

**В.П. Шульга, Л.Н. Беркман, Н.В. Галаган,
К.С. Нестеренко, О.Г. Варфоломеєва, Н.С. Хаб'юк**

СУЧАСНІ МЕТОДИ ПЕРЕДАЧІ ТА ОБРОБКИ ІНФОРМАЦІЇ В ІНФОКОМУНІКАЦІЙНИХ МЕРЕЖАХ

Київ - 2024

Автори:

В.П. Шульга, Л.Н. Беркман, Н.В. Галаган,
К.С. Нестеренко, О.Г. Варфоломєєва, Н.С. Хаб'юк

Рецензенти:

д-р техн. наук, проф. КНУ ім. Тараса Шевченка **О.М. Ткаченко**,
д-р техн. наук, проф. НТУУ «КПІ» ім. Ігоря Сікорського **С.І. Отрох**

Шульга В.П.

C91 Сучасні методи передачі та обробки інформації в інфокомунікаційних мережах:
Підручник / В.П. Шульга, Л.Н. Беркман, Н.В. Галаган, К.С. Нестеренко,
О.Г. Варфоломєєва, Н.С. Хаб'юк. — К.: ФОП Білецький Р.Г., 2024. — 536 с.

ISBN 978-617-8530-18-1

В підручнику викладено загальні закономірності передавання інформації в інфокомунікаційних мережах, потєційні можливості способів передавання й приймання сигналів. Розглянуто основні напрямки побудови різних систем передавання інформації з урахуванням призначення, оцінку їх ефективності, шляхи оптимізації, удосконалення. Наведено методи обробки та оптимального прийому багатопозиційних сигналів з багатократною модуляцією та фазорізничевою модуляцією високих порядків.

Для студентів вищих навчальних закладів, аспірантів за спеціальністю «Електронні комунікації та радіотехніка».

ISBN 978-617-8530-18-1

ВСТУП

Загальні відомості про теорію інформації і передачі сигналів

Під інформацією розуміють будь-які відомості про які-небудь події, факти, явища природи, гіпотези і т.д. Передача інформації за допомогою мови, листів, газет, книг, телеграфу, телефону, радіо, кіно, телебачення, а також незліченної множини інших способів є однією з найважливіших сторін повсякденного життя людини.

Однією з основних умов ефективної життєдіяльності людини є наявність надійної належної інформації. Зовнішній світ впливає на людину через його органи почуттів, що дають людині інформацію про оточуюче його навколишнє середовище. Отримана в такий спосіб інформація перетворюється нервовою системою і мозком людини, а потім, після відповідного накопичення, перетворення і відбирання, використовується їм для активного впливу на зовнішній світ. Таким чином, взаємодія між людиною і оточуючим його зовнішнім світом відбувається на основі процесів передачі, перетворення і використання інформації для цілеспрямованих дій.

Процеси передачі, перетворення і використання інформації для цілеспрямованих дій, чи, як ми домовимося називати їх надалі, інформаційні процеси, можуть відбуватися не тільки при взаємодії людини з оточуючим його середовищем. Нагромадження досвіду (інформації) дає можливість людині створювати машини, що допомагають йому ефективно впливати на природу і накопичувати нову інформацію. Тому обмін інформацією може відбуватися між людиною і створеними ним машинами, і по суті, навіть між машинами.

Науково-технічний прогрес супроводжується інтенсивним зростанням обсягів інформації, необхідної для керування промисловістю, сільським господарством, транспортом та іншими галузями народного господарства. Про це можна судити наприклад, по збільшенню обсягу інформації, зв'язаної з роботою сучасного інформаційного обчислювального центра, пов'язаного з обробкою інформації в системі банків.

У технічних пристроях інформація завжди передається у формі сигналів. Слово «сигнал» походить від латинського «сигнум», що означає знак. Сигнали – це знаки, про зміст яких заздалегідь домовилися. Наприклад, зелене світло світлофора означає сигнал «Шлях вільний», червоне світло – сигнал «Стій!» і т.д. Вид сигналу являє собою продукт умовної угоди, і не має безпосереднього відношення до тієї інформації, що передається з його допомогою.

У всіх інформаційних процесах сигнали, що несуть інформацію, завжди втілені у вигляді будь-якої матеріальної форми. Інформація передається по телеграфу, телефону і радіо у вигляді електричних сигналів; вона записується в книгах, статтях і на магнітній стрічці; інформація міститься в показаннях приладів, формулах, таблицях і т.д. Одне і те ж саме повідомлення може передаватися за допомогою різних сигналів; наприклад за допомогою тексту, мови людини, фотографії, електричних коливань і т.д. Однак, яка б не була фізична природа сигналів, що використовуються для передачі інформації, при ідеальній передачі (за відсутності перешкод) завжди існує взаємно однозначна відповідність між переданим повідомленням і сигналом, який несе це повідомлення.

