа належать задачі приймання телеграфних сигналів, сигналів при імпульсно-кодovій модуляції, радіолокаційних сигналів. У цих випадках важливим параметром, що характеризує якість виявлення, є відношення сигнал–завада. Лінійний фільтр, який максимізує це відношення, називається *оптимальним погодженим фільтром*.

Нехай на вході фільтра діє сума сигналу $s(t)$ і завади $\omega(t)$, тобто коливання $x(t) = s(t) + \omega(t)$. Корисний сигнал $s(t)$ розглядається не як випадковий процес, а як функція відомої форми; завада має спектральну щільність $G_3(\omega) = \frac{1}{2} N_0$.

Коефіцієнт передачі лінійного фільтра запишемо так:

$$K(i\omega) = K(\omega)e^{i\Psi(\omega)}. \quad (2.42)$$

Сигнал на виході фільтра дорівнює сумі корисного сигналу $y_c(t)$ і завади $y_3(t)$:

$$y(t) = y_c(t) + y_3(t).$$

Корисний сигнал на виході можна записати так:

$$y_c(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(i\omega)K(i\omega)e^{i\omega t} d\omega = \int_{-\infty}^{\infty} S(\omega)K(\omega)e^{i[\varphi(\omega)+\Psi(\omega)+\omega t]} d\omega,$$

де $S(i\omega)$ – спектральна щільність вхідного сигналу.

Пікова потужність сигналу в деякий момент t_0 дорівнюватиме:

$$P_{\text{с.п}} = |y_c(t_0)|^2 = \frac{1}{4\pi^2} \left| \int_{-\infty}^{\infty} S(i\omega)K(i\omega)e^{i\omega t_0} d\omega \right|^2,$$

а потужність завади:

$$P_3 = \frac{N_0}{4\pi} \int_{-\infty}^{\infty} K^2(\omega) d\omega.$$

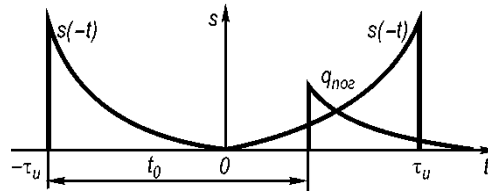


Рисунок 2.7 – Побудова імпульсної характеристики погодженого фільтра [7]

Тоді відношення сигнал–завада в момент часу t_0

$$q = \frac{P_{\text{с.п.}}}{P_3} = \frac{\left| \int_{-\infty}^{\infty} S(i\omega)K(i\omega)e^{i\omega t_0} d\omega \right|^2}{\pi N_0 \int_{-\infty}^{\infty} K^2(\omega) d\omega}. \quad (2.43)$$

Необхідно визначити, яким має бути коефіцієнт передачі фільтра, щоб відношення сигнал–завада q на його виході було максимальним. На підставі нерівності Буняковського–Шварца:

$$\left| \int_{-\infty}^{\infty} f_1(x)f_2(x)dx \right|^2 \leq \int_{-\infty}^{\infty} |f_1(x)|^2 dx \int_{-\infty}^{\infty} |f_2(x)|^2 dx \quad (2.44)$$

одержуємо:

$$q \leq \frac{\int_{-\infty}^{\infty} |S(i\omega)|^2 d\omega \int_{-\infty}^{\infty} |K(i\omega)|^2 d\omega}{\pi N_0 \int_{-\infty}^{\infty} |K(i\omega)|^2 d\omega} = \frac{1}{\pi N_0} \int_{-\infty}^{\infty} |S(i\omega)|^2 d\omega. \quad (2.45)$$

Отже, при будь-якій характеристиці фільтра $K(i\omega)$ максимальне відношення сигнал–завада:

$$q_{\text{max}} = \frac{1}{\pi N_0} \int_{-\infty}^{\infty} |S(i\omega)|^2 d\omega = \frac{2E}{N_0}, \quad (2.46)$$

де E – повна енергія сигналу.

Величина q_{max} досягається в тому випадку, коли коефіцієнт передачі фільтра:

$$K(i\omega) = cS(-i\omega)e^{-i\omega t_0} = cS(\omega)e^{-i[\omega t_0 + \varphi(\omega)]}, \quad (2.47)$$

де $S(-i\omega) = S(\omega)e^{-i\varphi(\omega)}$ – функція, комплексно-спряжена зі спектром сигналу $S(i\omega)$; c – довільна стала.

Вираз (2.47) можна записати у вигляді двох рівностей:

$$K(i\omega) = cS(\omega), \quad \varphi(\omega) = -[\varphi(\omega) + \omega t_0], \quad (2.48)$$

з яких випливає, що амплітудно-частотна характеристика погодженого фільтра з точністю до постійного множника збігається з амплітудним спектром сигналу, а фазочастотна характеристика визначається фазовим спектром сигналу $\varphi(\omega)$ і лінійною функцією частоти ωt_0 . Таким чином, частотна характеристика оптимального фільтра цілком визначається спектром сигналу, тобто «погоджена» з ним. Звідси й назва – *погоджений фільтр*.

Фаза сигналу на виході погодженого фільтра дорівнює:

$$\theta(t) = \omega t + \varphi(\omega) + \psi(\omega) = \omega t + \varphi(\omega) - \varphi(\omega) - \omega t_0 = \omega(t - t_0). \quad (2.49)$$

При $t = t_0$ $\theta = 0$, тобто в момент t_0 усі гармонічні складові сигналу мають однакову фазу і складаються арифметично, утворюючи в цей момент пік сигналу на виході фільтра. Спектральні ж складові завади на виході фільтра мають випадкову фазу. Цим і пояснюється те, що погоджений фільтр максимізує відношення сигнал–завада на виході.

Імпульсний відклик погодженого фільтра легко визначається на підставі перетворення Фур'є. Згідно з (2.47) маємо:

$$\begin{aligned} q(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} K(i\omega) e^{i\omega t} d\omega = \frac{c}{2\pi} \int_{-\infty}^{\infty} S(-i\omega) e^{-i\omega(t-t_0)} d\omega = \\ &= \frac{c}{2\pi} \int_{-\infty}^{\infty} S(i\omega) e^{i\omega(t-t_0)} d\omega = cs(t - t_0). \end{aligned} \quad (2.50)$$

Отже, імпульсним відкликом погодженого фільтра є дзеркальне відображення сигналу відносно t_0 в масштабі «с» (рис. 2.7). Значення t_0 не може бути менше моменту закінчення сигналу T . Це означає, що для практично реалізованого фільтра має виконуватися умова $t_0 \geq T$. Для зменшення часу аналізу доцільно взяти $t_0 = T$.

Погоджений фільтр як корелятор. Нехай $u_{\text{вх}}(t)$ – деякий вхідний сигнал, який у загальному випадку не збігається з сигналом $s(t)$ і стосовно якого розглянутий лінійний фільтр є погодженим. Відгук фільтра на даний вхідний вплив:

$$u_{\text{вх}}(t) = \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) q_{\text{пог}}(t - \tau) d\tau = k \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) s[t - (t - \tau)] d\tau = k \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) s[\tau - (t - t_0)] d\tau. \quad (2.51)$$

Останній інтеграл у (2.51) являє собою взаємкореляційну функцію сигналів $u_{\text{вх}}(t)$ і $s(t)$, тобто $u_{\text{вх}}(t) = kB_{us}(t - t_0)$.

У момент часу t_0 миттєве значення вихідного сигналу з точністю до коефіцієнта пропорційності дорівнює скалярному добутку обох сигналів:

$$u_{\text{вих}}(t) = k \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) s(\tau) d\tau. \quad (2.52)$$

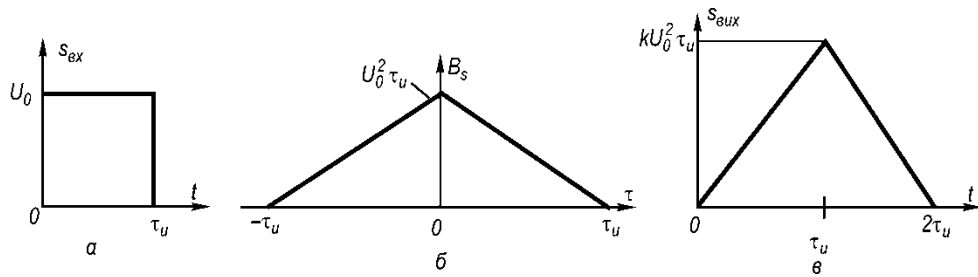


Рисунок 2.8 – Побудова сигналу на виході фільтра, погодженого з прямокутним відеоімпульсом: *a* – сигнал на вході; *б* – його автокореляційна функція; *в* – сигнал на виході, коли максимум вихідного коливання досягається в момент закінчення імпульсу на вході

Припустимо тепер, що $u_{\text{вх}} = s(t)$, тобто на вході фільтра присутній сигнал, стосовно якого цей фільтр погоджений. З формули (2.51) випливає, що в цьому випадку:

$$s_{\text{вих}}(t) = kB_s(t - \tau_0),$$

тобто вихідний сигнал пропорційний автокореляційній функції вхідного сигналу. Як приклад на рисунку 2.8 зображена побудова сигналу на виході конкретного погодженого фільтра.

Погоджений фільтр для прямокутного відеоімпульсу. Розглянемо імпульсний сигнал $s(t)$, що є відеоімпульсом прямокутної форми з відомою тривалістю τ_u і довільною амплітудою U_0 . Щоб знайти структуру фільтра, погодженого з таким сигналом, використовуємо спектральний метод. Насамперед обчислимо спектральну щільність корисного сигналу:

$$s(\omega) = \int_{-\infty}^{\infty} s(t)e^{-j\omega t} dt = U_0 \int_0^{\tau_u} e^{-j\omega t} dt = \frac{U_0}{j\omega} (1 - e^{-j\omega \tau_u}). \quad (2.53)$$

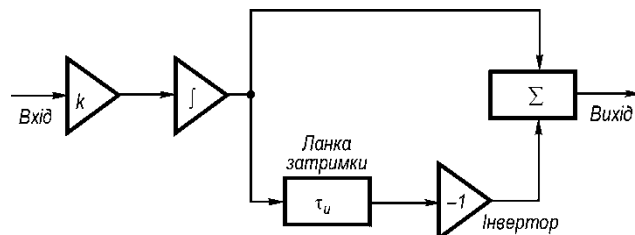


Рисунок 2.9 – Структурна схема погодженого фільтра для прямокутного відео імпульсу [8]

Звідси знаходимо частотний коефіцієнт передачі погодженого фільтра, взявши для конкретизації $t_0 = \tau_u$, тобто максимальність відклику фільтра в момент закінчення імпульсу:

$$K_{\text{пор}}(j\omega) = k \frac{1 - e^{-j\omega \tau_u}}{-j\omega} e^{-j\omega \tau_u} = \frac{k}{j\omega} (1 - e^{-j\omega \tau_u}). \quad (2.54)$$

Отриманий результат дозволяє синтезувати погоджений фільтр. Дійсно, відповідно до (2.54) такий фільтр має бути каскадним з'єднанням трьох лінійних

ланок: а) масштабного підсилювача з коефіцієнтом підсилення k ; б) ідеального інтегратора; в) пристрою з коефіцієнтом передачі $K'(j\omega) = 1 - \exp(-j\omega\tau_u)$. Останній пристрій реалізується за допомогою ланки затримки сигналу на час τ_u , інвертора, що змінює знак сигналу, і суматора. Структурна схема фільтра зображена на рисунку 2.9.

Погоджений фільтр для пачки однакових відеоімпульсів. У радіолокації часто, прагнучи збільшити енергію корисного сигналу, обробляють імпульси окремими пачками. Припустимо, що на виході амплітудного детектора приймача є пачка з N однакових відеоімпульсів тривалістю τ_u кожний; інтервал між імпульсами дорівнює T . Якщо $S_0(\omega)$ – спектральна щільність окремого імпульсу, то спектральна щільність пачки імпульсів:

$$S_n(\omega) = S_0(\omega) [1 + e^{-j\omega T} + e^{-j2\omega T} + \dots + e^{-j(N-1)\omega T}]. \quad (2.55)$$

Синтезуючи структуру погодженого фільтра для пачки імпульсів, поставимо вимогу, щоб максимальний відклик виникав у момент закінчення останнього імпульсу пачки, тобто $t_0 = (N-1)T + \tau_u$. Знаходимо частотний коефіцієнт передачі погодженого фільтра:

$$\begin{aligned} K_{\text{пор}}(j\omega) &= kS_0^*(\omega) e^{-j\omega t_0} [1 + e^{j\omega T} + e^{j2\omega T} + \dots + e^{j(N-1)\omega T}] e^{j(N-1)\omega T} = \\ &= K_{0\text{пор}}(j\omega) [1 + e^{-j\omega T} + e^{-j2\omega T} + \dots + e^{-j(N-1)\omega T}], \end{aligned} \quad (2.56)$$

де $K_{0\text{пор}}(j\omega)$ – коефіцієнт передачі погодженого фільтра для одиночного відеоімпульсу.

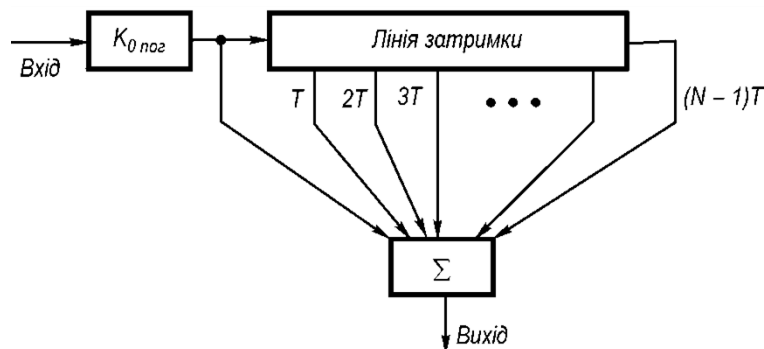


Рисунок 2.10 – Структурна схема погодженого фільтра для пачки відео імпульсі

Формула (2.56) безпосередньо визначає структурну схему погодженого фільтра, зображену на рисунку 2.10. На вході розміщений погоджений фільтр для одиночного відеоімпульсу. Основою пристрою служить багатовідвідна лінія затримки, що забезпечує запізнювання сигналів на відрізок часу $T, 2T, \dots, (N-1)T$. Сигнали з усіх відводів надходять у суматор. Максимальний відклик на виході суматора спостерігатиметься тоді, коли корисні сигнали від всіх імпульсів пачки одночасно з'являться на всіх його входах. Ефективність роботи пристрою тим вища, чим довша пачка.

Практично пристрої для виявлення радіолокаційних сигналів містять також спеціальний нелінійний граничний елемент, вхід якого з'єднаний з виходом

суматора погодженого фільтра. Рівень порога дещо перевищує середньоквадратичне значення шуму за відсутності корисного сигналу.

Якщо сплеск вихідного сигналу фільтра досягає граничного рівня, на пристрій індикації надходить керуючий сигнал, який свідчить про наявність імпульсу, відбитого від цілі.

Погоджений фільтр для прямокутного радіоімпульсу. Нехай сигнал являє собою радіоімпульс вигляду:

$$s(t) = \begin{cases} 0 & \text{при } t < 0, \\ U_0 \sin \omega_0 t & \text{при } 0 \leq t \leq \tau_u, \\ 0 & \text{при } t > \tau_u. \end{cases} \quad (2.57)$$

Синтезуємо погоджений фільтр для такого сигналу, використовуючи відомості про імпульсну характеристику фільтра.

Як було показано, імпульсна характеристика погодженого фільтра $q_{\text{пор}}(t) = ks(t_0 - t)$. Візьмемо $t_0 = \tau_u$ і вважатимемо для простоти, що тривалість імпульсу є кратною періоду високочастотного заповнення: $\sin \omega_0 \tau_u = 0$.

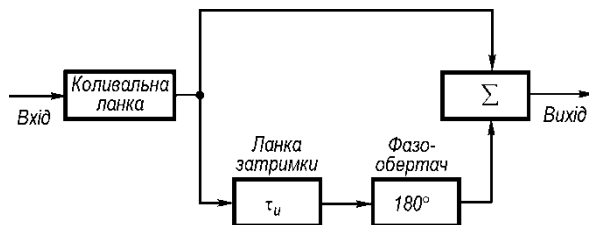


Рисунок 2.11 – Структурна схема погодженого фільтра для прямокутного радіоімпульсу [8]

Тоді

$$q_{\text{пор}}(t) = \begin{cases} 0 & \text{при } t < 0, \\ k \sin \omega_0 t & \text{при } 0 \leq t \leq \tau_u, \\ 0 & \text{при } t > \tau_u, \end{cases} \quad (2.58)$$

тобто імпульсна характеристика погодженого фільтра з точністю до амплітудного множника повторює вхідний сигнал.

Таку імпульсну характеристику можна наближено реалізувати за допомогою системи, структурна схема якої наведена на рисунку 2.11. На вході фільтра розміщується коливальна ланка (наприклад, високоякісний коливальний контур) з імпульсною характеристикою:

$$q_{\text{пор}}(t) = \begin{cases} 0 & \text{при } t < 0, \\ b \sin \omega_0 t & \text{при } t \geq 0, \end{cases}$$

де b – стала величина.

Для того щоб імпульсна характеристика погодженого фільтра дорівнювала нулю при $t > \tau_u$, передбачені суматор і фазообертач. На один із входів суматора сигнал з виходу коливальної ланки подається безпосередньо, а на іншій – через ланку затримки на τ_u секунд. Фазообертач змінює фазу сигналу на 180° . У разі

такого вмикання елементів, починаючи з моменту часу $t = \tau_u$, до входів суматора прикладені два гармонічні коливання з однаковими амплітудами і протилежними фазами, що перетворює в нуль сигнал на виході суматора.

Погоджений фільтр для сигналу Баркера. Достоїнство сигналів Баркера – високе значення головної пелюстки автокореляційної функції і гранично низький рівень бічних пелюсток.

На рисунку 2.12 зображена структурна схема погодженого фільтра, призначеного для виявлення M -позиційного сигналу Баркера з фазовим кодуванням. Такий сигнал виглядає як послідовність відрізків гармонічних коливань з фазовими зсувами $\varphi_1, \varphi_2, \dots, \varphi_M$, рівними 0 чи 180° .

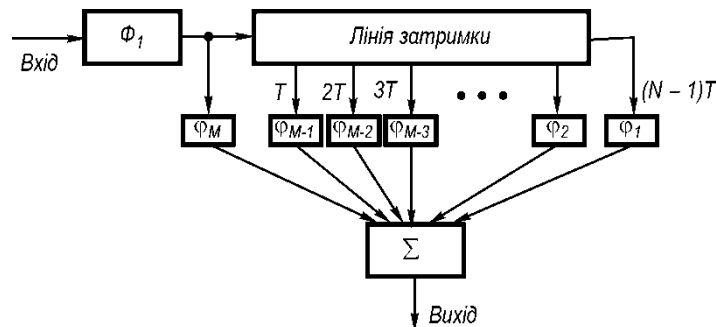


Рисунок 2.12 – Структурна схема погодженого фільтра для сигналу Баркера [8]

При синтезі погодженого фільтра виходять з того, що його імпульсна характеристика має бути «дзеркальною» копією сигналу, котрий виділяють з оберненим у часі порядком проходження окремих позицій.

На вході пристрою є допоміжний фільтр Φ_1 , погоджений до однієї позиції складного фазоманіпульованого сигналу, тобто до прямокутного радіоімпульсу. На виході фільтра під впливом вхідного дельта-імпульсу виникає радіоімпульс із обвідної прямокутної форми. Цей імпульс подається на лінію затримки з відводами, яка являє собою хвильову (розподілену) систему. Затримка в часі між відводами дорівнює тривалості T кожної позиції сигналу.

Для правильного функціонування пристрою необхідно, щоб послідовність фазових зсувів $\varphi_M, \varphi_{M-1}, \dots, \varphi_1$ відповідала значенням фаз в окремих позиціях сигналу Баркера, рахуючи від кінця сигналу до початку.

Прямокутний радіоімпульс, який переміщається вздовж лінії затримки, по черзі збуджує входи суматора, на виході якого виникає «дзеркальна» копія сигналу, що виділяється.

Контрольні запитання для самооцінки рівня знань

1. Які основні задачі виникають при прийманні сигналів?
2. У чому полягає процес приймання сигналів із накопиченням?
3. Які різновиди методу накопичення Ви знаєте?
4. У чому полягає суть кореляційного приймання сигналів?
5. Чим автокореляційний метод приймання сигналів відрізняється від кореляційного, а когерентний – від некогерентного?
6. Як змінюється відношення сигнал–завада на виході детектора при різних значеннях відношення сигнал–завада на вході?