Характерною особливістю передачі інформації є те, що інформація в процесі її передачі залишається незмінною, незважаючи на те, що несучий її сигнал може перетерплювати значні зміни. Наприклад, передана по радіо інформація спочатку існує у виді тексту, який перетворюється диктором у звуки мови, що у свою чергу перетворюються в електричні коливання, передані по радіоканалу зв'язку, і відтворюються

на виході радіоприймача у вигляді звуків мови. Незважаючи на багаторазові перетворення сигналів, які несуть інформацію, в остаточному підсумку на виході радіоприймача ми одержуємо початково передану інформацію.

Таким чином передача повідомлень (а отже, і інформації) на відстань здійснюється за допомогою якого-небудь матеріального носія (паперу, магнітної стрічки і т.д.) чи фізичного процесу (звукових або електромагнітних хвиль, струму і т.д.). Фізичний процес, що відображає (несучий) передане повідомлення, називається сигналом.

В якості сигналу можна використовувати будь-який фізичний процес, що змінюється відповідно до повідомлення, яке переноситься. У сучасних системах управління і зв'язку найчастіше використовують електричні сигнали. Фізичною величиною, що визначає такий сигнал, є струм чи напруга. Сигнали формуються шляхом зміни тих чи інших параметрів фізичного носія відповідно до переданого повідомлення. Цей процес (зміни параметрів носія) прийнято називати модуляцією.

Сигнал передає (розгортає) повідомлення в часі. Отже, він завжди є функцією часу, навіть якщо повідомлення (наприклад, нерухоме зображення) таким не є. Якщо

сигнал являє собою функцію $x(t)$, що приймає тільки визначені дискретні значення x (наприклад, 1 і 0), то його називають дискретним чи дискретним за рівнем (амплітудою). Точно так само і повідомлення, що приймає тільки деякі визначені рівні, називають дискретним. Якщо ж сигнал (чи повідомлення) може приймати будь-як рівні в деякому інтервалі, то вони називаються неперервними чи аналоговими.

У деяких випадках повідомлення чи сигнал задають не на всій осі часу, а у визначені моменти t . Такі повідомлення (сигнали) називають дискретними за часом на відміну від неперервних за часом, заданих на всій осі t . Наприклад, мова є повідомленням неперервним як за рівнем, так і за часом, а датчик температури, що видає її значення через кожні 5 хв, слугує джерелом повідомлень, неперервних по величині, але дискретних за часом. На рис.В.1 наочно проілюстровані різні види сигналів. Сигнал з кінцевим числом дискретних рівнів часто називають цифровим, оскільки рівні можна пронумерувати числами з кінцевим числом розрядів.

Рішення проблем ефективності і надійності систем зв'язку в теорії інформації ґрунтується на вивченні статистичних властивостей переданих повідомлень і аналізі загальних закономірностей передачі повідомлень по каналах зв'язку при наявності перешкод. Розроблені в даний час у теорії інформації методи дозволяють істотно підвищити надійність і ефективність систем зв'язку. Особливо широке застосування ці методи знаходять у радіоелектроніці й автоматичі.

Класифікація об'єктів інформаційної техніки. За функціональною ознакою можна виділити наступні основні класи об'єктів інформаційної техніки:

- мережі та системи електрозв'язку (телеграфні, телефонні, телевізійні мережі та системи, системи радіомовлення, системи передачі даних і т.п.);
- інформаційно-вимірювальні системи й комплекси (радіолокаційні, радіонавігаційні, радіотехнічні системи автоматичного літаководіння, радіотелеметричні й т.п.);
- системи перетворення інформації (цифрові й аналогові обчислювальні машини, аналого-цифрові і цифро-аналогові перетворювачі і т.п.);

- підсистеми інформаційного забезпечення автоматизованих систем управління;
- інформаційно-пошукові системи;
- системи збереження інформації;
- мережі та системи експериментального спостереження чи дослідження (астрономічні, фізичні, медичні і т.п.).

Найбільший розвиток методи теорії інформації й передачі сигналів одержали для аналізу й синтезу систем електрозв'язку, тому що саме з задач аналізу найпростіших систем електрозв'язку почалася сама теорія інформації. Цей напрямок теорії інформації передачі сигналів називають статистичною теорією зв'язку. В даний час методи теорії інформації й передачі сигналів усе успішніше застосовують для аналізу й синтезу інших об'єктів інформаційної техніки.

Не слід думати, що дискретні повідомлення обов'язково перетворюються в дискретні сигнали, а неперервні повідомлення – у неперервні сигнали. Найчастіше саме неперервні сигнали використовують для передачі дискретних повідомлень (у якості їх переносників, несучої). Дискретні ж сигнали можуть використовуватися для передачі неперервних повідомлень (після їх дискретизації).

Повідомлення за допомогою спеціальних пристроїв (датчиків) зазвичай перетворюється в електричну величину $b(t)$ - первинний сигнал. При передачі мови таке перетворення виконує мікрофон, при передачі зображення - телевізійна камера. У більшості випадків первинний сигнал є низькочастотним коливанням, що відображає передане повідомлення. Варто звернути увагу на умовність цього терміну. Перший телевізійний сигнал, наприклад, займає область частоти від 0 до 6 МГц.

У деяких випадках первинний сигнал безпосередньо передають по лінії. Так надходять, наприклад, при звичайному міському телефонному зв'язку. Для передачі на великі відстані (по кабелю чи радіоканалу) первинний сигнал перетворюють у високочастотний.

Якби передане повідомлення було детермінованим, тобто заздалегідь відомим з повною вірогідністю, то передача його не мала б змісту. Таке детерміноване повідомлення не містить інформації. Тому повідомлення варто розглядати як випадкові події (чи випадкові величини, випадкові функції). Іншими словами, повинна існувати деяка множина варіантів повідомлення (наприклад, множина різних значень температури, видаваних датчиком), з яких реалізується з визначеною імовірністю одне. Тому і сигнал є випадковою функцією. Детермінований сигнал не може бути носієм інформації. Його можна використовувати лише для випробування системи зв'язку або окремих її елементів.

Випадковий характер повідомлень, сигналів, а також завад зумовив найважливіше значення теорії імовірностей у побудові теорії зв'язку. Як буде показано в наступних главах, імовірнісні властивості сигналів і повідомлень, а також середовища, у якому передається сигнал, дозволяють визначити кількість переданої інформації та її втрати.

Передача інформації завжди зв'язана з енергетичними процесами, тобто процесами передачі та перетворення енергії. Однак при передачі інформації об'єктом передачі є не енергія, а повідомлення. Саме ця сторона передачі інформації є загальною для живих організмів, машин і суспільства.

Теорія інформації – наука про загальні закони передачі інформації - як самостійна галузь знання виникла наприкінці 40-х років нашого сторіччя. Особливість цієї теорії насамперед полягає в тім, що в ній з єдиної точки зору розглядаються закони

передачі інформації в різних за принципом дії та фізичною природою системах (живі організми, машини і т.д.). З цієї причини теорія інформації відразу привернула до себе увагу фахівців самих різних профілів: радіотехніків, математиків, біологів, лінгвістів, соціологів і фахівців ряду інших областей науки й техніки.

Зародження і розвиток теорії інформації як науки стало можливим після того, як була визначена кількісна міра інформації, що дозволила поставити рішення всіх основних проблем передачі інформації на строгу математичну основу. Першими роботами, в яких були зроблені спроби кількісного вимірювання інформації, є роботи Р. Хартлі (1927 р.) і В.А. Котельнікова (1937 р.). Наукові основи теорії інформації були закладені в роботах К. Шеннона, опублікованих у 1948 р., і розвинуті в працях Д.А. Агеева, А.Я. Хінчіна, В.А. Котельнікова, Н.А. Колмогорова, С. Гольдмана, Д. Міддлтона, Л. Бріллю-ена, Н. Вінера та ряду інших вітчизняних і закордонних учених.

Основними проблемами теорії інформації є проблеми ефективності і надійності систем передачі інформації, тобто систем зв'язку. Перша з цих проблем – проблема ефективності систем зв'язку – полягає в тому, щоб забезпечити передачу найбільшої кількості інформації по каналах зв'язку найбільш економічним шляхом, тобто при мінімальних витратах технічних засобів у системі зв'язку. Друга проблема – проблема надійності систем зв'язку – полягає в тому, щоб забезпечити найбільшу вірогідність передачі інформації в системі зв'язку при діях на неї завад, тобто забезпечити найбільшу завадостійкість системи зв'язку.

Структура системи зв'язку

Сукупність елементів, що забезпечують передачу повідомлень від джерела повідомлень, прийнято називати системою зв'язку. Узагальнена структурна схема системи зв'язку, що відображає процес передачі інформації від джерела повідомлень до приймача повідомлень, показана на рис. В.2.

До складу системи зв'язку входять перетворювач повідомлення в сигнал, лінія зв'язку, по якій відбувається передача сигналу, і перетворювач сигналу в повідомлення, що надходить у приймач повідомлень.