7. Який фільтр називається погодженим?
8. Яке максимальне відношення сигнал–завада на виході погодженого фільтра?
9. Нарисуйте структурні схеми погодженого фільтра для одиночного відеоімпульсу та для прямокутного радіоімпульсу.
10. Нарисуйте структурну схему погодженого фільтра для сигналу Баркера.

3 КОДУВАННЯ ДЖЕРЕЛ ІНФОРМАЦІЇ ТА КАНАЛІВ ЗВ'ЯЗКУ

3.1 Основні поняття та означення

Перетворення дискретного повідомлення в сигнал складається з двох операцій: кодування і модуляції. *Кодування* визначає закон побудови сигналу, а *модуляція* – вид формованого сигналу, що має передаватися каналом зв'язку.

Найпростішим прикладом дискретного повідомлення є текст. Будь-який текст складається зі скінченного числа елементів: букв, цифр, розділових знаків. Для європейських мов число елементів коливається від 52 до 55, для східних – може обчислюватися сотнями і навіть тисячами. Оскільки число елементів у дискретному повідомленні скінченне, то їх можна пронумерувати і тим самим звести передавання повідомлень до передавання послідовності чисел.

Множина знаків (символів) і система правил побудови складених знаків називається *кодом*. Кінцева послідовність кодових знаків називається *словом* або *ковою комбінацією*. З цього погляду алфавіт української або іншої мови і сукупність правил побудови слів є кодом, за допомогою якого усне мовлення перетворюється в текст. Слова (кодові комбінації) можуть розглядатися як символи деякої іншої множини; разом із установленими правилами утворення речень вони теж утворюють код, яким ми користуємося для обміну інформацією. Коди поділяються:

- за числом m символів у множині – на двійкові ($m=2$) і недвійкові ($m \neq 2$). Залежно від числа символів у множині коди називають, наприклад, трійковими, вісімковими. Кількість символів у множині є *основою коду*;

- за довжиною кодових комбінацій – на нерівномірні, у яких слова мають нерівну кількість символів, і рівномірні – з однаковою кількістю символів у словах;

- за принципом використання кодових комбінацій – на коди з використанням усіх дозволених комбінацій заданої довжини n для подання повідомлень і коди з використанням частини можливих комбінацій. Останні у свою чергу поділяються на коди, що дають змогу виявляти або виправляти спотворені символи (помилки), які виникають через завади, – коректувальні коди, і на коди, що не дають такої можливості.

У загальному випадку процес перетворення повідомлення в кодову комбінацію прийнято називати кодуванням. При цьому мається на увазі взаємооднозначне перетворення. В окремому випадку кодування можна визначити як операцію встановлення однозначної відповідності між символами групи символів деякого коду з символами групи символів іншого коду. Таке кодування є переведенням з однієї системи числення в іншу (наприклад, десяткових чисел у двійкові). Слова «код» і «кодування» походять від латинського слова *codex* – кодекс, тобто книга, що містить знаки, систему правил.

Повідомлення, які є текстом природною мовою, мають значний надлишок, тому для його зменшення використовують кодування, що враховує статистичні особливості повідомлення. Для зменшення надлишку доцільно кодувати повідомлення так, щоб середня довжина L кодових комбінацій була найменшою, тобто здійснювати *статистичне кодування*. Одним з кодів, що використовують статистичні властивості повідомлень, є код *Шеннона – Фано*.

Зазначимо, що поліпшення якісних характеристик каналів для підвищення надійності приймання завжди пов'язано зі значними матеріальними витратами, а

іноді й дуже високими. Тому величезне значення мають широко застосовувані в техніці передавання методи підвищення надійності приймання, які не потребують поліпшення якості каналу. Ці методи засновані на внесенні в переданий сигнал значного надлишку. Надлишок, що вводиться в переданий сигнал, накладає на нього певні додаткові умови, перевірка дотримання яких при прийманні дає змогу встановити факт спотворення сигналу, а також ототожнити прийнятий спотворений сигнал з відповідним неспотвореним.

Будь-які методи внесення надлишку в переданий сигнал пов'язані зі збільшенням об'єму сигналу, тобто зі збільшенням або потужності сигналу, або ширини спектра, або часу передавання. Можливості підвищення надійності приймання збільшенням потужності і ширини спектра при передаванні дискретної інформації зі стандартних каналів зв'язку досить обмежені, тому переважне застосування дістав метод уведення надлишку *збільшенням часу передавання сигналу*. Цей метод можна реалізувати у двох різновидах – використанням для передавання дискретної інформації, зниженої стосовно номінального значення швидкості, і застосуванням коректувальних кодів. *Зниження швидкості передавання інформації* по каналах невисокої якості набуло широкого застосування. Так, у багатьох типах апаратури передавання даних (АПД), що працюють по каналах тональної частоти (ТЧ), передбачаються дві і більше швидкості модуляції (наприклад, 600 і 1200 бод), причому з меншою швидкістю звичайно ведеться передавання на великі відстані, а також робота по каналах зі звуженою смугою пропускання (300...2700 Гц).

Застосування коректувальних кодів є ефективнішим методом підвищення надійності, ніж застосування зниженої швидкості. Зазначимо, що обидва методи не слід протиставляти один одному: зниження швидкості передавання можна вважати окремим випадком застосування коректувального коду.

Розглянемо деякі поняття, пов'язані з коректувальними кодами [5].

Звичайний (простий) код характеризується тим, що окремі його кодові комбінації відрізняються одна від одної лише одним розрядом. Тому навіть один помилково прийнятий розряд спричинює заміну однієї кодової комбінації іншою і, отже, неправильне приймання повідомлення. *Коректувальні (надлишкові, завадостійкі) коди* будують так, що для передавання інформації використовується лише частина кодових комбінацій (дозволені комбінації), які відрізняються одна від одної більше ніж одним розрядом. Усі інші комбінації для передавання не використовуються і належать до числа недозволених (заборонених). Таким чином, при застосуванні коректувальних кодів помилка в одному розряді спричинює заміну дозволеної кодової комбінації недозволеною, що дає змогу знайти помилку. При досить великій відмінності дозволених комбінацій одна від одної можливе виявлення дворазової, триразової і т. д. помилки, оскільки вони зумовляють утворення недозволених комбінацій, а перехід однієї дозволеної комбінації в іншу відбуватиметься під впливом помилок більшої кратності, що є результатом найбільш інтенсивних завад.

Пояснимо це на прикладі. Використаємо для передавання інформації чотирирозрядні кодові комбінації, що відрізняються одна від одної не менш ніж двома розрядами: 0011, 0110, 1001, 1010, 1100, 1111, 0101, 0000. Нехай при передаванні кожної з цих комбінацій (наприклад, 0011) сталася одинична помилка, внаслідок чого спотворився перший розряд, і прийнято комбінацію 1011. Ця

комбінація є недозвеною, що свідчить про наявність помилки. Підберемо далі чотирирозрядні комбінації, що відрізняються всіма чотирма розрядами: 0011 і 1100. Легко переконатися, що при використанні цих комбінацій виявляються одно-, дво- і триразові помилки, а не виявляється лише чотириразова помилка.

Цей самий код, що складається з двох кодових комбінацій, може використовуватися і для виправлення одиничних помилок. Нехай, наприклад, прийнято комбінацію 1011. Ця комбінація відрізняється від дозволеної комбінації 0011 одним розрядом, а від іншої дозволеної 1100 – трьома. Отже, прийнята комбінація «ближча» до комбінації 0011, ніж до комбінації 1100; це дає підставу вважати, що була передана комбінація 0011.

Бачимо, що підвищена завадостійкість двох розглянутих кодів пов'язана з наявним в них надлишком. Так, перший код містить вісім комбінацій, що становлять чотири розряди кожна. В той самий час у простому коді для утворення восьми комбінацій достатньо трьох розрядів, а не чотирьох. Отже, підвищення завадостійкості потребує введення додаткового розряду. Другий код має ще більшу завадостійкість, і це потребує ще більшого надлишку – трьох додаткових розрядів.

Коректувальні коди так само, як і прості, можуть бути *рівномірними* або *нерівномірними*, *двійковими* або *багатопозиційними*. Використання нерівномірних або багатопозиційних кодів спричинює значне ускладнення апаратури передавання даних, тому застосовуються вони дуже рідко. У зв'язку з цим надалі розглядатимемо лише двійкові рівномірні коректувальні коди. Останні поділяються на два класи – блокові і неперервні (рис. 3.1). При використанні блокових кодів передана інформаційна послідовність розбивається на окремі кодові комбінації (блоки), що кодуються і декодуються незалежно одна від одної. Неперервні коди – це неперервна послідовність розрядів, і поділ її на окремі блоки неможливий. Блокові коди, у свою чергу, поділяються на роздільні та нероздільні.

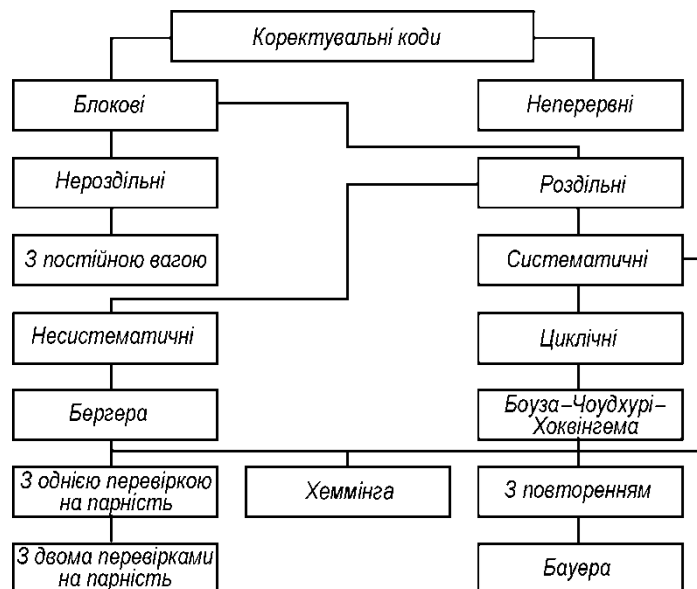


Рисунок 3.1 – Класифікація коректувальних кодів [5]

Роздільними називаються коди, в яких одні розряди є інформаційними, інші – перевірними. Останні і вносять у код надлишок, необхідний для виявлення або

виправлення помилок. У роздільних кодах інформаційні і перевірні розряди займають завжди одні й ті самі позиції в кодовій комбінації. Роздільні коди позначаються як (n, k) -коди, де n – довжина або число розрядів коду; k – число інформаційних розрядів.

Нероздільні коди утворюють зараз нечисленну групу. До них, зокрема, належать рекомендований МСЕ стандартний телеграфний код № 3 – семирозрядний код, кожна кодова комбінація якого містить три одиниці і чотири нулі.

Серед роздільних кодів розрізняють коди систематичні і несистематичні. *Систематичними* називаються такі блокові роздільні (n, k) -коди, в яких перевірні розряди – це лінійні комбінації інформаційних. Систематичні коди утворюють велику групу кодів і дуже широко застосовуються на практиці. Тому головну увагу приділяють саме цим кодам і, зокрема, їх найбільш відомим різновидам – кодам Хеммінга і циклічним.

Кількість розрядів, якими різняться дві кодові комбінації, називається *кодовою відстанню* між двома комбінаціями. Так, кодова відстань між комбінаціями 11011 і 00010 дорівнює трьом, оскільки вони різняться трьома розрядами – першим, другим і п'ятим. Найменша з кодових відстаней у коді називається *мінімальною кодовою* або *хеммінговою відстанню*. Наприклад, у трирозрядному коді з дозволеними комбінаціями 101, 110, 011, 000 мінімальна кодова відстань d_0 дорівнює двом, для простих кодів – одиниці.

Мінімальна кодова відстань d_0 пов'язана з кількістю помилок, що виявляються, і кількістю помилок, що виправляються, так:

$$d_0 \geq \sigma + 1, \quad (3.1)$$

$$d_0 \geq 2t + 1, \quad (3.2)$$

де σ – кількість помилок, що виявляються; t – кількість помилок, що виправляються.

Коректувальні коди можна використовувати для виправлення помилок і одночасного виявлення помилок більшої кратності. Можна показати, що при цьому (коли $\sigma > t$):

$$d_0 \geq \sigma + t + 1. \quad (3.3)$$

Мінімальна кодова відстань лише частково характеризує коректувальні властивості коду, оскільки звичайно забезпечується виправлення і виявлення помилок і більш високої кратності, ніж обумовлене співвідношеннями (3.1)–(3.3).

Загальна кількість комбінацій коду завдовжки n дорівнює 2^n . Число дозволених кодових комбінацій визначається числом інформаційних розрядів і дорівнює $M = 2^k = 2^{n-r}$, де r – число перевірних розрядів. Отже, число дозволених кодових комбінацій у 2^r рази менше загального числа комбінацій.

Надлишок коду називають відношення r/n . Питання про мінімально необхідний надлишок коду при заданих мінімальних кодовій відстані і довжині коду в загальному випадку не вирішено. Існує лише ряд верхніх і нижніх оцінок. Найбільше наближення звичайно забезпечує критерій Варламова:

$$r \geq \log_2 \left(1 + \sum_{i=1}^{d_0-2} C_{n-1}^i \right), \quad (3.4)$$

де C_{n-1}^i – число сполучень з $n-1$ елементів по i елементах.

Для деяких кодів отримано точні залежності між числом перевірних і інформаційних розрядів при заданій мінімальній кодовій відстані. Так, при $d_0 = 3$ має місце співвідношення:

$$r \geq \log_2(n+1), \quad (3.5)$$

причому r – найменше ціле число, при якому задовольняється нерівність (3.5).

3.2 Коректувальні коди

Принцип побудови систематичних кодів. Нехай комбінація $a_1 a_2 \dots a_k b_1 b_2 \dots b_r$, де $a_1 a_2 \dots a_k$ – інформаційні, а $b_1 b_2 \dots b_r$ – перевірні розряди, є дозволеною комбінацією систематичного (n, k) -коду. В систематичних кодах перевірні розряди є лінійною комбінацією інформаційних. Це означає, що значення будь-якого перевірного розряду $b_i = c_{i1} a_1 \oplus c_{i2} a_2 \oplus \dots \oplus c_{ik} a_k$, де $c_{i1}, c_{i2}, \dots, c_{ik}$ – числа, які дорівнюють 0 або 1. Нульова (тобто така, що складається з одних нулів) комбінація в будь-якому систематичному коді є дозволеною, оскільки лінійною комбінацією нулів є нуль.

Складемо за модулем 2 дві дозволені кодові комбінації систематичного коду:

$$\begin{aligned} & a_1 a_2 \dots a_k b_1 b_2 \dots b_r \oplus a'_1 a'_2 \dots a'_k b'_1 b'_2 \dots b'_r = \\ & = (a_1 \oplus a'_1)(a_2 \oplus a'_2) \dots (a_k \oplus a'_k)(b_1 \oplus b'_1)(b_2 \oplus b'_2) \dots (b_r \oplus b'_r), \end{aligned}$$

де

$$\begin{aligned} b_i \oplus b'_i & = c_{i1} a_1 \oplus c_{i2} a_2 \oplus \dots \oplus c_{ik} a_k \oplus c_{i1} a'_1 \oplus c_{i1} a'_2 \oplus \dots \oplus c_{ik} a'_k = \\ & = c_{i1} (a_1 \oplus a'_1) \oplus c_{i2} (a_2 \oplus a'_2) \oplus \dots \oplus c_{ik} (a_k \oplus a'_k). \end{aligned}$$

Як бачимо, перевірні розряди суми за модулем 2 двох дозволених комбінацій утворюються за тим самим правилом, що і для кожної дозволеної комбінації. Звідси сума двох дозволених комбінацій систематичного коду також є дозволеною комбінацією.

Таке положення є правдивим і при підсумовуванні за модулем 2 будь-яких кількостей дозволених комбінацій систематичного коду, що дає можливість визначити всі дозволені кодові комбінації, користуючись лише їх обмеженою кількістю.

Усі дозволені комбінації систематичного коду можна визначити так. Виберемо кілька (g) ненульових дозволених кодових комбінацій. Складемо їх у деяких сполученнях: по дві, по три, по чотири, ..., по g комбінацій. Кожне додавання дає нам нову дозволена комбінація.

Процес побудови множини кодових комбінацій можна зобразити як $x_i = c_1 x_1 \oplus c_2 x_2 \oplus \dots \oplus c_g x_g$, де x_1, x_2, \dots, x_g – початкові кодові комбінації; c_1, c_2, \dots, c_g – коефіцієнти, що набувають значення 1 або 0. Надаючи різним коефіцієнтам c_i значення 1 або 0, можна скласти вихідні комбінації в різних сполученнях. Загальна кількість отриманих комбінацій становитиме $c_g^2 + c_g^3 + \dots + c_g^g$. Враховуючи g вихідних кодових комбінацій і одну нульову, можна прийти до висновку, що загальна кількість дозволених комбінацій дорівнює:

$$c_g^0 + c_g^1 + c_g^2 + \dots + c_g^g = 2^k, \quad (3.6)$$

оскільки систематичний код містить 2^k дозволених кодових комбінацій. Ця умова задовольняється при $g = k$.

Співвідношення (3.6) є слушним тоді, коли комбінації, отримані внаслідок додавань, не збігаються ні з початковими комбінаціями, ні з нульовою. Для цього початкові комбінації треба вибирати так:

– усі початкові комбінації мають бути різні; у протилежному випадку при додаванні виходитиме нульова комбінація;

– нульова комбінація не повинна входити в число початкових, оскільки внаслідок додавання нульової комбінації і якої-небудь початкової виходитиме та сама початкова комбінація;

– усі початкові кодові комбінації мають бути лінійно незалежні, тобто має виконуватися рівність $c_1x_1 \oplus c_2x_2 \oplus \dots \oplus c_kx_k = 0$ при всіх значеннях c_i , за винятком $c_1 = c_2 = \dots = c_k = 0$. У протилежному разі похідні кодові комбінації збігатимуться з початковими;

– кожна початкова кодова комбінація, як і будь-яка ненульова дозволена комбінація, повинна мати вагу не меншу ніж d_0 . (Нагадаємо, що вагою кодової комбінації називається кількість одиниць у ній.) Слушність цієї умови легко зрозуміти, якщо врахувати, що нульова комбінація також є дозволеною, а кодова відстань між парою будь-яких кодових комбінацій коду не повинна бути меншою, ніж d_0 ;

– кодова відстань між будь-якими парами початкових комбінацій не повинна бути меншою, ніж d_0 .

Підібрані певним чином k початкових кодових комбінацій однозначно визначають систематичний код. Початкові комбінації прийнято записувати одна під одною у вигляді таблиці (матриці) із k рядків і n стовпців, яка називається *породжувальною матрицею*:

$$G = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{21} & a_{22} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Звичайно інформаційні розряди займають перші k позицій у кодовій комбінації. При цьому породжувальну матрицю зручно будувати, починаючи з одиничної матриці E_k , що має k стовпців і k рядків:

$$E_k = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix};$$

праворуч до неї приписується матриця $C_{r;k}$, що має r стовпців і k рядків:

$$G = [E_k, C_{r;k}] = \begin{pmatrix} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1r} \\ 0 & 1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2r} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & b_{k1} & b_{k2} & \dots & b_{kr} \end{pmatrix}. \quad (3.7)$$

Матриця G при належним чином підібраній матриці $C_{r;k}$ також є твірною. Рядки приписаної матриці $C_{r;k}$ знаходять перебором різних r -розрядних комбінацій, що містять не менше ніж $d_0 - 1$ одиниць, при цьому сума за модулем 2 двох будь-яких рядків матриці $C_{r;k}$ не повинна мати менше, ніж $d_0 - 2$ одиниць. Матриця G записана в канонічній формі [10].

Розглянемо тепер приклад побудови систематичного коду. Нехай треба побудувати код $n=7$, що забезпечує виправлення одиничної помилки. Відповідно до формули (3.2) $d_0 = 3$. Користуючись нерівністю (3.5), знайдемо $r=3$. Відповідно $k=4$. Побудуємо твірну матрицю коду $(7, 4)$:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Рядки приписаної частини твірної матриці (обведено штриховою лінією) містять не менше, ніж дві одиниці. Сума за модулем 2 двох будь-яких рядків приписаної матриці не менша, ніж 1.

Оскільки $k=4$, то код $(7, 4)$ має 16 дозволених комбінацій: перші чотири комбінації є рядками породжувальної матриці, п'ята – нульової, інші одинадцять комбінацій знайдемо підсумовуванням за модулем 2 різних сполучень рядків породжувальної матриці:

$$\begin{array}{ll} 1 \oplus 2 & 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \\ 1 \oplus 3 & 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \\ 1 \oplus 4 & 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \\ 2 \oplus 3 & 0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \\ 2 \oplus 4 & 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \\ 3 \oplus 4 & 0 \ 0 \ 1 \ 1 \ 0 \ 0 \ 1 \\ 1 \oplus 2 \oplus 3 & 1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \\ 2 \oplus 3 \oplus 4 & 0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \\ 1 \oplus 3 \oplus 4 & 1 \ 0 \ 1 \ 1 \ 0 \ 1 \ 0 \\ 1 \oplus 2 \oplus 4 & 1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \\ 1 \oplus 2 \oplus 3 \oplus 4 & 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \end{array}$$

Неважко переконатися, що побудований код має мінімальну кодову відстань, яка дорівнює 3.

У систематичному коді процес кодування зводиться до визначення r перевірних розрядів на основі відомих k інформаційних. Кожен перевірний розряд визначається за допомогою *перевірного співвідношення*, а визначення всіх r перевірних розрядів вимагає r перевірних співвідношень. Перевірні співвідношення прийнято записувати одне під одним у вигляді таблиці (матриці), яку називають *перевірною матрицею*. Перевірна матриця містить r рядків і n стовпців, причому кожен рядок є перевірним співвідношенням для знаходження значення одного з перевірних розрядів. Перевірна матриця утворюється так: будується одинична матриця:

$$E_r = \begin{vmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{vmatrix},$$

після чого до неї ліворуч приписується матриця $D_{k;r}$, що містить k стовпців і r рядків, причому кожен її рядок відповідає стовпцю перевірних розрядів породжувальної матриці в канонічній формі:

$$H = \left| D_{k;r}, E_r \right| = \begin{vmatrix} b_{11} & b_{21} & \dots & \dots & \dots \\ b_{12} & b_{22} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ b_{1r} & b_{2r} & \dots & \dots & \dots \end{vmatrix}. \quad (3.8)$$

Матриця (3.8) є перевіркою. За її допомогою операція кодування здійснюється дуже просто. Позиції, займані одиницями в i -му рядку приписаної частини перевіркою матриці, визначають ті інформаційні розряди, що мають брати участь у формуванні i -го перевірного розряду. Наприклад, якщо перший рядок перевіркою матриці має вигляд 1011100, то перевірений розряд визначиться зі співвідношення $a_1 \oplus a_3 \oplus a_4 \oplus a_5 = 0$.

Декодування також зручно здійснювати за допомогою перевірних матриць; при цьому виконується r перевірок на парність відповідно до (2.8). Якщо хоча б одна з перевірок не дорівнює 0, то це означає, що в прийнятій комбінації є помилки.

Перевірку кодової комбінації при прийманні можна виконати зіставленням прийнятих перевірних розрядів і перевірних розрядів, обчислених на підставі прийнятих інформаційних. Їх сума за модулем 2 називається *синдромом*. Характерною рисою синдрому є те, що він не залежить від виду переданої комбінації, а цілком визначається помилками прийнятої комбінації. Між синдромом і комбінацією, що спричинила помилки, немає взаємно однозначної відповідності – тому самому синдрому відповідають 2^k різних комбінацій помилок. Так, нульовому синдрому відповідає нульова комбінація помилок (тобто безпомилкове приймання), а також інші $2^k - 1$ комбінації помилок, що збігаються з дозволеними кодовими комбінаціями (невиявлені помилки). Тільки одну з комбінацій помилок, що відповідають ненульовому синдрому, можна виправити кодом. При цьому за кожним синдромом закріплюється така виправна комбінація помилок, поява якої в каналі найбільш імовірна. Зазначимо, що синдром збігається з комбінацією результатів перевірки на парність. Це легко перевірити практично.

Для побудованого раніше коду (7, 4) перевірна матриця:

$$H = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix}. \quad (3.9)$$

Тому

$$\begin{aligned} a_5 &= a_2 \oplus a_3 \oplus a_4; \\ a_6 &= a_1 \oplus a_3 \oplus a_4; \\ a_7 &= a_1 \oplus a_2 \oplus a_4, \end{aligned}$$

де a_1, \dots, a_4 – інформаційні розряди; a_5, a_6, a_7 – перевірні.

Декодування відбувається обчисленням перевірних співвідношень:

$$\begin{aligned} a_2 \oplus a_3 \oplus a_4 \oplus a_5 \\ a_1 \oplus a_3 \oplus a_4 \oplus a_6; \\ a_1 \oplus a_2 \oplus a_4 \oplus a_7. \end{aligned} \quad (3.10)$$

Якщо код використовується для виправлення помилок (а даний код може виправити одну помилку), то на приймальному боці має бути заздалегідь визначена відповідність між видом синдрому і видом однократної помилки, що виправляється. Встановимо цю відповідність. Нехай помилка є в першому розряді (комбінація помилки 1000000). Перевіримо цю комбінацію з урахуванням (3.10):

$$\begin{aligned} a_2 \oplus a_3 \oplus a_4 \oplus a_5 &= 0 \oplus 0 \oplus 0 \oplus 0 = 0, \\ a_1 \oplus a_3 \oplus a_4 \oplus a_6 &= 1 \oplus 0 \oplus 0 \oplus 0 = 1, \\ a_1 \oplus a_2 \oplus a_4 \oplus a_7 &= 1 \oplus 0 \oplus 0 \oplus 0 = 1. \end{aligned}$$

Отже, при наявності помилки в 1-му розряді синдром дорівнює 011. Аналогічно можна одержати види синдромів і у разі наявності всіх інших можливих однократних помилок:

Помилка в розряді	1	2	3	4	5	6	7
Синдром	001	101	110	111	100	010	001

Нехай, наприклад, у кодувальній пристрій надійшла інформаційна послідовність 1101. Згідно з (3.9) ця послідовність кодується так: 1101001. Тепер припустимо, що в процесі передавання сталася помилка в 2-му розряді і комбінація 1101001 прийнята як 1001001. При декодуванні відповідно до (3.10) будуть отримані такі результати:

$$\begin{aligned} 1\text{-ша перевірка } & 0 \oplus 0 \oplus 1 \oplus 0 = 1, \\ 2\text{-га } & \gg \quad 1 \oplus 0 \oplus 1 \oplus 0 = 1, \\ 3\text{-тя } & \gg \quad 1 \oplus 0 \oplus 1 \oplus 0 = 1. \end{aligned}$$

Отже, синдром дорівнює 101, що свідчить про те, що помилковим є саме 2-й розряд у прийнятій комбінації.

Використання систематичних кодів. Одним із найпростіших систематичних кодів, що одержали практичне застосування, є код з повторенням, який має два різновиди. В одному з них є S -кратне повторення комбінації простого коду a_1, a_2, \dots, a_k :

$$\underbrace{a_1 a_1}_{1} \underbrace{a_2 a_2}_{2} \dots \underbrace{a_k a_k}_{S}$$

Інший різновид коду з повторенням характеризується S -кратним передаванням кожного розряду:

$$\underbrace{a_1 a_1 a_1}_{S \text{ разів}} \underbrace{a_2 a_2 a_2}_{S \text{ разів}} \dots \underbrace{a_k a_k a_k}_{S \text{ разів}}.$$

Код з повторенням має довжину $n = S_k$, число перевірних розрядів $r = k(S-1)$, мінімальну кодову відстань $d_0 = S$. Надлишок цих кодів дорівнює $(S-1)/S$. Звичайно застосовується перший різновид коду з повторенням, що має в умовах корельованих помилок підвищену завадостійкість. Це зумовлено тим, що вхідні в

одну перевірку на парність розряди досить далеко стоять один від одного і з малою ймовірністю вражаються одним пакетом помилок. Число повторень звичайно дорівнює 2 ($d_0 = 2$) і набагато рідше 3 ($d_0 = 3$). Велика кратність повторень практично не використовується.

Код з повторенням характеризується досить високими властивостями, що проявляються при дії пакетів помилок. Так, при $S=2$ завжди виявляються пакети помилок завдовжки до $n/2$, а також усі помилки непарної кратності. При $S=2$ перевірна матриця має вигляд:

$$H = |E_k, E_k|. \quad (3.11)$$

Недоліком кодів з повторенням є дуже великий надлишок. Навіть при дворазовому повторенні він становить 0,5.

Існує особливий різновид коду з дворазовим повторенням, що забезпечує вдвічі більшу мінімальну кодову відстань ($d_0 = 4$). Це так званий *код Бауера* (інверсний код). При використанні даного коду комбінації з парним числом одиниць повторюються в незмінному вигляді, а комбінації з непарним числом одиниць – в інвертованому. Перевірна матриця:

$$H = |\overline{E}_k, E_k|, \quad (3.12)$$

де \overline{E}_k – матриця, отримана з одиничної матриці E_k заміною одиниць нулями, а нулів – одиницями.

При невеликих довжинах кодових комбінацій (до 10–14) інверсний код за завадостійкістю має такий самий надлишок, що і складніші коди.

Код з парним числом одиниць (код з однією перевіркою на парність, паритетний код), так само, як і код із дворазовим повторенням, забезпечує $d_0 = 2$, але має набагато менший надлишок. Незалежно від довжини кодової комбінації цей код має один перевірний розряд, значення якого вибирають з умови одержання парного числа одиниць у кодовій комбінації, тобто він визначається зі співвідношення:

$$b_1 = a_1 \oplus a_2 \oplus \dots \oplus a_k.$$

Перевірна матриця коду з парним числом одиниць містить лише один рядок:

$$H = |111\dots 1|. \quad (3.13)$$

Як і код з дворазовим повторенням, розглянутий код виявляє не тільки однократні помилки, але і взагалі всі помилки непарної кратності. Не слід, однак, думати, що обидва коди мають однакову завадостійкість. У разі дії пакетних помилок завадостійкість коду з дворазовим повторенням набагато вища.

Існує також *код із двома перевірками на парність*. Незалежно від довжини кодової комбінації цей код має два перевірних розряди, один із яких вибирають з умови парності всіх інформаційних розрядів, а другий – з умови парності всіх непарних (або парних) за номером інформаційних розрядів. Так, при $k=5$ значення перевірних розрядів:

$$b_1 = a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5, \quad b_2 = a_1 \oplus a_3 \oplus a_5.$$

Як і код з парним числом одиниць, розглянутий код має мінімальну кодову відстань $d_0 = 2$, але більшу завадостійкість, оскільки виявляє частину помилок парної кратності – суміжних, тобто розташованих поруч.

Перевірні матриця коду з двома перевірками на парність має вигляд:

$$H = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & \dots \\ 1 & 0 & 1 & 0 & 1 & \dots \end{vmatrix}.$$

Коди Хеммінга. Кодами Хеммінга називають коди з мінімальною кодовою відстанню $d_0 = 3$, що виправляють усі одиничні помилки, і коди з відстанню $d_0 = 4$, що виправляють усі одиничні й усі подвійні помилки. Коди Хеммінга з мінімальною кодовою відстанню $d_0 = 3$ мають довжину $n \leq 2^r - 1$. Відповідно:

$$2^k \leq \frac{2^n}{n+1}. \quad (3.14)$$

Характерною рисою перевіркої матриці цього коду з $d_0 = 3$ є те, що її стовпці – це будь-які різні ненульові комбінації завдовжки r . Наприклад, для коду Хеммінга (7, 4) можлива така перевірна матриця:

$$H = \begin{vmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{vmatrix}.$$

Стовпці перевіркої матриці можна довільно переставляти, при цьому коректувальні властивості коду не зміняться.

Коди Хеммінга з мінімальною кодовою відстанню $d_0 = 4$ утворюються на основі кодів з $d_0 = 3$ введенням ще одного перевірного розряду, що доповнює кодову комбінацію до парного числа одиниць. Перевірні матриця даного коду також утвориться з перевіркої матриці коду Хеммінга з $d_0 = 3$ приписуванням праворуч нульового стовпця і приписуванням додаткового рядка, що складається тільки з одиниць. Перевірні матриця, наприклад коду (8, 4), має вигляд:

$$H = \begin{vmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{vmatrix}.$$

При декодуванні кодової комбінації, переданої кодом Хеммінга з $d_0 = 4$, можливі три випадки:

- помилок немає (всі перевірки означають відсутність помилок);
- одинична помилка (всі перевірки означають наявність помилки);
- подвійна помилка (остання перевірка означає відсутність помилок, а інші – їх наявність).

Як видно, подвійна помилка, яка не виправляється кодом, має місце тільки тоді, коли результат хоча б однієї з перших чотирьох перевірок ненульовий, а результат останньої перевірки дорівнює нулеві. Це дає змогу запобігти видаванню одержувачу спотвореного повідомлення, тобто знайти подвійні помилки.

Іноді під кодами Хеммінга розуміють особливий різновид розглянутих кодів з $d_0 = 3$, який відрізняється тим, що комбінація синдрому збігається з номером спотвореного розряду, записаним у двійковій формі. Так, при спотворенні,

наприклад, третього розряду синдром має вигляд 011, четвертого – 100 і т. д. Ці коди характеризуються особливим порядком розташування перевірних розрядів – на першому, другому, четвертому, восьмому, шістнадцятому і т. д. місцях. Великого поширення ці коди не одержали, тому докладно їх не розглядатимемо.

3.3 Циклічні коди

Циклічні коди є різновидом систематичних кодів і мають усі їх властивості. Спочатку вони були створені з метою спрощення схем кодування і декодування. Згодом виявилися їх високі коректувальні властивості, що й забезпечило їм широке використання на практиці.

При побудові циклічних кодів кодові комбінації подають у вигляді поліномів:

$$G(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0, \quad (3.15)$$

де a_0, a_1, \dots, a_{n-1} – коефіцієнти, що набувають значень 0 або 1. Наприклад, комбінацію 1100101 можна записати як $G(x) = x^6 + x^5 + x^2 + 1$.

Основна властивість розглянутих кодів полягає в тому, що циклічний зсув дозволеної кодової комбінації також є дозволеною кодовою комбінацією. Отже, якщо комбінація 1000111 є дозволеною, то комбінації 0001111, 0011110 і т. д. теж дозволені. Нульова комбінація є дозволеною, оскільки циклічний код належить до класу систематичних. Зазначимо, що циклічний зсув нульової комбінації є також нульовою комбінацією.

Можна показати, що циклічний зсув є еквівалентним множенню на x кодової комбінації, записаної у вигляді полінома. Дійсно,

$$xG(x) = a_{n-1}x^n + a_{n-2}x^{n-1} + \dots + a_1x^2 + a_0x.$$

Оскільки в кодовій комбінації, що має довжину n , степінь полінома не може перевищувати $n-1$ (у протилежному разі довжина кодової комбінації перевищить n), то x^n замінюють на 1. При цьому

$$xG(x) = a_{n-2}x^{n-1} + \dots + a_1x^2 + a_0x + a_{n-1}.$$

Отже, $xG(x)$ є циклічним зсувом комбінації $G(x)$.

Циклічні коди прийнято визначати за допомогою твірних поліномів $P(x)$ степеня n . Твірну матрицю циклічного коду можна утворити з твірного полінома циклічним зсувом останнього (або, що те саме, множенням його на x, x^2, \dots, x^{n-1}):

$$G = \begin{vmatrix} P(x) \\ xP(x) \\ x^2P(x) \\ \vdots \\ x^{k-1}P(x) \end{vmatrix}. \quad (3.16)$$

Безпосередньо з твірної матриці (2.16) випливає, що всі дозволені кодові комбінації циклічного коду без остачі діляться на похідний поліном. Нагадаємо, що тут і далі мається на увазі ділення за модулем 2. Останнє виконується майже так само, як звичайне ділення, з тією різницею, що в процесі ділення операція

віднімання замінюється додаванням за модулем 2. Поліном $x^6 + x^5 + x^3 + 1$ на поліном $x^2 + x + 1$ (відповідні двійкові форми 1101001 і 111) ділиться так:

$$\begin{array}{r} \oplus \begin{array}{r} 1101001 \\ \underline{111} \end{array} \left| \begin{array}{r} 111 \\ \hline 10101 \end{array} \right. \\ \oplus \begin{array}{r} 110 \\ \underline{111} \end{array} \\ \oplus \begin{array}{r} 101 \\ \underline{111} \end{array} \\ 10 \end{array}$$

Двочлен 10 є остачею від ділення полінома 1101001 на поліном 111.

Оскільки кожна дозволена комбінація без остачі ділиться на твірний поліном, то ця обставина стає основою дуже зручного способу визначення приналежності прийнятої кодової комбінації до дозволених.

Розглянемо принцип побудови циклічних кодів. Кожну кодову комбінацію $G(x)$ простого k -елементного коду помножимо на x^r , а потім поділимо на твірний поліном степеня r . Внаслідок множення степінь кожного члена x_i , що входить у поліном $G(x)$, підвищується на r . При діленні добутку $x^r G(x)$ на $P(x)$ виходить частка $Q(x)$ такого самого степеня, що $G(x)$. Крім того, якщо добуток $x^r G(x)$ не ділиться на $P(x)$ без остачі, то остачею є $R(x)$:

$$\frac{x^r G(x)}{P(x)} = Q(x) \oplus \frac{R(x)}{P(x)}. \quad (3.17)$$

Оскільки частка $Q(x)$ має той самий степінь, що $G(x)$, то вона також є комбінацією простого k -елементного коду.

Помноживши обидві частини рівності (3.17) на $P(x)$, маємо

$$F(x) = Q(x)P(x) = x^r G(x) \oplus R(x). \quad (3.18)$$

Отже, кодову комбінацію циклічного коду можна отримати способами:

– множенням k -елементної комбінації простого коду на твірний поліном $P(x)$;

– множенням кодової комбінації простого коду на одночлен x^r і додаванням до цього добутку остачі від ділення добутку $x^r G(x)$ на $P(x)$.

Зазначимо, що перший спосіб спричинює утворення *нероздільного* коду. Нероздільність значно ускладнює процес декодування, тому на практиці використовується *інший спосіб* побудови кодових комбінацій. Дуже важливо, що цей спосіб дає можливість одержати твірну матрицю відразу в канонічній формі:

$$G = | E_k; C_{r;k} |,$$

де $C_{r;k}$ – матриця, що складається з r стовпців і k рядків, причому кожен рядок є остачею від ділення рядка одиничної матриці, доповненої r -нулями, на твірний поліном.

Кодують і декодують циклічні коди не на основі обчислення перевірок на парність, а на основі ділення на твірний поліном. Проте за необхідності перевірку матрицю можна побудувати обчисленням перевірного полінома:

$$h(x) = \frac{x^n + 1}{P^{-1}(x)}, \quad (3.19)$$

де $P^{-1}(x)$ – поліном, об'єднаний з твірним поліномом $P(x)$.

Нагадаємо, що в об'єднаних поліномах члени розташовані в зворотному порядку. Так, поліноми 100111 і 111001 є об'єднаними. Перший рядок перевірної матриці циклічного коду є перевірним поліномом $h(x)$, помноженим на x^{r-1} (тобто доповнений праворуч $r-1$ нулями). Наступні рядки перевірної матриці є циклічним зсувом першої.

Перевірна матриця може бути також побудована звичайним способом, виходячи з твірної матриці. У такому разі перевірна матриця зовні може відрізнятися від побудованої за допомогою перевірного полінома, однак обидві матриці завжди можуть бути зведені до одного вигляду.

Наведемо приклад побудови семирозрядного циклічного коду з $d_0=3$. Для цього потрібні три перевірних розряди ($d_0=3$), а твірний поліном має бути третього степеня. Нехай твірний поліном $P(x)=x^3+x+1=1011$. (Принципи вибору твірного полінома розглянемо пізніше.) Твірна матриця має вигляд

$$G = \begin{vmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{vmatrix}.$$

Нагадаємо, що перевірні розряди 101 маємо внаслідок ділення комбінації 1000000 на твірний поліном 1011, а 111 – внаслідок ділення 100000 на той самий твірний поліном, і т. д.

Перевірний поліном:

$$h(x) = \frac{x^7 + 1}{(x^3 + x + 1)^{-1}} = \frac{10000001}{1101} = 11101,$$

отже, перевірна матриця :

$$H = \begin{vmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{vmatrix}.$$

Якщо будувати твірну матрицю з породжувальної, то одержимо матрицю:

$$H' = \begin{vmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix},$$

яка відрізняється останнім рядком від раніше отриманої матриці H . Однак цю відмінність можна ліквідувати, якщо до останнього рядка матриці H' додати перший рядок; при цьому матриця H' збігається з матрицею H .