Реальні системи зв'язку завжди працюють в умовах впливу завад, що можуть мати як природне, так і штучне походження. Тому в систему зв'язку входять джерела завад, що на узагальненій структурній схемі системи зв'язку об'єднані в одне еквівалентне джерело завад.

Процес перетворення повідомлення в сигнал прийнято називати кодуванням.

Лінія зв'язку являє собою фізичне середовище, по якій відбувається передача (поширення) сигналу.

Процес перетворення сигналу в повідомлення на прийомній стороні системи зв'язку прийнято називати декодуванням.

Джерело повідомлень і приймач повідомлень не входять безпосередньо в систему зв'язку. Однак їх властивості істотно впливають на властивості і параметри системи, і з цієї причини джерело повідомлень і приймач повідомлень розглядаються разом із системою зв'язку.

На практиці мають місце випадки, коли по одній лінії зв'язку ведеться передача сигналів одночасно від декількох джерел повідомлень до декількох приймачів повідомлень. У таких лініях зв'язку сигнали від різних джерел повідомлень

розділяються між собою, і як би йдуть окремими каналами зв'язку. Тому такі лінії зв'язку прийнято називати багатоканальними. Прикладами багатоканальних ліній зв'язку можуть слугувати лінії телевізійного мовлення, у яких по одній лінії зв'язку передаються сигнали зображення і сигнали звукового супроводу; проводові лінії зв'язку з ущільненням, в яких по одній парі проводів одночасно передається кілька телефонних розмов, і т.д.

При дослідженні передачі сигналів по багатоканальних лініях зв'язку необхідно враховувати не тільки умови передачі сигналів по кожному з каналів, але й взаємний вплив між каналами зв'язку.

Основні характеристики сигналів

Повідомлення, передані системою зв'язку, можуть бути величинами, неперервно змінюються в часі, або мати дискретний характер. Відповідно до цього сигнали, що відображають повідомлення, можуть мати також дискретний чи неперервний характер.

Дискретні сигнали являють собою послідовність символів, що розрізняються між собою, обраних з деякої кінцевої множини елементарних символів

$A_1, \dots, A_n$
 A_i має визначену тривалість у часі τ_i секунд. Передбачається, що кожний із символів A_i має визначену тривалість у часі τ_i секунд. Найпростішим прикладом дискретних сигналів є телеграфні сигнали, що застосовуються при передачі чи тексту чисел.

Неперервні сигнали являють собою неперервні функції часу, два різних значення яких можуть відрізнитися друг від друга як завгодно мало. Прикладами неперервних сигналів можуть слугувати телефонні сигнали, сигнали зображення в телебаченні й т.д.

У залежності від характеру переданих сигналів системи зв'язку підрозділяються на дискретні системи зв'язку і неперервні системи зв'язку.

Основними параметрами сигналу, що визначають його властивості, як переносника інформації в системах зв'язку, є тривалість сигналу T_c , смуга частот, займана спектром сигналу F_c , і перевищення сигналу над завадами D_c .

Перевищення сигналу над завадами D_c є енергетичною мірою сигналу і визначається як логарифм відношення середньої потужності сигналу до середньої потужності завад на вході лінії зв'язку, тобто:

$$D_c = \log \frac{P_c}{P_n}$$

Добуток основних параметрів сигналу

$$V_c = T_c F_c D_c \quad (B.1)$$

прийнято називати обсягом сигналу. Обсяг сигналу є узагальненою характеристикою сигналу як носія інформації. На практиці іноді використовується

геометричне представлення обсягу сигналу у вигляді прямокутного паралелепіпеда в тривимірному просторі з координатами T, F і D (рис. В.2).

За аналогією із сигналами канал зв'язку, застосовуваний для передачі сигналів, прийнято характеризувати наступними основними параметрами: часом роботи каналу

зв'язку T_k , смугою частот перепуску каналу зв'язку F_k і перевищенням середньої потужності сигналу над середньою потужністю завад на виході каналу зв'язку:

$$D_k = \log \frac{P_c}{P_n}$$

Добуток основних параметрів каналу зв'язку

$$V_k = T_k F_k D_k$$

(В.2)

прийнято називати обсягом, чи ємністю каналу зв'язку.

Основною умовою узгодження каналу зв'язку із сигналом, при виконанні якого забезпечується можливість неспотвореної передачі сигналу, є умова:

$$T_k \geq T_c; \quad F_k \geq F_c \quad \text{і} \quad D_k \geq D_c.$$

(В.3)

Перетворення повідомлення в сигнал повинно бути виконано таким чином, щоб

обсяг сигналу в тривимірному просторі T, F, D був викладений в обсязі каналу в тому ж тривимірному просторі, тобто час передачі сигналу повинний збігатися із часом роботи каналу зв'язку, смуга частот сигналу повинна збігатися зі смугою частот каналу зв'язку, і перевищення сигналу над завадою на вході і виході каналу зв'язки повинні бути рівні між собою. Тільки в цьому випадку канал зв'язку може забезпечити неспотворену передачу сигналу.

Якщо обсяг сигналу менший чи рівний обсягу каналу зв'язку, то завжди можна здійснити таке перетворення сигналу, при якому будуть виконуватися умови (В.3). Таким чином, умова можливості узгодження сигналу і каналу в'язі має вигляд:

$$V_k \geq V_c.$$

(В.4)

Ця умова повинна виконуватися для всіх сигналів, переданих через даний канал зв'язку.

Відношення

$$r_k = \frac{V_k}{V_c}$$

(В.5)

прийнято називати резервом ємності каналу зв'язку. Резерв ємності каналу зв'язку характеризує можливості підвищення надійності каналу зв'язку.

Дискретні сигнали. Для опису дискретних сигналів і повідомлень у теорії інформації широко використовується поняття абстрактного алфавіту та слів в абстрактному алфавіті.

Абстрактним алфавітом називається будь-яка закінчена сукупність об'єктів, які називають літерами даного алфавіту. Літерами абстрактного алфавіту можна вважати, наприклад, літери якої-небудь мови, цифри, ієрогліфи, малюнки і т.д. При бажанні можна ввести абстрактний алфавіт, літерами якого будуть вважатися цілі слова якої-небудь мови. Важливо лише те, щоб розглянутий алфавіт був закінченим, тобто складався з закінченого числа літер. Любий абстрактний алфавіт може бути розширений шляхом включення до нього нових додаткових літер.

Словом в абстрактному алфавіті називається будь-яка закінчена упорядкована послідовність літер. Наприклад, в алфавіті $A(x, y)$, що складається з двох літер x і y словами будуть вважатися будь-які послідовності, що складаються з цих літер: $x, xy, yy, xxxy, \dots$ і т.д. В алфавіті $A(0, 1, 2, 3, 4, 5, 6, 7, 8, 9)$, що складається з десяти літер, словом буде будь-яке ціле число. Число літер у слові називається довжиною цього слова. Наприклад, слово $xyxy$ має довжину, рівну трьом, слово 13572 – довжину, рівну п'яти, і т.д.

Кожний із символів, які входять у дискретне повідомлення, можна розглядати як літеру деякого абстрактного алфавіту, а саме повідомлення – як слово чи сукупність слів у цьому алфавіті. Прикладами дискретних повідомлень можуть слугувати слова тексту, команди, закодовані числами, і т.д. Не порушуючи спільності, можна розглядати інтервали між словами як літеру вихідного алфавіту, і розглядати повідомлення у виді слів у цьому розширеному алфавіті та довжину слова як довжину повідомлення.

Дискретні сигнали аналогічно з дискретними повідомленнями також можуть розглядатися як слова в деякому абстрактному алфавіті. Абстрактний алфавіт сигналів у загальному випадку може не збігатися з алфавітом, використовуваним для представлення повідомлень.

Раніше вказувалося, що сигнали являють собою відображення повідомлень, тобто існує однозначна відповідність між словами, представленими в алфавіті повідомлень, і словами, представленими в алфавіті сигналів. Таке однозначне відображення слів може бути задане за допомогою алфавітного оператора.

Алфавітним оператором чи алфавітним відображенням називається будь-яка відповідність (функція), що зіставляє словам у тому чи іншому алфавіті слова в тому ж самому чи деякому іншому фіксованому алфавіті. Перший алфавіт називається при цьому вхідним, а другий вихідним алфавітом даного оператора. Сукупність усіх слів, на яких алфавітний оператор визначений, називається областю визначення алфавітного оператора. Алфавітний оператор може бути визначений на всіх словах чи частині слів вхідного алфавіту.

Дискретні сигнали, передані по лініях зв'язку, являють собою алфавітні відображення повідомлень, що співставляють між собою слова в алфавіті повідомлень і слова в алфавіті сигналів. Зв'язок між повідомленнями і сигналами задається відповідним алфавітним оператором. Алфавітний оператор визначає алгоритм перетворення повідомлень у сигнали, тобто сукупність правил, за якими здійснюється це перетворення.

Кодування дискретних повідомлень. Процес перетворення повідомлень у сигнали в теорії інформації прийнято називати відображенням, що кодує, або кодуванням, а алгоритм, що визначає правила, за якими здійснюється це перетворення – кодом.