Коректувальна здатність циклічного коду цілком визначається виглядом твірного полінома. Для деяких циклічних кодів можна сформулювати досить прості принципи вибору вигляду твірного полінома.

Циклічний код з $d_0=2$. Твірний поліном має вигляд $x+1$. Цей поліном дає змогу будувати код будь-якої довжини. Циклічний код з $d_0=2$ виявляє будь-яку непарну кількість помилок і повністю тотожний коду з парним числом одиниць.

Твірним поліномом для циклічного коду з $d_0=2$ може бути також поліном x^2+x+1 . Код при цьому має підвищену завадостійкість – виявляються не тільки будь-які помилки непарної кратності, але й будь-які парні суміжні помилки (тобто

пакети помилок завдовжки 2), а також усі парні помилки, поділені одним неспотвореним розрядом.

Степінь	Вигляд	Циклічний код з $d_0 = 3$ є різновидом кодів Хеммінга.
полінома	полінома	Довжину кодової комбінації вибирають з умови $n = 2^r - 1$. Твірним поліномом може бути будь-який незвідний поліном степеня r . (Поліном називається незвідним, якщо він ділиться без остачі тільки на одиницю і на самого себе.) Незвідні поліноми до п'ятого степеня включно наведені поруч на цій самій сторінці.
1	$1+x$	Циклічний код з $d_0 = 4$ є також різновидом кодів Хеммінга і будується на основі твірних поліномів для кодів з $d_0 = 3$. Твірний поліном циклічного коду з $d_0 = 4$ є добутком двочлена $x+1$ на незвідний поліном, що придатний як твірний для коду з $d_0 = 3$. Довжину кодової комбінації вибирають з умови $n = 2^m - 1$; число перевірних розрядів $r = m + 1$. Наприклад, при $n = 7$ твірний поліном має вигляд $(x+1)(x^3+x+1) = x^4+x^3+x^2+1$.
2	x^2+x+1	
3	x^3+x+1	
4	x^3+x^2+1	
5	x^4+x+1	
	x^4+x^3+1	
	$x^4+x^3+x^2+x+1$	
	x^5+x^2+1	
	x^5+x^3+1	
	$x^5+x^3+x^2+x+1$	
	$x^5+x^4+x^2+x+1$	
	$x^5+x^4+x^3+x+1$	
	$x^5+x^4+x^3+x^2+1$	

3.4 Коди Боуза–Чоудхурі–Хоквінгема

Дуже потужними кодами, що мають високу коректувальну здатність, є коди Боуза–Чоудхурі–Хоквінгема. Довжина кодової комбінації становить $n = 2^m - 1$. Твірний поліном знаходять як найменше спільне кратне мінімальних поліномів $a_i(x)$

$$P(x) = \text{НСК}\{a_1(x)a_3(x)\dots a_i(x)\dots a_{d_0-2}(x)\}. \quad (3.20)$$

Мінімальні поліноми для $m \leq 7$ наведені в табл. 3.1, причому всі вони є незвідними.

Таблиця 3.1 – Мінімальні поліноми

i	Значення m					
	2	3	4	5	6	7
1	x^2+x+1	x^3+x+1	x^4+x+1	x^5+x^2+1	x^6+x+1	x^7+x^3+1
3	—	—	$x^4+x^3+x^2+x+1$	$x^5+x^4+x^3+x^2+1$	$x^6+x^4+x^2+x+1$	$x^7+x^3+x^2+x+1$
5	—	—	—	$x^5+x^4+x^2+x+1$	$x^6+x^5+x^2+x+1$	$x^5+x^4+x^3+x^2+1$
7	—	—	—	—	x^6+x^3+1	$x^7+x^6+x^5+x^4+x^2+x+1$

Нехай потрібно побудувати код Боуза–Чоудхурі–Хоквінгема з $n = 31$ і $d_0 = 5$.
Маємо:

$$P(x) = \text{НСК}\{a_1(x)a_3(x)\} = a_1(x)a_3(x).$$

Оскільки в даному випадку $m = 5$, то значення мінімальних поліномів випишемо з 5-го стовпця таблиці 3.1:

$$a_1(x) = x^5 + x^2 + 1,$$

$$a_3(x) = x^5 + x^4 + x^3 + x^2 + 1,$$

$$P(x) = x^{10} + x^9 + x^8 + x^6 + x^5 + x^3 + 1.$$

Коди Боуза–Чоудхурі–Хоквінгема мають непарні значення мінімальної кодової відстані d_0 . При бажанні мінімальну кодову відстань можна збільшити на одиницю, застосувавши твірний поліном, що дорівнює добутку твірного полінома коду Боуза–Чоудхурі–Хоквінгема на двочлен $x+1$. Так, у розглянутому кодi з $d_0 = 5$ мінімальну кодову відстань можна підвищити до 6, якщо використовувати твірний поліном:

$$P(x) = (x+1)(x^{10} + x^9 + x^8 + x^6 + x^5 + x^3 + 1) = x^{11} + x^8 + x^7 + x^5 + x^4 + x^3 + x + 1.$$

Зазначимо, що такий спосіб збільшення мінімальної кодової відстані можна застосувати до будь-яких систематичних кодів з непарною мінімальною кодовою відстанню. Для цього в циклічних кодах змінюється описаним чином твірний поліном, а в нециклічних кодах уводиться додаткова перевірка на парність, що охоплює всі інформаційні розряди аналогічно тому, як це було зроблено в циклічних і нециклічних кодах Хеммінга з $d_0 = 4$.

Циклічні коди поширені в системах передавання даних і використовуються у різній за призначенням апаратурі передавання даних (АПД). Незважаючи на те, що умови застосування кожної конкретної АПД різні, а використовувані канали можуть сильно відрізнятися за своїми характеристиками, є тенденція до стандартизації методів підвищення надійності приймання за допомогою коректувальних кодів і, зокрема, стандартизації твірних поліномів. Це дає можливість у ряді випадків спільно використовувати апаратуру різного виробництва, комплектувати АПД різного призначення тими самими функціональними вузлами і т. д. Зараз для середньошвидкісних систем передавання даних існує рекомендація МСЕ V.41, відповідно до якої для підвищення надійності приймання пропонується використовувати виявлення помилок за допомогою циклічного коду, що має довжину кодової комбінації 260, 500 і 980 розрядів, причому у всіх випадках береться твірний поліном $x^{16} + x^{12} + x^5 + 1$. Це код Боуза–Чоудхурі–Хоквінгема з мінімальною кодовою відстанню $d_0 = 4$. Численні випробування коду з зазначеним твірним поліномом підтвердили його високу ефективність. Навіть при використанні цього коду для передавання даних по комутованій телефонній мережі загального користування (коефіцієнт помилок по одиничних елементах порядку 10^{-3} і більше) частота помилкового приймання восьмирозрядних знаків, з яких складена інформаційна частина кодової комбінації, не перевищує 10^{-6} .

Тепер розглянемо питання про вибір довжини інформаційної частини кодової комбінації циклічного коду. Електронні обчислювальні машини, що є джерелом або споживачем інформації, переданої АПД, обмінюються з зовнішніми пристроями або безпосередньо машинними словами (16, 24, 32 і т. д. розрядів) або складами (байтами), що містять 8 розрядів. Тому інформаційна частина кодової комбінації має дорівнювати або бути кратною довжині машинного слова, якщо ЕОМ обмінюється з зовнішніми пристроями машинними словами, або дорівнювати довжині байта, якщо ЕОМ обмінюється з зовнішніми пристроями байтами. Проте можливості циклічних кодів не забезпечують вільного варіювання числом інформаційних розрядів при заданій мінімальній кодовій відстані або заданому

числі перевірних розрядів. У зв'язку з цим на практиці часто скорочують циклічний код.

Укорочені циклічні коди отримують з повних циклічних кодів, що мають необхідну мінімальну кодову відстань (або число перевірних розрядів), але більше, ніж потрібно, число інформаційних розрядів. Укорочення повного циклічного коду полягає в тому, що для передавання інформації використовуються не всі комбінації повного коду, а тільки ті, які містять ліворуч нулі, причому ці нулі в канал зв'язку зовсім не передаються. Практично, щоб укоротити код на 1 розряд, треба викреслити з відповідної матриці один верхній рядок і один лівий стовпець.

Наведемо приклад побудови укороченого циклічного коду. Нехай необхідно побудувати код, що має 5 інформаційних розрядів і мінімальну кодову відстань, яка дорівнює 3. Як відомо, при $d_0 = 3$ довжина кодової комбінації становить $n = 2^r - 1$, що дає змогу будувати коди зі значеннями n , k і r , наведеними нижче:

n	3	7	15	31	63	...
k	1	4	11	26	57	...
r	2	3	4	5	6	...

Отже, не існує циклічних кодів з $d_0 = 3$, що містять 5 інформаційних розрядів. Однак код з необхідними параметрами можна утворити з коду (15, 11), укоротивши його на 6 розрядів. Твірним поліномом цього коду може бути будь-який поліном четвертого степеня. Твірна матриця циклічного (15, 11)-коду має вигляд

$$G(15,11) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Твірну матрицю вкороченого циклічного (9, 5)-коду отримують з твірної матриці (15, 11)-коду викреслюванням 6 верхніх рядків і 6 лівих стовпців:

$$G(9,5) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Коректувальна здатність укороченого циклічного коду не нижча за коректувальну здатність початкового повного циклічного коду. Техніка кодування і декодування в обох випадках та сама. Однак циклічний зсув кодової комбінації укороченого циклічного коду не завжди спричинює утворення дозволеної комбінації, тому вкорочені коди належать до псевдоциклічних.

3.5 Несистематичні коди

Несистематичні коди в техніці передавання даних використовуються рідше, ніж систематичні. Одним з найбільш відомих кодів цього типу є код з постійною вагою, що називають також кодом з постійним відношенням числа одиниць і нулів. У коді з постійною вагою дозволеними є тільки ті комбінації, що містять визначене число одиниць, однакове у всіх дозволених комбінаціях. Серед кодів з постійною вагою найбільш поширений на практиці семирозрядний код з постійною вагою 3, кожна комбінація якого містить три одиниці і чотири нулі. Цей код рекомендований МСЕ для використання при передаванні телеграфних повідомлень по короткохвильових радіоканалах і відомий як *міжнародний телеграфний код № 3*.

Міжнародний телеграфний код № 3 містить 35 дозволених кодових комбінацій, мінімальна кодова відстань $d_0 = 2$. Код виявляє всі одиничні помилки, а також усі помилки непарної кратності. З різних помилок парної кратності не виявляються тільки помилки типу зсувів (транспозицій), при яких від однієї або кількох одиниць у кодовій комбінації переходять у нулі, а така сама кількість нулів – в одиниці. У цілком асиметричних каналах (тобто каналах, у яких переходять або тільки одиниці у нулі, або тільки нулі в одиниці) транспозиції не можуть мати місця, тому коди з постійною вагою в таких каналах виявляють усі помилки. Саме з цієї причини коди з постійною вагою поширені в апаратурі, що працює на короткохвильових радіоканалах, які є асиметричними.

Істотним недоліком кодів з постійною вагою є те, що вони належать до числа нероздільних кодів, у комбінаціях яких неможливо виділити інформаційні і перевірні розряди. Це значно ускладнює кодування і декодування, внаслідок чого ускладнюється і дорожчає апаратура. Тому останнім часом дані коди поступово витісняються систематичними і, зокрема, циклічними. Цьому сприяє наявна тенденція до створення досить універсальних пристроїв захисту від помилок, придатних для роботи на різних каналах, як провідних, так і короткохвильових.

Коди Бергера (коди з підсумовуванням) також призначені для використання в асиметричних каналах. Мінімальна кодова відстань у цих кодах $d_0 = 2$. Існує ряд варіантів побудови кодів Бергера. У найпростішому варіанті кодування відбувається так: в інформаційній частині кодової комбінації підраховується число одиниць, після чого формуються перевірні розряди, що являють собою запис цього числа в двійковій формі. Так само формуються перевірні розряди під час приймання і порівнюються з прийнятими перевірними розрядами. Підвищення надійності за допомогою кодів Бергера дає приблизно такі самі результати, як і використання коду № 3, однак найважливішою перевагою коду Бергера є його *роздільність*, що різко спрощує побудову кодувальних і декодувальних пристроїв.

У техніці передавання даних поширені коди з інвертованими перевірками на парність. Це несистематичні коди, одержувані із систематичних інвертуванням одного або декількох перевірних розрядів. Так, широко застосовується код з непарним числом одиниць, що відрізняється від коду з парним числом одиниць інвертованим перевірним розрядом. Коректувальна здатність кодів з інвертованими перевірками цілком збігається з коректувальною здатністю систематичних кодів. Значною перевагою розглянутих кодів є те, що вони не містять кодових комбінацій, які складаються з одних нулів. Ця обставина дуже

сприятлива для побудови систем передавання даних, оскільки відсутність у коді нульових комбінацій підвищує стійкість синхронізації (у синхронних системах) і тим самим поліпшує роботу систем передавання даних. Крім того, у ряді випадків застосування кодів з інвертованими перевірками, отриманих з циклічних кодів, дає змогу набагато ефективніше виявляти порушення циклової синхронізації, ніж при використанні звичайних циклічних кодів. Така властивість кодів з інвертованими перевірками значною мірою підвищує стійкість роботи системи передавання даних.

3.6 Ітеративні коди

Ітеративні (матричні) коди характеризуються наявністю двох перевірок усередині кожної кодової комбінації. Розглянемо принцип побудови ітеративного коду на конкретному прикладі. Запишемо всі інформаційні розряди блоку, що підлягає передаванню, у вигляді таблиці, що, наприклад, може мати такий вигляд:

1	0	1	1	1
0	0	1	0	0
1	1	1	0	0
0	1	0	0	1
1	1	1	1	0

Закодуємо спочатку кожен рядок таблиці яким-небудь кодом, а потім (не обов'язково тим самим кодом) – кожен стовпець. За перший код візьмемо код з парним числом одиниць, а за другий – з непарним.

Тоді одержимо:

1	0	1	1	1	0
0	0	1	0	0	1
1	1	1	0	0	1
0	1	0	0	1	0
1	1	1	1	0	0
0	0	1	1	1	1

Отримана комбінація є кодовою комбінацією найпростішого двовимірного ітеративного коду, перевірни розряди якого зосереджені в правому стовпці і нижньому рядку. Кожен інформаційний розряд цього коду входить у комбінацію двох ітеративних кодів – коду з парним числом одиниць і коду з непарним числом одиниць. Передавання комбінації ітеративного коду звичайно відбувається по рядках послідовно – від першого рядка до останнього.

Можуть бути утворені також багатовимірні ітеративні коди, в яких кожен інформаційний розряд входить у комбінації трьох, чотирьох і т. д. ітеративних кодів, однак багатовимірні ітеративні коди менш поширені.

Властивості ітеративного коду визначаються його параметрами, залежно від яких код може бути як систематичним, так і несистематичним, як роздільним, так і нероздільним. Довжина кодової комбінації, число інформаційних розрядів і мінімальна кодова відстань ітеративного коду дуже просто виражаються через відповідні параметри цих кодів:

$$n = \prod_{i=1}^S n_i, \quad k = \prod_{i=1}^S k_i, \quad d_0 = \prod_{i=1}^S d_{0i}, \quad (3.21)$$

де n_i, k_i, d_{0i} – параметри ітеративних кодів; S – кратність ітерування.

Отже, найпростіший ітеративний код, утворений перевіркою на парність (непарність) рядків і стовпців, має мінімальну кодову відстань $d_{0i} = 4$ і тому дає можливість виявляти всі помилки кратністю до 3. Крім того, виявляються всі помилки непарної кратності. Не виявляються чотириразові помилки, що розташовуються у вершинах правильного чотирикутника, а також деякі шестиразові, восьмиразові і т. д. помилки (рис. 3.2).

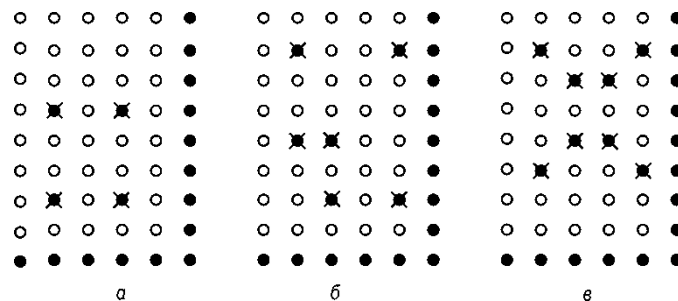


Рисунок 3.2 – Деякі помилки, що не виявляються найпростішим ітеративним кодом: ● – чотириразові помилки; ✕ – шести-, восьмиразові і т. д. помилки

Найпростіший ітеративний код має досить високу виявляючу здатність – при дії пакетних помилок виявляється будь-який пакет помилок завдовжки $l+1$ і менше, де l – довжина рядка.

На практиці найпоширеніші *двовимірні ітеративні коди*. Довжину рядка звичайно вибирають такою, що дорівнює довжині одного знака первинного коду. Як ітеративні коди найчастіше використовуються коди з однією і двома перевітками на парність і набагато рідше коди Хеммінга. При використанні досить довгих блоків (завдовжки в кілька десятків знаків і більше) проста перевітка на парність по рядках і стовпцях забезпечує на реальних каналах зв'язку коефіцієнт помилок на знак порядку 10^{-6} , а застосування кодів із двома перевітками на парність і кодів Хеммінга – порядку $10^{-8} - 10^{-10}$.

Істотним недоліком ітеративних кодів, які використовують для перевірок по рядках і стовпцях кодів з однією або двома перевітками на парність, є їх порівняно високий надлишок, що звичайно становить 15–20 % і значно перевищує за інших рівних умов надлишок циклічних кодів. Однак кодування і декодування за допомогою ЕОМ таких ітеративних кодів звичайно набагато простіше, ніж циклічних. Тому найпростіші ітеративні коди, незважаючи на їх високий надлишок, застосовують у системах передавання даних, що використовують програмні способи підвищення надійності (зокрема, у системах з комутацією повідомлень).

3.7 Ефективність систематичних коректувальних кодів

Розглянемо ефективність використання систематичних коректувальних кодів, що виявляють помилки. Найважливішою характеристикою таких кодів є коефіцієнт підвищення правильності:

$$K_{п.п} = \frac{P_{пом}}{P_{н.пом}}, \quad (3.22)$$

де $P_{\text{пом}}$ – частота появи помилкових кодових комбінацій у дискретному каналі (тобто перед пристроєм захисту від помилок); $P_{\text{н.пом}}$ – частота появи кодових комбінацій з невиявленими помилками в каналі передавання даних (тобто після пристрою захисту від помилок).

Характерною рисою використання коректувальних кодів є те, що їх ефективність залежить від особливостей розподілу помилок у каналах. Код, що виявляє якісь певні сполучення помилок, буде досить ефективним при його використанні в каналах, де ці сполучення є переважаючими. Так, код з однією перевіркою на парність виявляє всі помилки непарної кратності і може при нежорстких вимогах до надійності приймання бути цілком придатним для роботи з каналами з частотною модуляцією (ЧМ). Проте цей код практично непридатний під час роботи з каналами з відносною фазовою модуляцією (ВФМ), характерною рисою яких є помилки парної кратності. Причина неоднакового поведіння коду в різних каналах – в різних закономірностях розподілу помилок у них.

Відомо, що коректувальний код з мінімальною кодовою відстанню d_0 виявляє всі сполучення помилок кратністю до $d_0 - 1$ включно. Тому ймовірність невиявлення помилки кодом можна визначити як:

$$P_{\text{н.пом}} = \frac{B(d_0)}{C_n^{d_0}} P(d_0, n) + \frac{B(d_0 + 1)}{C_n^{d_0 + 1}} P(d_0 + 1, n) + \dots = \sum_{i=d_0}^n \frac{B(i)}{C_n^i} P(i, n), \quad (3.23)$$

де $B(i)$ – кількість i -кратних помилок, що не виявляються кодом; C_n^i – загальна кількість можливих i -кратних помилок у кодовій комбінації завдовжки n (дорівнює числу сполучень з n по i); $P(i, n)$ – імовірність появи i -кратних помилок у кодовій комбінації завдовжки n . Ця ймовірність цілком залежить від властивостей використовуваного каналу.