Алфавітний оператор, що визначає код перетворення повідомлень у сигнали, може бути заданий у вигляді таблиці сукупності чи правил, які визначають порядок перетворення повідомлень у сигнали.

При табличному способі завдання алфавітного оператора складається таблиця, у лівій частині якої вписуються всі слова, що входять в область визначення даного алфавітного оператора, а в правій частині – слова, що виходять у результаті застосування даного алфавітного оператора до кожного слова, що знаходиться в лівій частині таблиці.

Правила, за допомогою яких визначається порядок перетворення повідомлень у сигнали, можуть бути задані у вигляді формули чи інструкції, яка визначає порядок одержання сигналу з повідомлення.

Найпростішим видом кодування є так зване кодування по літерам. При такому

кодуванні кожній літері алфавіту повідомлення $A(x_1, \dots, x_n)$ відповідає деяка закінчена послідовність літер (слово) в алфавіті сигналу $B(y_1, \dots, y_m)$, називана кодом відповідної літери. Різним літерам алфавіту повинні відповідати різні коди.

При передачі повідомлень по системах зв'язку обов'язковою вимогою, пропонованим до кодування повідомлень, є вимога оборотності кодування, що полягає в тім, щоб по коду будь-якого слова можна було б відновити спочатку закодоване слово. Тільки при виконанні цієї вимоги можливо по сигналі відновити передане повідомлення.

Для того щоб визначити умови оборотності кодування слів за літерами, уведемо попереднє поняття про початковий відрізок слова, заданого в якому-небудь алфавіті.

Умовимося називати слово p початковим відрізком слова q , якщо слово q можна представити у виді $q = pl$, де l - будь-яке слово, у тому числі і слово нульової довжини. Наприклад, слово $p = xy$ є початковим відрізком слова $q = xyux$,

тому що слово q можна представити у виді $q = pl$, де $l = ux$.

Для того, щоб кодування слів за літерами в алфавіті $A(x_1, \dots, x_n)$ було зворотним, необхідно і досить, щоб виконувалися наступні умови:

1. Коди різних літер вихідного алфавіту повинні відрізнятися між собою.
2. Код кожної з літер вихідного алфавіту не повинний збігатися ні з яким із початкових відрізків кодів інших літер цього алфавіту.

Доведемо справедливність цієї теореми. Припустимо, що дві ці умови виконуються, і слово q є кодом якого-небудь слова p у вихідному алфавіті

$A(x_1, \dots, x_n)$. Покажемо, що по коду q можна однозначно відновити вихідне слово p . У відповідності з другою умовою, тільки один початковий відрізок слова

q може збігатися з кодом якої-небудь літери вихідного алфавіту, тобто слово q можна представити у вигляді $q_1 = s_1 q_1$, де s_1 - код першої літери слова p .

Відкидаючи початковий відрізок слова q , одержимо слово q_1 , яке можна представити у вигляді $q_1 = s_2 q_2$, де s_2 - код другої літери слова p .

Відкидаючи початковий відрізок слова q_1 , одержимо слово q_2 , за яким можна відновити третю літеру слова p , і т.д. Таким чином можна однозначно відновити одну за одною всі літери вихідного слова p .

Друга умова теореми про зворотність кодування завжди виконується, якщо коди всіх літер вихідного алфавіту мають однакову довжину. У цьому випадку в коді слова

q досить виділити початковий відрізок s_1 , що має довжину, рівну довжині коду літери, і по ньому визначити першу літеру слова p . Аналогічно можуть бути знайдені й всі інші літери слова p . Кодування за літерами, при якому коди всіх літер вихідного алфавіту мають рівну довжину, називається нормальним або рівномірним кодуванням.

Кодування дозволяє здійснювати перетворення слів, представлених у будь-якому довільному алфавіті, у слова, представлені в якому-небудь раз і назавжди обраному стандартному алфавіті. Найчастіше в якості такого стандартного алфавіту вибирається двійковий алфавіт, що складається з двох літер, які зазвичай ототожнюються з цифрами 0 і 1.

Покажемо принципову можливість такого перетворення. Нехай $A(x_1, \dots, x_n)$ - довільний алфавіт, що містить n літер, а $B(y_1, \dots, y_m)$ - стандартний алфавіт, що містить m літер ($m > 1$). У цьому випадку завжди можна вибрати ціле позитивне число k так, щоб задовольнялася нерівність:

$$m^k > n. \quad (B.6)$$

Оскільки число різних слів довжиною k в алфавіті, що складається з m літер, дорівнює m^k , то нерівність (B.6) показує, що можна закодувати всі літери алфавіту

$A(x_1, \dots, x_n)$ словами довжиною k і в алфавіті $B(y_1, \dots, y_m)$ таким чином, що коди різних літер будуть відрізнятися між собою. Будь-яке таке кодування буде нормальним а, отже, таке кодування за літерами буде зворотним.

Найпростішими прикладами кодування за літерами повідомлень є коди Бодо і Морзе, що застосовуються при телеграфуванні.

Алфавіт сигналів коду Бодо складається з двох літер (0 і 1). Код кожної літери повідомлення складається з п'яти літер алфавіту сигналу, що дозволяє передати 32 різні

літери ($2^5 = 32$). Коди деяких літер російського алфавіту для цього випадку приведені в таблиці В.1.

Таблиця В.1

Літери російського алфавіту	А	Б	У	Г	Д
Код Бодо	10000	01100	01101	01010	11110

Код Бодо є рівномірним кодом. При побудові сигналів у кодї Бодо для представлення одиниці використовується послілка струму в лінію зв'язку, а для представлення нуля — відсутність послілки струму (рис. В.4, а). Можливі й інші варіанти побудови сигналів у кодї Бодо (рис. В.4, б).

У кодї Морзе алфавіт сигналу складається з чотирьох букв: точки, тире, паузи між буквами і паузи між словами. Цей код є нерівномірним, і коди букв вихідного алфавіту (наприклад, російнина) мають у ньому різну довжину. Наприклад, слово «КИЇВ» у цьому кодї має вигляд:

К И Ї В

- • - • • • • — —

При побудові сигналів для коду Морзе точці відповідає коротка послілка струму і рівна їй по тривалості пауза, тире — послілка струму потроєної тривалості і коротка пауза, інтервалу між буквами — пауза подвоєної тривалості й інтервалу між словами — пауза збільшеної в чотири рази тривалості (рис. В.4, в).

Для передачі прописних букв російського алфавіту (їх 32) необхідно передати числа від 0 до 31. Для передачі будь-якого числа, записаного в десятковій формі, потрібна передача десяти цифр — від 0 до 9. Практично для цього потрібно десять сигналів, які відповідають різним цифрам. Систему передачі дискретних повідомлень можна істотно спростити, якщо скористатися при кодуванні двійковою системою числення.

У десятковій системі основою числення є число 10. Тому будь-яке ціле число K можна представити у вигляді:

$$K = a_n 10^n + \dots + a_2 10^2 + a_1 10^1 + a_0 10^0, \quad (\text{В.7})$$

де $a_0, a_1, \dots, a_n$ - коефіцієнти, що приймають значення від 0 до 9. Так, число 256 можна записати як $2 \cdot 10^2 + 5 \cdot 10^1 + 6 \cdot 10^0$. Очевидно, в якості числення можна прийняти будь-яке ціле число t і представити число N як

$$K = a_n m^n + \dots + a_2 m^2 + a_1 m^1 + a_0 m^0, \quad (B.8)$$

де $a_0, a_1, \dots, a_n$ - коефіцієнти, що приймають значення від 0 до $t - 1$. Задаючи величиною t , можна побудувати будь-яку систему числення.

При $T = 2$ одержимо двійкову систему, у якій числа записуються за допомогою двох цифр — 0 і 1. Наприклад, число 13 у двійковій системі записується 1101, що відповідає виразу $1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0$. Арифметичні дії в двійковій системі дуже прості. Так, додавання здійснюється за наступними правилами: $0 + 0 = 0$; $0 + 1 = 1$; $1 + 0 = 1$; $1 + 1 = 10$. Розрізняють ще порозрядне додавання без переносу в старший розряд, так зване "додавання за модулем 2". Правила цього додавання наступні: $0 \otimes 0 = 0$; $0 \otimes 1 = 1$; $1 \otimes 0 = 1$; $1 \otimes 1 = 0$.

Якщо перетворити послідовність елементів повідомлення в послідовність двійкових чисел, то для передачі останніх по каналу зв'язку досить передавати всього лише два різних сигнали. Наприклад, символи 0 і 1 можуть передаватися коливаннями з різними частотами чи імпульсами струму різної полярності. Завдяки своїй простоті двійкова система числення широко застосовується при кодуванні дискретних повідомлень.

При кодуванні відбувається процес перетворення елементів повідомлення у відповідні їм числа (кодові символи). Кожному елементу повідомлення привласнюється визначена сукупність кодових символів, яка називається кодовою комбінацією. Сукупність кодових комбінацій, що відображають дискретні повідомлення, утворює код. Правило кодування може бути виражено кодовою таблицею, в якій приводяться алфавіт повідомлень, що кодуються, і відповідні їм кодові комбінації. Множина можливих кодових символів називається кодовим алфавітом, а їх кількість T — основою коду. У загальному випадку при основі коду T правила кодування K елементів повідомлення зводяться до правил запису K різних чисел у T -ій системі числення. Число розрядів n , що утворюють кодову комбінацію, називається розрядністю коду чи довжиною кодової комбінації. У залежності від системи числення, використовуваної при кодуванні, розрізняють двійкові і T -ві (недвійкові) коди.