Визначення відношення B_i / C_n^i у багатьох випадках є дуже трудомісткою задачею, особливо при великих значеннях n . Доведено, що для багатьох систематичних кодів це відношення порівняно мало залежить від конкретного різновиду коду. Існує наближене співвідношення:

$$\frac{B(i)}{C_n^i} \approx 2^{-r}, \quad (3.24)$$

яке визначає відношення кількості комбінацій помилок, що не виявляються, кратності k до загальної кількості можливих комбінацій помилок цієї кратності для коду в середньому. Зазначимо, що дане співвідношення дуже близьке до відношення кількості комбінацій помилок, які не виявляються кодом, до загальної кількості можливих помилок без урахування їх кратності. Дійсно, загальне число можливих комбінацій помилок різної кратності від 1 до n включно дорівнює $2^n - 1$, причому з них не виявляються тільки ті комбінації, що збігаються з ненульовими кодовими комбінаціями, число яких дорівнює $2^n - 1$. Їх відношення:

$$\frac{2^k - 1}{2^n - 1} \approx \frac{2^k}{2^n} = 2^{-r}.$$

Отже,

$$P_{\text{н.пом}} \approx 2^{-r} \sum_{i=d_0}^n P(i, n) = 2^{-r} P(\geq d_0, n), \quad (3.25)$$

де $P(\geq d_0, n)$ – імовірність того, що в кодовій комбінації завдовжки n буде помилок більше, ніж d_0 , або їх кількість дорівнюватиме d_0 . Нагадаємо, що ймовірність появи в кодовій комбінації $P_{\text{пом}}$ помилок записується як $P_{\text{пом}} = P(\geq 1, n)$.

Тоді коефіцієнт підвищення правильності коду можна визначати як:

$$K_{\text{п.п}} = \frac{P_{\text{пом}}}{P_{\text{н.пом}}} = \frac{P(\geq 1, n)}{2^{-r} P(\geq d_0, n)} = 2^r \frac{P(\geq 1, n)}{P(\geq d_0, n)}. \quad (3.26)$$

Відношення $P(\geq 1, n)/P(\geq d_0, n)$ завжди більше за 1. Тому коефіцієнт підвищення правильності коду $K_{\text{п.п}}$ майже завжди перевищує 2^r . Для більшості систематичних кодів величина 2^r є досить надійною нижньою оцінкою коефіцієнта підвищення правильності $K_{\text{п.п}}$.

Співвідношення (3.26) визначає коефіцієнт підвищення надійності коду за допомогою ймовірності $P(\geq d_0, n)$, тобто ймовірності того, що в кодовій комбінації буде помилок більше, ніж d_0 , або їх кількість дорівнюватиме d_0 . Іноді зручніше користуватися не ймовірностями появи помилок визначеної кратності, а ймовірностями появи пакетів помилок визначеної довжини. При цьому нагадаємо, що пакетом помилок називається комбінація помилок, котра починається і закінчується помилковими розрядами, між якими можуть розташовуватися як помилкові, так і безпомилкові розряди. Циклічний код, утворений поліномом степеня r , виявляє будь-який пакет помилок завдовжки r і менше. Правдивість цього положення легко зрозуміти, якщо врахувати, що пакету помилок завдовжки r і менше відповідає поліном помилок степеня, меншого за r , оскільки він не ділиться без остачі на твірний поліном $P(x)$. Отже, кодова комбінація, що є сумою за модулем 2, передані кодові комбінації і комбінації завади не ділитимуться без остачі на твірний поліном. Унаслідок цього будь-який пакет помилок завдовжки r і менше завжди виявляється.

Наведемо без доведення положення, що характеризують властивості циклічних кодів:

- кількість пакетів завдовжки $r+1$, що виявляють циклічним кодом, становить $1/2^{r-1}$ частини всіх пакетів завдовжки $r+1$;
- кількість пакетів завдовжки більше $r+1$, що виявляють циклічним кодом, становить $1/2^r$ частини всіх пакетів помилок завдовжки від $r+2$ до n включно.

Отже, для циклічних кодів можна записати, що

$$P_{\text{н.пом}} = \frac{1}{2^{r-1}} P(r+1, n) + \frac{1}{2^r} P(\geq r+2, n) \approx \frac{1}{2^r} P(> r, n), \quad (3.27)$$

де $P(> r, n)$ – імовірність появи в кодовій комбінації пакету помилок завдовжки більше r розрядів. Імовірність появи в кодовій комбінації пакету помилок будь-якої довжини $P(\geq 1, n)$ дорівнює, і в цьому легко переконатися, ймовірності появи в кодовій комбінації помилки.

Тому коефіцієнт підвищення правильності при використанні циклічного коду:

$$K_{\text{п.п}} \approx 2^r \frac{P(\geq 1, n)}{P(> r, n)}. \quad (3.28)$$

Величина 2^r для більшості випадків є досить надійною нижньою оцінкою коефіцієнта підвищення правильності циклічних кодів.

Отже, при використанні систематичних кодів можна вважати, що ймовірність невиявлення кодом помилкового блоку в 2^r рази менша, ніж ймовірність появи помилкових блоків на вході декодувального пристрою. Властивості виявляючих систематичних кодів цілком визначаються кількістю перевірних розрядів. Досить істотне значення має вигляд перевірних співвідношень, а для циклічних кодів – вигляд твірного полінома. Серед систематичних кодів з однаковим числом перевірних розрядів існують так звані «сильні» коди, що забезпечують на даних каналах коефіцієнт підвищення правильності, який значно перевищує 2^r . Є також «слабкі» коди, коефіцієнт підвищення правильності яких менше ніж 2^r . Тому при створенні систем передавання даних завжди відбирають «сильні» коди, які щонайкраще відповідають розподілу помилок у дискретних каналах, і найбільші значення коефіцієнта, що забезпечують підвищення правильності. Саме таким «сильним» кодом є згаданий вище циклічний код з твірним поліномом $x^{16} + x^{12} + x^5 + 1$.

Зазначимо, що аналітичні методи вибору «сильних» кодів далеко не завжди дають можливість здійснити такий вибір. Складність задачі визначається специфічним характером розподілу помилок, описати який аналітично досить точно не завжди вдається. Тому на практиці поширені методи добору кодів, засновані на моделюванні реальних потоків помилок за допомогою ЕОМ.

У цьому розділі ми обмежимося розглядом ефективності кодів, що виявляють помилки, і не розглядатимемо ефективність кодів, що виправляють помилки, у зв'язку з порівняно рідким застосуванням виправлення помилок на практиці.

3.8 Мажоритарне декодування

В основі мажоритарного способу виправлення помилок – визначення, який розряд кодової комбінації, що декодує, за більшістю результатів перевірок на парність значно спрощує декодери. Мажоритарне декодування застосовують не для всіх кодів, а для тих, структура яких має певні особливості. До таких кодів належать, зокрема, деякі циклічні коди, а також більш прості коди, у тому числі коди з повторенням при $S > 2$. Розглянемо дуже простий випадок мажоритарного декодування стосовно використання кодів з розділеними перевірками.

Як відомо, кожен рядок перевіркової матриці будь-якого систематичного коду

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} & \dots \\ h_{21} & h_{22} & h_{23} & \dots \\ \dots & \dots & \dots & \dots \\ h_{r1} & h_{r2} & h_{r3} & \dots \end{pmatrix}$$

складається з коефіцієнтів, що входять у співвідношення перевірки на парність: $h_{i1}x_1 \oplus h_{i2}x_2 \oplus \dots \oplus h_{in}x_n = 0$. Лінійні комбінації рядків перевіркової матриці також

утворюють перевірни співвідношення. Виконавши μ лінійних операцій над рядками перевірної матриці, можна (для деяких кодів) одержати нову матрицю:

$$L = \begin{pmatrix} l_{11} & l_{12} & l_{13} & \dots & \dots \\ l_{21} & l_{22} & l_{23} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ l_{\mu 1} & l_{\mu 2} & l_{\mu 3} & \dots & \dots \end{pmatrix},$$

що характеризується двома важливими властивостями:

- один із стовпців містить тільки одиничні елементи;
- всі інші стовпці містять не більше ніж по одному одиничному елементу.

Матриця L визначає μ перевірок на парність для розряду, що відповідає одиничному стовпцю. Додавши до цієї сукупності перевірок тривіальну перевірку $x_i = x_i$, одержимо $\mu + 1$ незалежних перевірних співвідношень для одного розряду x_i , причому властивості матриці L такі, що кожен розряд кодової комбінації входить тільки в одну перевірку.

Така сукупність перевірок називається *системою поділених (ортогональних) перевірок* щодо розряду x_i .

Мажоритарне декодування здійснюється так. Якщо в прийнятій кодовій комбінації помилки відсутні, то при визначенні розряду x усі $\mu + 1$ перевірки вкажуть одне й те саме значення (або 1, або 0). Одинична помилка в кодовій комбінації може спричинити спотворення лише однієї перевірки, подвійна помилка – двох і т. д. Значення розряду x_i вибирають за більшістю (тобто мажоритарно) однойменних результатів перевірок.

При цьому декодування безпомилкове, якщо число помилок у кодовій комбінації не перевищує $\mu/2$, тобто спотворено не більше ніж $\mu/2$ перевірок. Якщо всі системи поділених перевірок для кожного розряду кодової комбінації містять не менше ніж $\mu + 1$ поділених перевірок, то реалізована мінімальна кодова відстань $d_0 = \mu + 1$.

Зазначимо, що далеко не всі коди допускають мажоритарне декодування, оскільки вимоги до структури перевірної матриці для такого декодування задовольняють не всі коди.

Пояснимо принцип мажоритарного декодування на конкретному прикладі. Нехай необхідно побудувати декодувальний пристрій для циклічного $(15, 7)$ -коду Боуза–Чоудхурі–Хоквінгема, що допускає мажоритарне декодування. Код має твірний поліном $P(x) = x^8 + x^7 + x^6 + x^4 + x + 1 \approx 111010001$.

Мінімальна кодова відстань $d_0 = 5$, і код може виправити дворазову помилку.

Обчислимо перевірний поліном:

$$h(x) = \frac{x^n + 1}{p^{-1}(x)} = x^7 + x^3 + x + 1 \approx 10001011.$$

Побудуємо перевірну матрицю.

При цьому як перший рядок використаємо перевірний поліном, помножений на x^{r-1} , а інші рядки одержимо циклічним зсувом першого:

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Перетворимо перевірну матрицю так. Додамо 1, 5, 7 і 8-й рядки матриці:

$$1 \oplus 5 \oplus 7 \oplus 8 = 100000001000101.$$

Аналогічно:

$$2 \oplus 3 \oplus 6 \oplus 7 \oplus 8 = 011000000010001,$$

$$4 \oplus 6 \oplus 7 \oplus 8 = 000101100000001.$$

Складемо матрицю L , використавши для її побудови три отримані суми і 8-й рядок перевірної матриці H :

$$L = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Як бачимо, в цій матриці один із стовпців складається тільки з одиниць, а всі інші стовпці містять не більше однієї одиниці. Матриця L дає чотири незалежних перевірних співвідношення з поділеними щодо члена a_0 перевітками:

$$a_0 = a_1 \oplus a_3 \oplus a_7;$$

$$a_0 = a_2 \oplus a_6 \oplus a_{14};$$

$$a_0 = a_4 \oplus a_{12} \oplus a_{13};$$

$$a_0 = a_8 \oplus a_9 \oplus a_{11}.$$

Додавши до цих співвідношень тривіальну перевірку $a_0 = a_0$, одержимо систему поділених відносно a_0 перевірок:

$$a_0 = a_0;$$

$$a_0 = a_2 \oplus a_6 \oplus a_{14};$$

$$a_0 = a_4 \oplus a_{12} \oplus a_{13};$$

$$a_0 = a_8 \oplus a_9 \oplus a_{11};$$

$$a_0 = a_1 \oplus a_3 \oplus a_7.$$

Нехай при передаванні спотворено розряд a_6 , який входить тільки в другу перевірку, тому чотири перевірки дадуть правильний результат, а друга перевірка – неправильний. Значення розряду не вибирають за критерієм більшості і тому воно буде правильним. Помилкова реєстрація розряду відбудеться при дії трьох і більше помилок, що спричиняють неправильні результати трьох і більше перевірок.

Ми одержали систему поділених перевірок щодо розряду a_0 . Системи поділених перевірок для інших розрядів отримують циклічним зсувом рядків матриці L . Зробивши, наприклад, зсув на один розряд, одержимо

$$L' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

звідки випливає:

$$a_{14} = a_0 \oplus a_2 \oplus a_6;$$

$$a_{14} = a_1 \oplus a_5 \oplus a_{13};$$

$$a_{14} = a_3 \oplus a_{11} \oplus a_{12};$$

$$a_{14} = a_7 \oplus a_8 \oplus a_{10}.$$

Крім того, $a_{14} = a_{14}$.

Аналогічно можна одержати системи поділених перевірок і для всіх інших розрядів кодової комбінації. Крім того, системи перевірок для циклічних кодів можна отримати циклічним зсувом якої-небудь однієї системи перевірок.

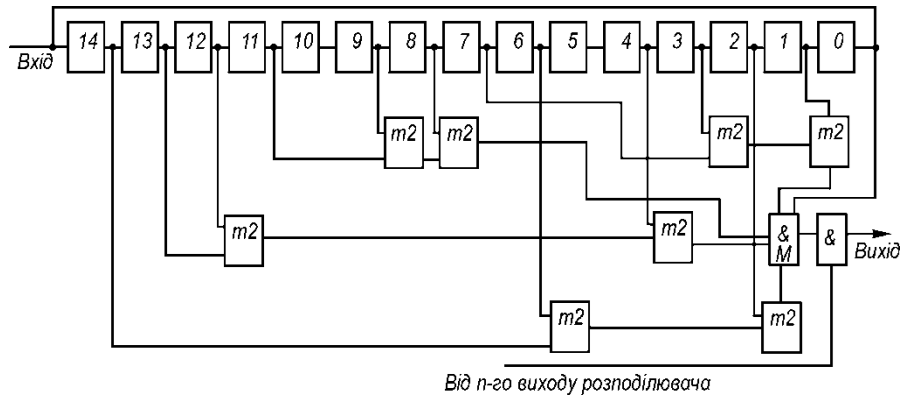


Рисунок 3.3 – Побудова мажоритарного декодера циклічного (15, 7)-коду з подільними перевірками [7]

Схема мажоритарного декодера циклічного (15, 7)-коду показана на рисунку 3.3. Декодер складається з n -розрядного регістра зсуву, набору двовхідних суматорів за модулем 2 і так званих мажоритарних органів M , завдяки яким вибирають значення декодованого розряду відповідно до критерію більшості. Підключають суматори до чарунок регістра зсуву згідно з системою поділених щодо розряду a_0 перевірок.

Діє декодер так. На виходах суматорів формуються результати поділених перевірок щодо розряду a_0 , і мажоритарний орган вибирає значення розряду a_0 . Далі в регістр подається ще один тактовий імпульс, і мажоритарний орган вибирає значення розряду a_1 , і т. д. до декодування розряду a_{14} . Отже, декодування кодової комбінації здійснюється за 2^n такти: протягом перших тактів заповнюється регістр зсуву, а протягом наступних визначається кожний розряд.

3.9 Узагальнення теорії кодування на недвійкові коди

Дотепер ми розглядали тільки двійкові лінійні коди. Однак це робилося лише для простоти. Насправді теорія лінійних кодів звичайно викладається відразу для m -кових кодів, де $m = p^l$ (p – просте число; l – натуральне число), тобто для випадку, коли символи коду утворюють кінцеве поле Галуа $GF(q)$ і над ними можуть бути здійснені всі арифметичні дії, що існують над дійсними чи комплексними числами. Для таких m -кових кодів можуть бути визначені всі поняття і доведені всі властивості, отримані раніше для двійкових кодів, а саме – породжувальна і перевірна матриці, систематичні та дуальні коди, границі мінімальних відстаней,

стандартні розташування, синдроми, вагові спектри і границі для ймовірностей невиявлених і виявлених помилок, циклічні і БЧХ-коди, алгебраїчні і мажоритарні алгоритми декодування. Найбільш важливим класом m -кових кодів є коди Ріда–Соломона (скорочено РС-коди). Вони можуть бути побудовані як систематичні циклічні (n, k) -коди при $n = q - 1, n - k = 2t$, де t – число помилок, що виправляються. Коди РС є частиною стандарту цифрового запису на компакт-диски.

За означенням вектор \mathbf{x} є словом m -кового (n, k) -коду РС, якщо відповідний йому многочлен $f_x(D)$ має корені, рівні елементам поля $GF(q): \alpha, \alpha^2, \dots, \alpha^{n-k}$, де α – примітивний елемент цього поля. Породжувальний многочлен коду РС має вигляд:

$$g(D) = (D - \alpha)(D - \alpha^2) \dots (D - \alpha^{n-k}). \quad (2.29)$$

Як бачимо з означення РС-коду, він є частковим випадком m -кових БЧХ-кодів, і, відповідно до доведеного раніше, мінімальна кодова відстань таких кодів буде в точності дорівнювати $d = n - k + 1$.

Легко показати, що жодний лінійний систематичний m -ковий ($m \geq 2$) код не може мати $d > n - k + 1$. Дійсно, якщо вибрати значення k мінус одного інформаційного символу рівними нулю, то це дасть ненульове кодове слово вагою не більше, ніж $n - k + 1$, що за властивістю лінійного коду і визначає верхню границю для d як $n - k + 1$. Оскільки РС-код реалізує верхню границю для мінімальної кодової відстані, то він є оптимальним серед усіх m -кових (n, k) -кодів щодо виправлення і виявлення помилок гарантованої кратності.

Можна дати простий опис РС-коду в несистематичному зображенні. Тоді кодовий вектор визначається як:

$$\mathbf{x} = [F(1), F(\alpha), F(\alpha^2), \dots, F(\alpha^{q-2})], \quad (3.30)$$

де $F(\alpha) = b_0 + b_1 D + \dots + b_{k-1} D^{k-1}$, b_0, b_1, \dots, b_{k-1} – значення інформаційних m -их символів.

Вибір довжини коду $n = d - 1$ є досить сильним обмеженням, особливо при великому порядку поля. Тому будують так звані *скорочені* коди РС, що мають довільну довжину $n \leq d - 1$. Їх можна одержати з повних РС-кодів, що мають довжину $n = d - 1$, якщо покласти частину інформаційних символів рівними нулю і викинути їх з кодових блоків. Оскільки укорочення коду не може зменшити мінімальної кодової відстані, то (n, k) -код при $n \leq d - 1$ буде, як і раніше, мати $d = n - k + 1$.

Коди РС, як окремий випадок БЧХ-кодів, мають алгебраїчний алгоритм виправлення помилок з поліноміальною складністю. Коди m -кові можуть бути використані разом із двійковими кодами для побудови каскадних кодів.

Ітеративні і каскадні коди. Потужні коди (тобто коди з довгими блоками і великою кодовою відстанню d) при порівняно простій процедурі декодування можна будувати, поєднуючи кілька коротких кодів. Так будується, наприклад, *ітеративний* код із двох лінійних систематичних кодів (n_1, k_1) і (n_2, k_2) , таблиця 3.2. Спочатку повідомлення кодується кодом 1-го ступеня (n_1, k_1) . Нехай k_2 блоків коду 1-го ступеня записані у вигляді рядків матриці. Її стовпці містять по

k_2 символів, які вважатимемо інформаційними для коду 2-го ступеня (n_2, k_2) , і допишемо до них $n_2 - k_2$ перевірних символів.

У результаті вийде блок (матриця $n_1 \times n_2$), що містить $n_1 n_2$ символів, з яких $k_1 k_2$ є інформаційними. Процес побудови коду можна продовжити в 3-му вимірі тощо.

Таблиця 3.2 – Побудова ітеративного коду [7]

	Інформаційні символи 1-го ступеня	Перевірні символи 1-го ступеня
Інформаційні символи 2-го ступеня	$b_{1,1}, b_{1,2}, \dots, b_{1,k_1}$ $b_{2,1}, b_{2,2}, \dots, b_{2,k_2}$ $b_{k_2,1}, b_{k_2,2}, \dots, b_{k_2,k_2}$	$b_{1,k_1+1}, b_{1,k_1+2}, \dots, b_{1,n_1}$ $b_{2,k_1+1}, b_{2,k_1+2}, \dots, b_{2,n_1}$ $b_{k_2,k_1+1}, b_{k_2,k_1+2}, \dots, b_{k_2,n_1}$
Перевірні символи 2-го ступеня	$b_{k_2+1,1}, b_{k_2+1,2}, \dots, b_{k_2+1,k_1}$ $b_{n_2,1}, b_{n_2,2}, \dots, b_{n_2,k_1}$	$b_{k_2+1}, b_{k_2+1,k_1+1}, \dots, b_{k_2+1,n_1}$ $b_{n_2,k_1+1}, b_{n_2,k_1+2}, \dots, b_{n_2,n_1}$

При декодуванні кожного блоку 1-го ступеня виявляють і виправляють помилки. Після того як прийнятий весь двовимірний блок, знову виправляють помилки і витирання, але вже по стовпцях, кодом 2-го ступеня, причому доводиться виправляти тільки ті помилки, що не були виправлені (чи були «виправлені» неправильно) кодом 1-го ступеня. Легко перекозатися, що мінімальна кодова відстань для двовимірного ітеративного коду $d = d_1 d_2$, де d_1 і d_2 – відповідно мінімальні кодові відстані для кодів 1-го і 2-го ступенів.

Дуже ефективний різновид потужних кодів – *каскадні* коди. Двокаскадний код (рис. 3.4) будується так: спочатку k_1 двійкових символів джерела розглядаються як збільшений символ багатопозиційного коду з основою $m = 2^{k_1}$. Потім до послідовності з k_2 таких збільшених символів додається $n_2 - k_2$ перевірних символів m -кового коду (кожен перевірний символ – це послідовність із k_1 двійкових символів). На цьому завершується утворення зовнішнього коду. Після цього формується внутрішній код з кодовою відстанню d_1 : до кожних k_1 елементарних двійкових символів зовнішнього коду додається $n_1 - k_1$ перевірних двійкових символів.

Результуюча кодова комбінація містить $n_1 n_2$ двійкових символів, з яких $k_1 k_2$ є інформаційними. Цим каскадний код схожий на ітеративний. Однак декодування каскадного коду виконується в такий спосіб: спочатку послідовно здійснюється декодування всіх блоків внутрішнього коду (з виявленням або виправленням помилок), потім декодується блок зовнішнього m -кового коду (n_2, k_2) , причому виправляються помилки і стирання, що залишилися після декодування внутрішнього коду.

Внутрішній код звичайно розрахований на виправлення одиничних помилок, зовнішній – на виправлення груп помилок, які є одиничними помилками в збільшених m -кових символах (рис. 3.4). Як зовнішній код звичайно використовується m -ковий код Ріда–Соломона (див. вище), який забезпечує найбільше можливе d_2 при заданих значеннях n_2 і k_2 , якщо $n_2 < m$.

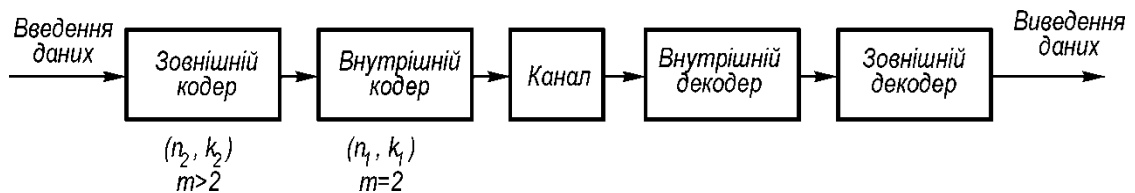


Рисунок 3.4 – Схема каскадного кодування та декодування [2]

Побудований каскадний код еквівалентний лінійному двійковому коду з мінімальною відстанню $d \geq d_1 d_2$. Фактично розглянутий вище алгоритм декодування каскадного коду є більш ефективним і простим, ніж алгоритм декодування двійкового коду з еквівалентною мінімальною відстанню d .

Використання каскадних кодів у каналах без пам'яті дозволяє забезпечити експонентне спадання ймовірності помилкового декодування при збільшенні повної довжини блока $n = n_1 n_2$, при цьому швидкість передавання $R = \frac{k_1 k_2}{n_1 n_2}$ може бути

як завгодно близькою до пропускну здатності каналу, а складність декодування поліноміально залежить від повної довжини блока n . Ще ефективнішим є застосування каскадних кодів у каналах з пам'яттю. Процес нарощування ступенів каскадного коду можна продовжити. Каскадні коди в багатьох випадках найперспективніші серед відомих блокових завадостійких кодів.

Кодування в каналах з пам'яттю. Хоча пам'ять і збільшує пропускну здатність каналів зв'язку, однак це не означає, що використання тих самих кодів у каналах з пам'яттю і без пам'яті дає меншу ймовірність помилкового декодування для каналів з пам'яттю. Більше того, навіть найкращі коди для каналів з пам'яттю можуть виявитися значно гіршими, ніж посередні коди в каналах без пам'яті.

Реальні фізичні канали зв'язку володіють, як правило, пам'яттю, що може пояснюватися кореляцією випадкових завад і параметрів каналів зв'язку. Типовими прикладами каналів з пам'яттю є канали з завмираннями. Універсальним методом зведення каналів з пам'яттю до каналів без пам'яті є переміжність символів на передачі та їх депереміжність на прийомі. Після переміжності символів можна використовувати алгоритми виправлення помилок для каналів без пам'яті. Платою за дану перевагу є затримки декодування, додаткова пам'ять, а також зменшення пропускну здатності каналу зв'язку. Остання обставина може бути «пом'якшена» переходом до алгоритму кодування-декодування з переміжністю і оцінкою ймовірності помилок. У цьому випадку пам'ять каналу використовується для оцінки ймовірностей помилок символів у кодових блоках, які потім враховуються при виправленні помилок.

Якщо пам'ять каналу проявляється в появі чітко виражених пачок помилок, тобто групи помилкових символів, що йдуть один за одним, то можна застосувати спеціальні коди, орієнтовані на виправлення помилок саме такої конфігурації, а не незалежних помилок. Ефективними в каналах з пам'яттю є й каскадні коди, якщо внутрішні коди в них використовуються в основному для виявлення пачок помилок і стирання блоків, на яких ці пачки виявлені, а зовнішній код використовується для виправлення стертих блоків.

Останнім часом обговорюються можливості використання в каналах з пам'яттю методу адаптивного кодування (декодування), коли за відсутності переміжності символів на передаванні декодер у режимі виправлення пачок помилок запускається лише тоді, коли зафіксовано групування помилок у каналі. В

інший час декодер працює в режимі виправлення одиничних помилок. Метод адаптивного кодування (декодування) істотно зменшує затримку прийняття рішення в порівнянні з методом переміжності символів на передачі і їх депереміжності на прийомі.

3.10 Системи зі зворотним зв'язком

Ми розглянули системи, в яких передавання інформації здійснювалося в одному напрямку: від передавача до приймача. Але існують системи, де між кінцевими пунктами можливий двосторонній обмін інформацією, що, природно, припускає розміщення в кожному з них передавача і приймача. Зворотний канал зв'язку в таких системах може використовуватися для передавання не тільки звичайної інформації, але й спеціальних повідомлень, призначених для підвищення завадостійкості прямого каналу. Системи зв'язку, в яких по зворотному каналу передаються сигнали, що автоматично коректують помилки в прямому каналі, називаються *системами зі зворотним зв'язком*.

Коректування помилок може здійснюватися двома способами. Перший з них характеризується тим, що помилки виявляються на приймальному кінці. У разі виявлення помилок по зворотному каналу передається сигнал запити для повторення спотвореного повідомлення. Якщо на передавальному кінці сигнал запити прийнятий правильно, то поточне передавання переривається й автоматично повторюється помилково прийняте повідомлення. Така система зветься *системою з керуючим зворотним зв'язком або системою з автозапитом* (перезапитом).

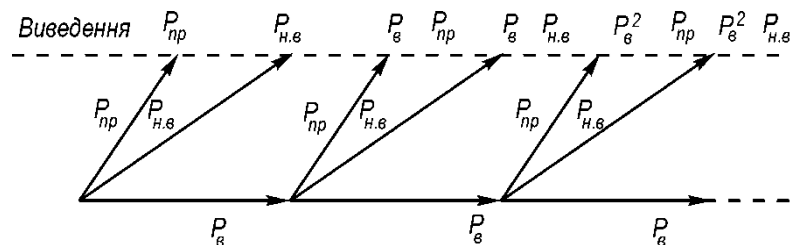


Рисунок 3.5 – Векторна діаграма станів системи з автозапитом [7]

При іншому способі коректування помилок по зворотному каналу на передавач передаються дані про кожне прийняте від нього повідомлення, і виявлення помилок, на відміну від системи з автозапитом, відбувається на передавальному кінці. Якщо в результаті аналізу виявляється помилка, то також відбувається повторення спотвореного повідомлення. Ця система називається *системою з інформаційним зворотним зв'язком* чи *системою з порівнянням*.

Зупинимось спочатку більш докладно на системах з автозапитом. Одним з методів виявлення спотворень у прийнятому повідомленні є застосування коректувальних кодів, що виявляють помилки. Визначимо ймовірність помилкових комбінацій у системі з автозапитом і виявляючим кодом. При прийманні кодової комбінації можливі три несумісні події, сумарна ймовірність яких дорівнює одиниці:

$$P_{пр} + P_{н.в} + P_{в} = 1,$$

де $P_{\text{пр}}$ – імовірність правильного прийому; $P_{\text{н.в}}$ – імовірність помилок, що не виявляються даним кодом; $P_{\text{в}}$ – імовірність помилок, що виявляються.

У разі відсутності чи невиявлення помилок комбінація надходить на вихід системи. Якщо помилки виявлені, то прийнята комбінація стирається і посилається запит для її повторення. Тому при правильному прийманні сигналу запиту ймовірність повторення комбінації дорівнює $P_{\text{в}}$. У результаті вторинного передавання комбінації знову виникає первісне положення з трьома можливими наслідками, і т. д. Процес передавання окремої комбінації в такій системі можна зобразити графічно за допомогою векторної діаграми. На рисунку 3.5 векторами показані ймовірні переходи системи з одного стану в інший при прийманні деякої комбінації.

Надалі вважатимемо, що в системі відсутнє обмеження кількості повторних передач однієї і тієї ж комбінації, а переходи системи в різні стани незалежні. За цих умов неважко визначити ймовірність помилкового переходу, що складається з l повторень комбінації, з наступним поданням її на вихід у випадку наявності невиявлених помилок. Як видно з рисунку 3.5, ця ймовірність дорівнює $P_{\text{в}}^l P_{\text{н.в}}$. Повна ймовірність помилки дорівнює сумарній імовірності всіх помилкових переходів:

$$P_{\text{пом}} = P_{\text{н.в}} + P_{\text{в}} P_{\text{н.в}} + P_{\text{в}}^2 P_{\text{н.в}} + \dots = P_{\text{н.в}} \sum_{l=0}^{\infty} P_{\text{в}}^l.$$

Обчисливши за відомою формулою суму геометричної прогресії, одержимо:

$$P_{\text{пом}} = \frac{P_{\text{н.в}}}{1 - P_{\text{в}}}. \quad (3.31)$$

Виводячи цей вираз, ми не враховували ймовірність помилки в сигналі запиту, що в багатьох випадках допустимо, оскільки за такі сигнали вибираються спеціальні кодові послідовності, котрі мають високу завадостійкість.

Якщо $P_{\text{в}} \ll 1$, то $P_{\text{пом}} \approx P_{\text{н.в}}$, і застосування системи з автозапитом дасть підвищення завадостійкості, обумовлене величиною $P_{\text{н.в}}$.

У розглянутих системах передавання повідомлень відбувається з деякою надлишковістю, обумовленою надлишковістю самого коду і повторним передаванням комбінацій, що еквівалентно збільшенню їхньої тривалості. Визначимо величину цієї надлишковості. Звичайно при виявленні помилки приймач переходить у режим очікування запитуваної комбінації, тривалість якого $\tau_{\text{оч}}$ не менша за сумарну тривалість проходження сигналів в обидва кінці. Тоді при l повтореннях час, затрачений на приймання однієї комбінації, дорівнює $\tau_{\text{к}} + l\tau_{\text{оч}}$, де $\tau_{\text{к}}$ – номінальна тривалість комбінації.

Ймовірність того, що комбінація надійшла на вихід після l повторень (рис. 3.5), визначається виразом:

$$P_l = P_{\text{в}}^l (P_{\text{пр}} + P_{\text{н.в}}) = P_{\text{в}}^l (1 - P_{\text{в}}).$$

Звідси еквівалентна середня тривалість комбінації з урахуванням втрат на повторне передавання може бути зображена сумою:

$$\begin{aligned}\bar{\tau}_k &= \sum_{l=0}^{\infty} (\tau_k + l\tau_{оч}) P_l = \sum_{l=0}^{\infty} (\tau_k + l\tau_{оч}) P_B^l (1 - P_B) = \\ &= \tau_k \left[1 + N(P_B + P_B^2 + P_B^3 + \dots) \right],\end{aligned}$$

де $N = \frac{\tau_{оч}}{\tau_k}$ – число додатково переданих комбінацій.

Використовуючи формулу суми геометричної прогресії, одержимо:

$$\bar{\tau}_k = \tau_k \left[1 + N \left(\frac{1}{1 - P_B} - 1 \right) \right] = \tau_k \frac{1 + (N - 1)P_B}{1 - P_B}.$$

Отже, повна надлишковість у системі з автозапитом при використанні n -значного виявляючого коду, буде дорівнювати :

$$x = 1 - \frac{k}{n\bar{\tau}_k / \tau_k} = 1 - \frac{k(1 - P_B)}{n[1 + (N - 1)P_B]},$$

де k – значність первинного коду; $n\bar{\tau}_k / \tau_k$ – еквівалентна значність виявляючого коду при наявності втрат, завданих повторним передаванням.

Розглянемо тепер характеристики систем з інформаційним зворотним зв'язком. У найпростішому випадку прийняті приймачем повідомлення повністю передаються по зворотному каналу на передавач, де відбувається порівняння з переданими повідомленнями. За наявності розбіжностей спотворені повідомлення повторюються. Цей різновид інформаційного зворотного зв'язку іноді називають *ретрансляційним зворотним зв'язком*. У подібній системі помилка не буде виявлена, якщо вона з'явиться в одному і тому ж символі двічі: при передаванні по прямому і зворотному каналах. Якщо ймовірності помилок у цих каналах відповідно дорівнюють P_0 і P_0^* , то можна показати, що ймовірність невиявленої помилки в одному символі при P_0 і P_0^* приблизно дорівнює їхньому добутку $P_0 P_0^*$. Оскільки P_0^* звичайно величина досить мала, виграш у завадостійкості може бути значним.

Перевага ретрансляційного зворотного зв'язку полягає в достатньо простій конструкції. Однак для ретрансляції необхідно повністю займати зворотний канал.

У більш складних системах інформаційного зворотного зв'язку застосовуються коди, що виявляють помилки. Відмінність від систем з автозапитом полягає в тому, що по прямому каналу передається тільки та частина кодової комбінації, що містить інформаційні символи, а по зворотному каналу передаються відповідні їм контрольні символи. На передавальному кінці зіставляються передані інформаційні та прийняті контрольні символи. Ті комбінації, в яких порушується відповідність між цими двома групами символів, вважаються помилковими, і вони знову повторюються. Підбираючи коди з різною надлишковістю і коректувальною здатністю, тут легше, ніж у ретрансляційній системі, забезпечити найкращі умови передавання повідомлень. Аналіз показує, що застосування кодів з надлишковістю, рівною 0,5, тобто кодів, контрольні елементи яких також повністю займають зворотний канал, дає можливість одержати більш високу завадостійкість, ніж у ретрансляційній системі.

Якщо ймовірності помилок символів у прямому і зворотному каналах однакові, повна ймовірність помилки кодової комбінації при інформаційному

зворотному зв'язку виражається тією ж формулою, що й у системі з автозапитом. Що стосується надлишковості (з урахуванням прямої і зворотної передач), ці системи також рівноцінні. Проте у випадку, коли ймовірність помилки в зворотному каналі значно менша, ніж у прямому, системи з інформаційним зворотним зв'язком мають більшу завадостійкість. У межі при нескінченно великому відношенні сигнал–завада в зворотному каналі ($P_{\text{пом}}^* = 0$) і використанні кодів з достатньою надлишковістю інформаційний зв'язок може забезпечити безпомилкове передавання в прямому каналі. У цьому легко переконатися на прикладі ретрансляційного зв'язку. Дійсно, при $P_{\text{пом}}^* = 0$ всі помилки в прямому каналі будуть виявлені, а отже, й виправлені. Близькі до цього умови мають місце на практиці при здійсненні зв'язку потужної наземної станції зі станцією якогось рухомого об'єкта: літака, ракети і т. д. Використовуючи наземну станцію для передавання сигналів зворотного зв'язку, можна забезпечити високу завадостійкість і в каналі рухомий об'єкт–Земля, де потужність сигналу, як правило, невелика.

На закінчення зазначимо, що доцільність застосування систем зі зворотним зв'язком залежить від рівня завад у прямому і зворотному каналах. При дуже сильних завадах зростає кількість повторень, і пропускна здатність системи різко падає. За цих умов більш раціональною може виявитися система без зворотного зв'язку.

3.11 Посаднання процедур демодуляції і декодування

Вище були розглянуті двохетапні процедури визначення на прийомі переданого кодового слова. На першому етапі в схемі, яку можна назвати першою розв'язувальною (1РС), визначався вид переданого елемента (символу) кодової комбінації, а на другому етапі в другій розв'язувальній схемі (2РС) за послідовністю $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ визначалося кодове слово. Друга розв'язувальна схема при двохетапній процедурі винесення рішення називається також *жорстким (дискретним) декодером*, а найчастіше просто декодером. Двохетапне декодування називають *поелементним (посимвольним)*.

Дискретний декодер виносить рішення про передане кодове слово тільки за послідовністю «0» і «1», не використовуючи інформацію про якість прийому кожного з елементів. Це, зрозуміло, не сприяє одержанню максимальної правильності рішення і, крім того, обмежує можливості декодера. Для пояснення даної тези розглянемо декодування для коду (7, 6), який при поелементному прийманні здатний тільки виявляти помилки. Якщо декодер матиме інформацію про якість прийому кожного з елементів, то можна і виправляти одиничні помилки. Якість (надійність) прийому i -го елемента визначається умовною ймовірністю його неправильного прийому $p_i(H/Y)$. Чим більше $p_i(H/Y)$, тим гірша якість прийнятого елемента (символу). Припустимо, що в послідовності двійкових елементів $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_7$ мала місце однократна помилка, яку необхідно виправити. Найбільш імовірно, що місцеположення помилки відповідає місцеположенню найменш надійного символу. Змінюючи його на протилежний, зробимо виправлення одиничної помилки. Декодування, коли використовується інформація про надійність прийнятих символів, називається *м'яким або аналоговим*. Чим повніша інформація

про надійність символів на декодері, тим менша ймовірність неправильного декодування. Найбільш повне використання інформації про надійність прийнятих символів відповідає *прийому в цілому*. Свою назву такий метод прийому одержав унаслідок розгляду тракту приймання як єдиної розв'язувальної схеми, що виносить рішення про передане кодове слово на підставі аналізу сигналу $Y(t) = \{y_1(t), y_2(t), \dots, y_n(t)\}$ без попереднього визначення виду переданих елементів, тобто аналізу всього сигналу в цілому, з використанням всієї інформації про сигнал на вході приймача $Y(t)$. Реалізувати прийом у цілому в умовах гауссових завад і дискретного каналу без пам'яті можна, використовуючи схему, аналогічну схемі оптимального приймача. Ця схема дозволяє здійснити розділення двох сигналів, що відповідають символам «0» і «1». У нашому випадку число сигналів, які потрібно розділити, дорівнює 2^k , де k – число інформаційних елементів у кодовому слові. Якщо $k = 20$, то треба розділити 1 048 576 сигналів і, отже, таке ж число гілок прийому. Реалізувати такий приймач, навіть на сучасній елементній базі, неможливо.

Застосування прийому в цілому порівняно з поелементним прийомом по гауссовому каналу без пам'яті дає енергетичний виграв приблизно в 3 дБ при фіксованій імовірності помилкового декодування. Для каналів зі змінними параметрами і пам'яттю цей виграв буде значно більшим і становитиме вже 10...20 дБ.

Складність прийому в цілому змушує шукати такі двохетапні процедури винесення рішення про передане кодове слово, які забезпечують більш просту реалізацію тракту приймання, ніж при прийомі в цілому, і в той же час використовують інформацію про надійність прийнятих елементів, тобто тією чи іншою мірою поєднують процедури демодуляції і декодування. Залежно від ступеня використання інформації про надійність елементів буде забезпечене більше чи менше наближення за завадостійкістю до прийому в цілому.

Розрізняють два підходи до проблеми м'якого декодування. При першому мінімізується ймовірність неправильного декодування кодового слова, при другому мінімізується середня ймовірність помилки на елемент. В останньому випадку декодована послідовність може і не бути кодовим словом (дозволеною кодовою комбінацією). Характеристики декодерів, що реалізують ці два правила, збігаються, оскільки мінімізація ймовірності помилки на елемент веде до мінімізації ймовірності помилки в послідовності елементів, і навпаки.

Найбільш часто на практиці розв'язується задача мінімізації ймовірності неправильного декодування. При цьому враховуються обмеження на складність реалізації приймального тракту. В усіх цих випадках робиться спроба породити відносно невелике число кодових слів, серед яких з великою ймовірністю міститься передане кодове слово, що знаходиться на найменшій відстані від прийнятої послідовності елементів (символів). Методи мінімізації ймовірності неправильного декодування розрізняються способом породження цих кодових слів. Нижче розглядаються деякі з них.

Приймання за найбільш надійними символами. Ідея методу декодування за найбільш надійними символами (елементами) запропонована Л. Ф. Бородінін і може бути пояснена в такий спосіб. Визначимо вид двійкових символів $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ та їх надійність $\gamma_i, i = \overline{1, N}$. Нехай найменш надійні символи розташовані на перевірних позиціях. Зітремо ці символи. Тоді, якщо інші k символів, що залишилися, прийняті

без помилок, то послідовність $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_k$ буде правильно декодована, тобто буде виправлено r стирань. Відомо, що будь-який код здатний виправити до $(d_0 - 1)$ стирань, якщо на нестертих позиціях не було помилок. Звідси випливає, що число символів n_1 , за якими може бути винесене рішення про передане кодове слово, визначається нерівністю $k \leq n_1 \leq n - (d_0 - 1)$. На кодувальній пристрій надходить послідовність одиниць і нулів, а також інформація про надійність кожного із символів прийнятої послідовності $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$.

Декодування починається з виділення k найбільш надійних символів і перевірки можливості однозначного декодування за цими символами. Якщо однозначне декодування неможливе, то додається ще один найбільш надійний з $n - k$ символів, що залишилися, і всі операції повторюються. При прийманні за методом Бородіна можуть бути виправлені всі помилки (стирання) кратності $t \leq d_0 - 1$, якщо кожний з $n - d_0 + 1$ безпомилково прийнятих символів має міру надійності більшу, ніж у кожного з неправильно прийнятих.

Для гауссових каналів метод Бородіна забезпечує в порівнянні з поелементним прийманням вигоду приблизно в 1,4 дБ, а для коду з $k = n - 1$ таку ж саму завадостійкість, як і метод прийому в цілому.

Більш складними в порівнянні з методом Бородіна є алгоритми декодування, запропоновані Чейзом.

Клас алгоритмів Чейза. Запропонована Чейзом процедура полягає в наступному.

1. Визначається значення кожного елемента (символу), що входить у кодову комбінацію. В результаті цього одержуємо вектор $\hat{A} = \{\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n\}$, який складається з 0 і 1.

2. На прийняту кодову комбінацію \hat{A} накладаються сформовані певним чином тестові послідовності T . Отримані послідовності $\hat{A}^* = \hat{A} \oplus T$ декодуються жорстким (дискретним) декодером за раніше розглянутими правилами. У результаті одержуємо кодові слова $\hat{A}'_1, \hat{A}'_2, \dots, \hat{A}'_N$, де N – число тестових послідовностей.

3. З урахуванням вірогідності (надійності) кожного з елементів визначаємо відстань послідовності \hat{A} до кожного з кодових слів \hat{A}'_i , де $i = 1, 2, \dots$. Прийнятим вважається те кодове слово, відстань до якого буде мінімальною. Відстань L_j до j -го кодового слова можна обчислити за формулою:

$$L_j = \sum_{i=1}^n \gamma_i (\hat{a}_i \oplus \hat{a}'_{ij}); \quad j = \overline{1, N}, \quad (3.32)$$

де γ_i – додатне число, що характеризує ступінь надійності прийнятого i -го елемента. Наприклад, при восьмирівневому квантуванні сигналів на виході демодулятора це можуть бути числа $0 \dots 7$; \hat{a}'_{ij} – i -й символ j -го кодового слова \hat{A}'_j ; послідовність $\hat{a}_i \oplus \hat{a}'_{ij}$, $i = \overline{1, n}$ визначає вектор помилки E'_j .

На схемі, що реалізує декодування за Чейзом (рис. 3.6), подвійними лініями зображені векторні зв'язки. На вхід схеми з виходу неперервного каналу подається сигнал $V(t)$, який демодулюється; рівень сигналу на виході демодулятора визначає

надійність прийнятих елементів. Вид елемента визначається в першій розв'язувальній схемі (1РС).

Три запропонованих Чейзом алгоритми відрізняються методом формування тестових послідовностей T . Алгоритм 1 потребує формування найбільшого числа послідовностей і забезпечує найкращу завадостійкість. В алгоритмах 2 і 3 формується менше число тестових послідовностей і відповідно забезпечується гірша завадостійкість.

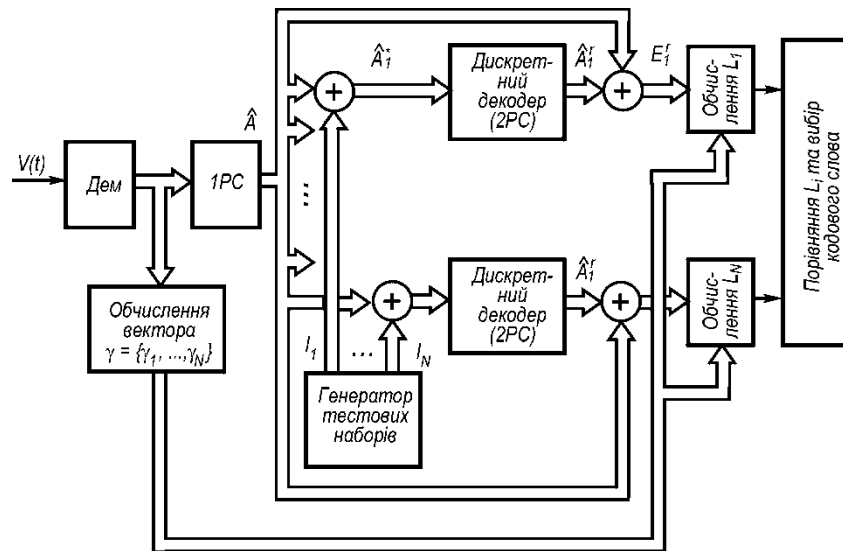


Рисунок 3.6 – Декодер Чейза [6]

Для алгоритму 1 потрібно взяти як тестові послідовності нульовий вектор і всі комбінації ваги $[d_0/2]$; кількість таких тестових комбінацій $C_n^{[d_0/2]}$. Для алгоритму 2 використовується $2^{[d_0/2]}$ тестові послідовності, що мають різні символи на $[d_0/2]$ найменш надійних позиціях і нульові символи – на інших. Для алгоритму 3 визначається $d_0 - 1$ найменш надійних символів. Тестові послідовності мають одиниці на i найменш надійних позиціях і нулі на інших ($i=0, 2, 4, \dots$ для непарного d_0 ; $i=0, 1, 3, \dots$ для парного d_0). Загальне число можливих тестових послідовностей $[d_0/2 + 1]$.

Приклад 3.1. Визначити необхідне число тестових послідовностей для коду $(7, 3)$ з кодовою відстанню $d_0 = 4$.

Для алгоритму 1 число тестових послідовностей $N = C_7^2 = 21$; для алгоритму 2 $N = 2^{[d_0/2]} = 2^2 = 4$; для алгоритму 3 $N = [d_0/2 + 1] = 2 + 1 = 3$. Отже, найменше число тестових послідовностей потрібно при використанні алгоритму 3.

Розглянемо на прикладі процедуру декодування за Чейзом для алгоритму 3.

Приклад 3.2. Код $(6, 3)$ заданий твірною матрицею. Передано кодову комбінацію 000000. Внаслідок дії завад прийнято кодову комбінацію з помилками на позиціях 2 і 3 $\hat{A} = 011000$. Оскільки $d_0 = 3$, то необхідно сформулювати дві тестові послідовності, перша з яких 000000, а друга містить 2 одиниці на позиціях найменш надійних символів. Нехай надійність символів визначається числами 732656, тобто найменш надійними є символи, що розміщені на позиціях 2 і 3. Тоді

друга тестова послідовність має вигляд 011000, і в результаті складання прийнятої послідовності з тестовими маємо

$$\begin{array}{r} 011000 \\ \oplus \\ 000000 \\ \hline 011000 = \hat{A}_1^* \end{array} \quad \begin{array}{r} 011000 \\ \oplus \\ 011000 \\ \hline 000000 = \hat{A}_2^* \end{array}$$

Після декодування жорстким декодером одержимо кодові слова $A_1' = 011100$ і $A_2' = 000000$. Обчисливши L за формулою (3.32) для першого і другого кодових слів, одержимо відповідно $L_1 = 6$ і $L_2 = 5$. Таким чином, робиться висновок про те, що передавалося кодове слово 000000. У розглянутій вище ситуації вдалося виправити дворазову помилку.

При розробці систем передавання дискретних повідомлень доводиться розв'язувати не тільки задачі поєднання процедур демодуляції і декодування, але й модуляції та кодування, тому що характеристики дискретного каналу залежать від виду модуляції. Найбільш загальний підхід до розв'язання цих задач зводиться до того, що кодування і модуляція розглядаються як єдиний процес формування найкращого сигналу, а демодуляція і декодування – як процес найкращої обробки прийнятого сигналу.

3.12 Згорткові (гратчасті) коди

Вище розглядалися блокові коди, коли значення елементів, що входять у різні блоки, не залежали одне від одного.

Для систематичних двійкових блокових (n, k) -кодів послідовність інформаційних символів джерела розбивається на блоки довжиною k біт, потім у кодері до кожного такого блока додається $r = n - k$ перевірних символів, після чого блоки довжиною n символів передаються в канал зв'язку. Декодування блоків також здійснюється незалежно одне від одного.

Однак можливий і інший, *неперервний* принцип кодування і декодування, коли на вхід кодера надходить *неперервна* послідовність інформаційних символів джерела, а з виходу кодера знімається також неперервна послідовність символів, котрі є функцією вхідних символів і структури кодера. У декодері такого типу на вхід надходить неперервна послідовність символів з каналу зв'язку (можливо, спотворена помилками), а на виході відновлюється (можливо, з помилками, але, як правило, меншими за канальні) послідовність інформаційних символів. Найпоширенішим класом неперервних кодів є *згорткові* коди, для яких операція формування вихідної послідовності за заданою вхідною послідовністю є лінійною. Згорткові коди були відкриті Л. Фінком і П. Елайєсом, незабаром Возенкрафт розробив метод послідовного декодування згорткового коду.

В 1967 р. з'явився алгоритм Вітербі для оптимального декодування згорткових кодів.

Структура двійкового згорткового коду зі швидкістю $R = \frac{k}{n}$ показана на рисунку 3.7, а. Згортковий кодер складається зі зсувного регістра, що містить v чарунок пам'яті, і блока суматорів за mod 2, входи кожного з яких зв'язані з деякими виходами чарунок пам'яті регістра, обумовленими коефіцієнтами $h_{ij} = (0, 1)$. Виходи суматорів зчитуються за допомогою комутатора K та подаються в канал зв'язку. Таким чином, на кожному такті в регістр зсуву послідовно надходить

черговий блок з k інформаційних символів джерела й одночасно він звільняється від k символів, що містяться в його крайніх правих чарунках пам'яті. На цьому ж такті формуються n вихідних символів, що послідовно зчитуються в канал зв'язку. Отже, якщо v_i – швидкість надходження символів у кодер, то для відсутності зростаючих затримок у часі швидкість передавання символів по каналу зв'язку має бути не меншою, ніж $v_k = \frac{n}{k} v_i$, звідки випливає, що відношення k/n дійсно визначає швидкість згорткового коду. Величина v (чи довжина регістра) звичайно називається *довжиною кодового обмеження*. У деяких роботах довжина кодового обмеження визначається числом чарунок зсувного регістру мінус 1.

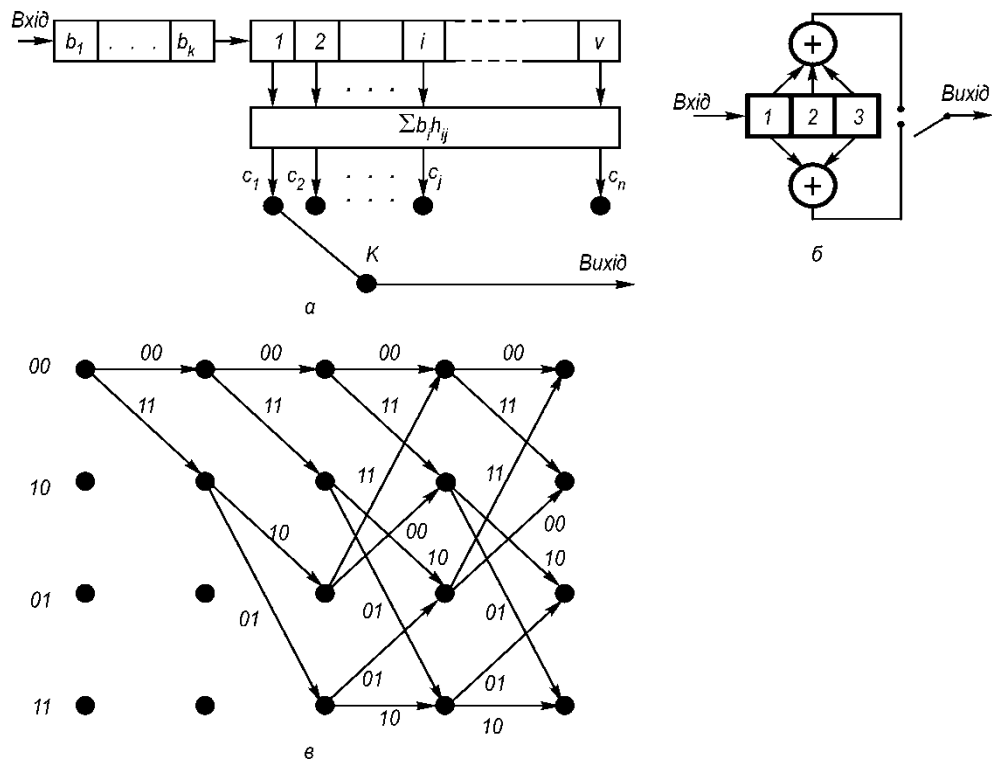


Рисунок 3.7 – Структури згорткових кодів зі швидкістю $R = k/n$ (а), зі швидкістю $R = 1/2$ (б) та ґратка кодера зі швидкістю $R = 1/2$ (в) [6]

Можливе й інше, більш загальне зображення згорткового кодера у вигляді схеми з k регістрами зсуву з кодовими обмеженнями v_i ($i=1, 2, \dots, k$). На вхід кожного з них подається один інформаційний символ за час одного такту.

На рисунку 3.7, б наведено окремий випадок згорткового коду зі швидкістю $R=1/2$ і довжиною кодового обмеження $v=3$. При нульовій інформаційній послідовності вихідна кодова послідовність також дорівнює нулю. Нижче розглянуто приклад формування вихідної послідовності для кодера, показаного на рисунку 3.7, б:

Вхід	0	0	1	0	0	0	1	1	0	0	0
Вихід			11	10	11	00	11	01	01	11	00

Вихідна послідовність кодера може бути подана як цифрова згортка вхідної інформаційної послідовності й імпульсного відклику кодера (звідси назва кодів – *згорткові*).

Згортковий код характеризується такими параметрами: *відносною швидкістю* коду $R = k/n$ і *надлишковістю* $\iota = 1 - R$, де k та n – число інформаційних і кодових символів, що відповідають одному такту роботи кодера (для кодера на рисунку 3.7, б $R = 1/2$); *довжиною кодового обмеження* ν (довжина регістра кодера); *породжувальним поліномом коду*, коефіцієнти яких описують зв'язки суматорів з чарунками регістра кодера (для верхнього суматора $g^{(1)} = 1 + D + D^2$, для нижнього суматора $g^{(2)} = 1 + D^2$). Поліноми звичайно записують скорочено, позначаючи кожні три відводи (двійкові коефіцієнти) як одну вісімкову цифру.

Крім названих параметрів згортковий код характеризується *вільною відстанню* d_v , під якою розуміють відстань за Хеммінгом між двома напівнескінченими кодовими послідовностями. Якщо дві однакові інформаційні послідовності кодувати за допомогою кодера, зображеного на рисунку 3.7, б, то відповідні їм кодові послідовності збігатимуться одна з одною. Якщо в деякий момент в одній інформаційній послідовності опиниться символ 0, а в іншій 1, то з цього моменту кодові послідовності будуть відрізнятися одна від одної незалежно від подальшого змісту інформаційних послідовностей. Мінімальна відстань за Хеммінгом між будь-якими двома напівнескінченими кодовими послідовностями з того моменту, як відповідні їм інформаційні послідовності починають відрізнятися, називається вільною відстанню згорткового коду d_v .

Вільна відстань d_v характеризує завадозахисні властивості згорткового коду (аналогічно тому, як мінімальна відстань d характеризує завадозахисні властивості блокових кодів). Вона показує, яке найменше число помилок має статися в каналі, щоб одна кодова послідовність перейшла в іншу і помилки не були виявлені. Для коду, наведеного в нашому прикладі, вільна відстань $d_v = 5$.

Пошук гарних згорткових кодів (з найбільшим d_v при заданих R і ν) звичайно здійснюється методом перебору всіх породжувальних поліномів на ЕОМ.

Згорткові коди є окремим випадком (лінійною реалізацією) *гратчастих* кодів. Можна також припустити, що ґратки є просто іншим (іноді більш зручним) способом подання і звичайних згорткових кодів.

ґратками називається орієнтований граф з періодично повторюваною структурою «чарунок». Кожна чарунка містить стовпчики з однакового числа вершин (вузлів), з'єднаних ребрами. Між процедурою кодування згортковим кодом і ґратками існує взаємно однозначна відповідність, яка задається такими правилами:

- кожна вершина (вузол) відповідає внутрішньому стану кодера;
- ребро, що виходить з кожної вершини, відповідає одному з можливих символів джерела (для двійкового джерела з кожної вершини виходять два ребра – верхнє для 0 і нижнє для 1);
- над кожним ребром проставлені значення символів, переданих у канал зв'язку, якщо кодер знаходився в стані, який відповідає даній вершині, і джерело видало символ, що відповідає даному ребру;
- послідовність ребер (шлях на ґратках) – це послідовність символів, виданих джерелом.

Так, якщо під станом кодера розуміти вміст двох останніх чарунок пам'яті (2, 3) у регістрі зсуву на рисунку 3.7, б, то ґратка з чотирма станами, що відповідає даному кодеру, матиме вигляд, показаний на рисунку 3.7, в (ґратка може

відображати і нелінійний кодер, коли вихідні символи не є лінійною функцією вхідних). Так само як і блокові коди, згорткові допускають зображення у вигляді напівнескінчених породжувальних чи перевірних матриць, однак зображення у вигляді ґратки є зручнішим для опису алгоритмів декодування.

Згорткові коди мають такі основні переваги над блоковими при їхньому використанні для виправлення помилок.

1. Вони не потребують синхронізації по блоках; необхідна лише синхронізація комутаторів K (на передачі та прийомі).

2. Якщо кодове обмеження v вибрати рівним довжині блокового коду, то виправна здатність згорткового коду буде більшою, ніж виправна здатність такого блокового коду (при найкращому виборі обох кодів).

3. Алгоритм декодування згорткових кодів допускає просте узагальнення на випадок м'якого декодування, що забезпечує додатковий енергетичний вигравш.

4. Згорткові коди допускають просте об'єднання кодування і модуляції (так звана *кодована модуляція* чи *сигнально-кодові конструкції*), що особливо важливо при побудові енергетично ефективних систем зв'язку для каналів з обмеженою смугою частот.

Для оптимального декодування згорткових кодів у каналах без пам'яті часто використовується рекурентний *алгоритм декодування Вітербі* (АВ). Розглянемо його на прикладі м'якого декодування в постійному каналі з адитивним білим гауссовим шумом.

Оскільки сигнал, що приймається на k -му тактовому інтервалі, нам відомий, то можна обчислити евклідові (або гільбертові) відстані між прийнятим сигналом і всіма можливими сигналами:

$$\Delta_{ki} = \int_0^T [z_k(t) - s_k^{(i)}(t)]^2 dt, \quad i = 0, 1, \dots, m-1, \quad k = 0, 1, \dots,$$

де $s_k^{(i)}(t)$ – очікуваний у місці прийому сигнал, що відповідає i -му символу (для двійкових сигналів $i = 0, 1$); $z_k(t)$ – сигнал, прийнятий на k -му тактовому інтервалі.

Тепер можна кожному ребру ґратки послідовно приписувати на k -х її ланках значення Δ_{ki} . Оптимальне за правилом максимальної правдоподібності декодування тоді відповідатиме вибору такого шляху на ґратках (тобто послідовності неперервно продовжуваних ребер), що $\sum_k \Delta_{ki}$ буде мінімальною.

Здавалося б, для ґратки завдовжки n (тобто послідовності переданих символів завдовжки n) потрібно перебрати 2^n можливі варіанти, але насправді це не так. Ключовий момент АВ полягає в тому, що для кожної вершини на даному кроці (такті) існує безліч метрик, які відповідають з'єднанням з нею ребрами вершинам. На попередньому кроці можна залишити тільки одне ребро, що мінімізує суму метрик на всіх попередніх кроках.

Найпростіше можна пояснити даний алгоритм на такому прикладі. Нехай ґратки мають усього два стани і структуру, показану на рисунку 3.8, *а*, де над ребрами поставлені відповідні метрики. Вважаємо, що перший інформаційний символ 0. Тоді шляхи, залишені (тобто такі, що «вижили») на різних кроках, показані на рисунку 3.8, *б*. Видно, що на 4-му кроці одержуємо «виживший» шлях, котрий в умовах наших позначень (орієнтація ребра вниз – 1, вгору – 0) відповідає інформаційній послідовності 0100.

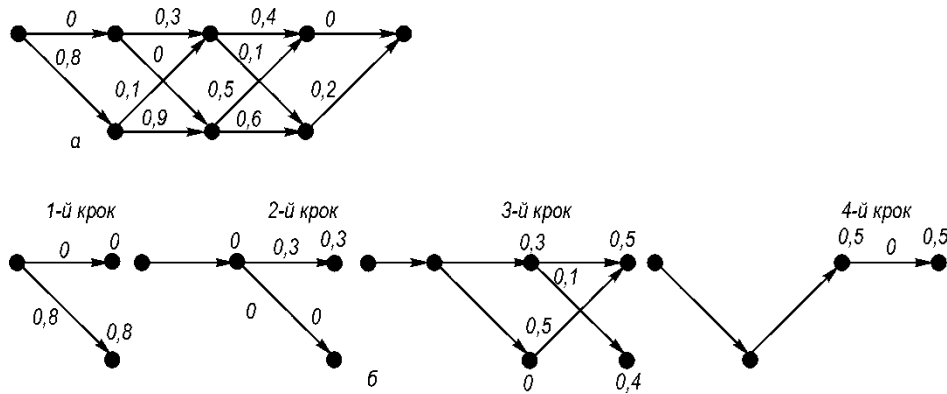


Рисунок 3.8 – Гратка з метриками (а) та побудова шляху, що «вижив», за алгоритмом Вітербі (б) [5]

Складність АВ визначається на кожному кроці числом порівнянь метрик, що з'єднують усі вершини, і воно обмежене величиною M^2 , де M – число станів гратки. Оскільки зі схеми згорткового кодера одержуємо, що $M = 2^{v-1}$, де v – число чарунок пам'яті регістра зсуву кодера, то бачимо, що складність АВ експоненціально залежить від довжини кодових обмежень, але лінійно залежить від довжини послідовності, що передається. Тому довжина кодових обмежень v при використанні АВ як алгоритму декодування звичайно вибирається не більшою за 10...15, що, утім, цілком достатньо для одержання великого енергетичного вигаду. АВ вимагає обробки всієї послідовності сигналів для оптимального декодування навіть першого інформаційного символу. Така процедура потребує значної пам'яті на прийомі та затримки для декодування елементів повідомлення. Для усунуння цих недоліків використовується модифікація АВ у вигляді *зрізаного алгоритму*, коли рішення про інформаційний символ на i -му такті приймається за результатами обробки по АВ послідовності символів на даному i -му і L наступних тактових інтервалах. Теорія й експеримент показують, що коли L вибрати порядку декількох довжин кодових обмежень, то енергетичні втрати при використанні такої модифікації будуть невеликими.

Зазначимо, що АВ є ефективним методом розв'язання значно загальнішої оптимізаційної задачі. Нехай потрібно знайти такий дискретний вектор $\mathbf{x}_N = (x_1, \dots, x_N)$, $x_i \in \mathbf{X}$, $|\mathbf{X}| = r$, що максимізує (мінімізує) функцію $\Lambda(\mathbf{x}_N)$. Тоді прямий метод вимагає перебору r^N різних векторів. Однак, якщо функція $\Lambda(\mathbf{x}_N)$ допускає зображення:

$$\Lambda(\mathbf{x}_N) = \sum_{k=1}^N \lambda_k(\sigma_k),$$

де $\lambda_k(\sigma_k)$ – довільні функції векторних аргументів вигляду:

$$\sigma_k = \begin{cases} (x_{k-v}, x_{k-v+1}, \dots, x_k), & k > v, \\ (x_1, \dots, x_k), & k \leq v, \end{cases}$$

то можна довести, що максимум такої функції знаходиться за допомогою наступного *узгальненого АВ*.

Крок 1. Знайти $\hat{x}_1(x_2, \dots, x_{v+1}) = \text{Arg max}_{x_1} \Lambda_{v+1}$ (при \hat{x}_{v+1}). (Оскільки $|\mathbf{X}| = r$, то для кожного значення аргументу (x_2, \dots, x_{v+1}) це потребує не більше ніж r обчислень.)

Крок 2. Знайти $\hat{x}_2(x_3, \dots, x_{v+2}) = \text{Arg max}_{x_2} \Lambda_{v+2}$ (при \hat{x}_{v+2}), де $\hat{x}_{v+2} = (\hat{x}_1, \hat{x}_2, \dots, x_{v+2})$, а \hat{x}_1 було знайдено на 1-му кроці, і т. д.

Крок S . Знайти $\hat{x}_S(x_{S+1}, \dots, x_{v+S}) = \text{Arg max}_{x_S} \Lambda_{v+S}$ (при \hat{x}_{v+S}), де $\hat{x}_{v+S} = (\hat{x}_1, \hat{x}_2, \dots, x_{S-1}, x_S, \dots, x_{v+S})$, а $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{S-1}$ були знайдені на попередніх кроках, і т. д.

Крок $N-v$. Знайти $N-v$ $(\hat{x}_{N-v}, \dots, \hat{x}_N) = \text{Arg max}_{(x_{N-v}, \dots, x_N)} \Lambda_N$ (при \hat{x}_N), де $\hat{x}_N = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{N-v-1}, x_{N-v}, \dots, x_N)$, а $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{N-v-1}$ були знайдені на всіх попередніх кроках.

Таким чином, узагальнений АВ має поліноміальну залежність від довжини послідовності N на відміну від експоненціальної залежності при прямому методі перебору всіх альтернативних гіпотез.

Вибираючи різні форми функцій $\lambda_k(\sigma_k)$, можна одержати досить прості методи розв'язання оптимізаційних задач, а отже, і оптимальні алгоритми обробки сигналів не тільки для згорткових кодів, але й для інших моделей каналів, наприклад для каналу з міжсимвольною інтерференцією.

Універсальність АВ полягає в тому, що він може бути використаний для різних розподілів сигналів і завад, неоднорідних каналів, для поєднання декодування й демодуляції і не тільки для незалежних розподілів, але й для випадків залежності, описуваних марковськими послідовностями.

У каналах із МСІ алгоритм Кловського–Ніколаєва (АКН) зі зворотним зв'язком за рішенням при фіксованій затримці прийняття рішення $D = Q$ (Q – пам'ять каналу) практично не поступається за завадостійкістю АВ з тією ж затримкою рішення, проте реалізується простіше. АКН, отриманий спочатку для субоптимального поелементного приймання дискретних повідомлень в каналах з розсіюванням (МСІ), можна використовувати і для спільної демодуляції–декодування при згортковому кодуванні, тому що згортковий кодер є аналогом деякого каналу з розсіюванням. Природно, що в цьому випадку затримка в прийнятті рішення L має враховувати не тільки пам'ять каналу (Q), але і довжину кодового обмеження.

Згорткові коди можуть декодуватися й іншими алгоритмами (наприклад, послідовного декодування і синдромного декодування), що не є, взагалі кажучи, оптимальними.

Послідовне декодування було ведене Возенкрафтом, однак найбільш широко використовуваний алгоритм належить Фано.

У той час, коли відповідно до алгоритму Вітербі виконується просування і відновлення метрики для всіх шляхів, які можуть здаватися нам кращими, послідовний декодер істотно обмежує число шляхів, які фактично відновляються. Основна ідея послідовного декодування полягає в тому, що використовуватися має лише той шлях, який виглядає найбільш імовірним. Через обмеженість пошуку при декодуванні ніколи не можна бути цілком упевненим, що цей шлях є найкращим. Цей підхід може розглядатися як метод спроб і помилок для пошуку правильного шляху на кодовому дереві. Такий пошук здійснюється послідовно, так, що в кожен момент відбувається обробка лише одного шляху. Однак декодер має можливість повернутися назад при пошуку найкращого рішення.

Два найбільш часто використовуваних методи синдромного декодування – декодування шляхом табличного пошуку і порогове декодування – застосовуються

при декодуванні як згорткових, так і блокових кодів. Про табличний метод синдромного декодування ми вже говорили. Порогове декодування – це метод, в основі якого лежить спеціальний вибір породжувальних многочленів коду, що допускають розв’язання перевірних рівнянь за допомогою мажоритарного вибору символу помилки за деякими оцінками. Останні отримуються додаванням одного чи кількох символів синдрому.

Для оцінки якості м’якого декодування за допомогою АВ скористаємося означенням: мінімальною евклідовою відстанню d_e згорткового коду при обраному методі модуляції називається мінімальна сума евклідових метрик помилкових шляхів на гратках, що починаються і закінчуються на правильному шляху.

Тоді для ймовірності помилки p_e в першому символі в детермінованому каналі без пам’яті з білим шумом при декодуванні по АВ буде слухна така границя:

$$p_e \leq N(d_e) e^{-\frac{d_e^2}{4N_0}},$$

де $N(d_e)$ – число шляхів з евклідовою відстанню d_e , які починаються і закінчуються на правильному шляху; N_0 – спектральна щільність потужності білого шуму.

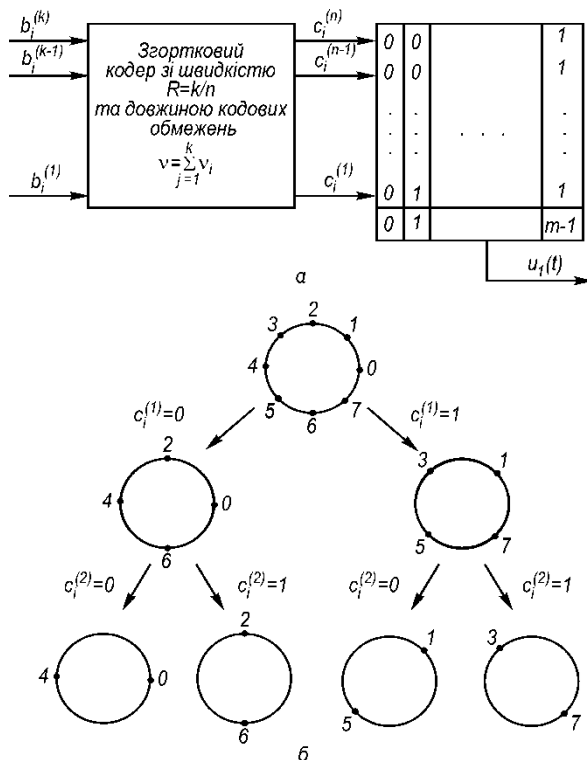


Рисунок 3.9 – Система з кодовою модуляцією (а) та розбиття вісімкових ФМ-сигналів на сузір’я (б) [5]

Використовуючи більш повно структуру гратки згорткового коду, можна оцінити також і середню ймовірність помилки на біт p_b , причому ця величина буде також експоненціально залежати від d_e^2 .

Для одержання найбільшої енергетичної ефективності, особливо в каналах з пам’яттю, доцільно використовувати каскадні коди з внутрішніми згортковими кодами і м’яким декодуванням по АВ і зовнішніми РС-кодами з використанням алгебраїчних методів декодування. Така конструкція дозволяє одержати

енергетичний вигравш, що досягає 5 дБ при еквівалентній імовірності помилки $p_e = 10^{-5}$ і прийнятній складності декодування. Зазначимо, що АВ можна використовувати і для декодування блокових кодів, якщо ці коди можуть бути описані за допомогою ґраток.

Для того щоб одержати одночасно найкращу енергетичну і частотну ефективність, використовується *кодована модуляція*, або за іншою термінологією – певні *сигнально-кодові конструкції* (СКК). Загальну ідею такого методу ілюструє рисунок 3.9, а, де індексом i позначається номер такту.

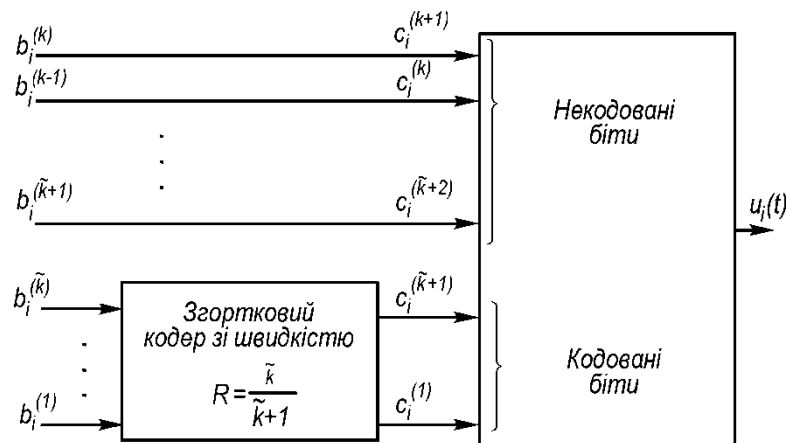


Рисунок 3.10 – Кодер Унгербоєка [5]

В цьому випадку кодер згорткового коду зі швидкістю k/n зручніше подавати як паралельний набір k регістрів зсуву з різними довжинами кодових обмежень v_i . На виході такого кодера у кожен момент часу з'являється n -вимірний двійковий вектор, який відображається в один з 2^n неперервних сигналів, що передаються в канал зв'язку. Для каналів з обмеженою смугою типовим є використання сигналів з багаторазовою фазовою або амплітудно-фазовою модуляцією, причому ключовим моментом тут є *розбиття множини сигналів на сузір'я*.

Приклад такого розбиття для вісімкової ФМ показаний на рисунку 3.9, б. Сузір'я знаходяться в нижньому ряді рисунка. На рисунку 3.10 наведена схема кодування Унгербоєка, в якій сигнали зображені у вигляді сузір'їв.

Тут двійкова послідовність символів джерела розбивається на блоки по k біт, і перші \tilde{l} біт цих блоків подаються на вхід згорткового кодера, а ті, що залишилися, надходять на модулятор у некодованому вигляді. Загальний принцип кодування полягає в тому, що некодовані біти вибирають сигнал у сузір'ї, а кодовані біти визначають вибір сузір'я. (Наприклад, для схеми розбиття, показаної на рисунку 3.9, б, кодер Унгербоєка, зображений на рисунку 3.10, має параметри: $k = 2$, $\tilde{l} = 1$, $R = 1/2$).

Оскільки, як видно з рисунку 3.9, б, сигнали в кожному із сузір'їв віддалені на значну евклідову відстань, а мінімальна евклідова відстань між сигналами різних сузір'їв може бути мала, то використання розглянутого вище принципу дозволяє максимізувати евклідову відстань між послідовностями кодованих сигналів при збереженні високої спектральної ефективності.

Розроблено спеціальну техніку, що дозволяє об'єднати згорткові коди й амплітудно-фазову модуляцію сигналу для забезпечення високої енергетичної та спектральної ефективності.

При практичному використанні завадостійких кодів головним обмеженням є складність пристрою декодування, яка може бути виражена або числом логічних схем у декодері, або кількістю обчислювальних операцій, необхідних для декодування. Тому серед кодів, що забезпечують заданий вигравш, слід вибирати ті, котрі допускають менш складну реалізацію, або навпаки, при заданій складності декодування слід вибирати коди, що забезпечують найбільший вигравш.

Контрольні запитання для самооцінки рівня знань

1. Дайте означення коду.
2. Дайте означення кодової відстані.
3. Як визначається кодова відстань?
4. Які коди називаються роздільними, а які нероздільними?
5. Як визначається вага кодової комбінації?
6. Яка матриця називається породжувальною?
7. Що називається синдромом?
8. Який код називається кодом Хеммінга?
9. Який код називається циклічним?
10. Які відмінні риси кодів БЧХ?
11. Які особливості вкорочених циклічних кодів?
12. Які коди називаються несистематичними?
13. Які особливості ітеративних кодів?
14. Які системи зі зворотним зв'язком Ви знаєте?
15. У чому полягають переваги і недоліки систем зі зворотним зв'язком?
16. Які особливості згорткових кодів?
17. Що таке сигнально-кодова конструкція?
18. У чому полягають особливості алгоритму Вітербі?
19. Що називається ґратками?

ПЕРЕЛІК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

1. Горєв В. М. Теорія електричних кіл та сигналів. Частина I. Навчальний посібник. Дніпро: Дніпровська політехніка, 2021. 104 с.
2. Горєв В. М. Теорія електричних кіл та сигналів. Частина II. Навчальний посібник. Дніпро: Дніпровська політехніка, 2021. 107 с.
3. Захаров І. П., Милютченко І. О. Основи аналізу лінійних електричних кіл: навч. посіб. Харків. нац. ун-т радіоелектроніки. Харків: ХНУРЕ, 2021. 102 с.
4. Панченко С. В., Ананьєва О. М., Бабаєв М. М., та ін. Теорія електричних і магнітних кіл: Підручник. 2-ге вид., випр. та допов. Харків: УкрДУЗТ, 2020. 246 с.
5. Bhooshan S. Fundamentals of Analogue and Digital Communication Systems. Springer Singapore. 2022. 645 p.
6. Harris J. F. Multirate Signal Processing For Communication Systems. Prentice Hall, 2021. 600 p.
7. Plonus M. Electronics and Communications for Scientists and Engineers. Elsevier Inc. 2020. 498 p.
8. Yakymchuk, N., Selepyna, Y., Yevsiuk, M., Prystupa, S., Moroz, S.: Monitoring of Link-Level Congestion in Telecommunication Systems Using Information Criteria. Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska, 12(4), 26-30. (2022).
9. Zablotskyi, V., Selepyna, Y., Lyshuk, V., Yakymchuk, N., Tkachuk, A.: Method for Evaluation Quality Parameters of Telecommunications Services. Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska, 12(2), 30-33 (2022).
10. Мороз С. А., Селепина Й. Р., Приступа С. О., Король О. О. Особливості забезпечення безпеки даних в GSM каналі мобільного зв'язку. Збірник наукових праць «Перспективні технології та прилади». № 17, Луцьк. 2020 С. 93–98.

С 66

Сигнали і коди телекомунікаційних та інформаційних систем: конспект лекцій для здобувачів другого (магістерського) рівня вищої освіти освітньої програми «Телекомунікації та радіотехніка» галузі знань галузі знань G Інженерія, виробництво та будівництво спеціальності G5 Електроніка, електронні комунікації, приладобудування та радіотехніка денної та заочної форм навчання / уклад. Й. Р. Селепина. Луцьк: ЛНТУ, 2025. 108 с.

Конспект лекцій з дисципліни «**Сигнали і коди телекомунікаційних та інформаційних систем**»: складений відповідно до діючої програми курсу.

Призначений для здобувачів вищої освіти спеціальності G5 Електроніка, електронні комунікації, приладобудування та радіотехніка освітньої програми «Телекомунікації та радіотехніка».

Комп'ютерний набір

Й. Р. Селепина

Редактор

Й. Р. Селепина

Підп. до друку «__» _____ 2025 р.
Формат 60x84/16. Папір офс. Гарнітура Таймс.
Ум. друк. арк. ____ . Тираж 10 прим. Зам. ____

Відділ іміджу та промоцій
Луцького національного технічного університету
43018, м. Луцьк, вул. Львівська, 75