Коди, в яких усі комбінації мають однакову довжину, називають рівномірними. Для рівномірного коду число можливих комбінацій дорівнює T^n . Прикладом такого коду є п'ятизначний код Бодо, що містить п'ять двійкових елементів ($t = 2$, $n = 5$). Число можливих кодових комбінацій дорівнює $2^5 = 32$, що вистачить для кодування всіх букв російського алфавіту. Однак цього недостатньо для передачі повідомлення, яке містить букви, цифри, різні умовні знаки (точка, кома, додавання, множення і т.п.). Тому на даний час використовується "Міжнародний код №2" (МТК-2). У коді МТК-2

використовується реєстровий принцип, відповідно до якого та сама пятиелементна кодова комбінація може використовуватися до трьох разів у залежності від положення регістра: російський, латинський, цифровий. Загальне число різних знаків при цьому дорівнює 84, що досить для кодування телеграми. Для передачі даних рекомендований семиелементний код МТК-5. Коди МТК-2 і МТК-5 є первинними (простими). Основними параметрами кодів є: основа коду T , довжина кодової комбінації n , відстань

між кодовими комбінаціями d_{ij} і вага кодової комбінації ω . Відстань d_{ij} характеризує розходження між двома кодовими комбінаціями і визначається по Лемінгу числом незбіжних у них розрядів, тобто числом одиниць у сумі двох комбінацій за модулем 2. Число ненульових елементів у кодовій комбінації визначає її вага ω . Застосування рівномірних кодів спрощує побудову автоматичних літеродрукувальних пристроїв, і не вимагає передачі розділових символів між кодовими комбінаціями.

Нерівномірні коди характерні тим, що в них кодові комбінації відрізняються друг від друга не тільки взаємним розташуванням символів, але й їхньою кількістю. Це приводить до того, що різні комбінації мають різну тривалість. Такі коди вимагають або спеціальних розділових знаків, що вказують кінець однієї й початок іншої кодової комбінації, або ж повинні будуватися так, щоб ніяка кодова комбінація не була початком іншої. Коди, які задовольняють цій умові, називаються тими, що не приводяться чи префіксними. Відмітимо, що рівномірний код також є таким, що не приводиться. Будування коду зручно представляти у виді графа (кодового дерева), у якому з кожного вузла виходить число галузей, рівне основі коду (для двійкового коду, наприклад, крок наверх означає 0, крок униз - 1).

Типовим прикладом нерівномірних кодів є код Морзе, в якому символи 0 і 1 використовуються тільки в двох сполученнях — як одиночні (1 і 0) або як потрійні (111 і 000). Сигнал, що відповідає одній одиниці, називається точкою, трьома одиницями — тире. Символ 0 використовується як знак, що відокремлює точку від тире, точку від точки і тире від тире. Сукупність 000 використовується як розділовий знак між кодовими комбінаціями.

По ознаці завадозахищеності коди поділяють на примітивні (первинні) і коригувальні. Коди, в яких усі можливі кодові комбінації використовуються для передачі інформації, називаються простими чи кодами без надмірності (примітивними). У простих рівномірних кодах перетворення одного символу комбінації в інший, наприклад 1 у 0 чи 0 у 1, приводить до появи нової дозволеної комбінації, тобто до помилки. Коригувальні коди будуються так, що для передачі повідомлення використовуються не всі кодові комбінації, а лише деяка їх частина (дозволені кодові комбінації). Тим самим створюється можливість виявлення і виправлення помилки при неправильному відтворенні деякого числа символів. Коригувальні властивості кодів досягаються введенням у кодові комбінації додаткових (надлишкових) символів.

Декодування полягає у відновленні повідомлення за прийнятими кодовими символами. Пристрої, що здійснюють кодування і декодування, називають відповідно кодером і декодером. Як правило, це логічні пристрої. На рис. В.5 зображена структурна схема системи передачі дискретних повідомлень, а на рис. В.6 пояснюється процес перетворення дискретного повідомлення в сигнал. Передане повідомлення

позначене буквою a_k , кодоване повідомлення (чи первинний цифровий сигнал) -

$d_y(t)$, його компоненти $d_l^{(i)}$ (l — номер послідовно переданого символу, i — номер позиції коду, $i = 0, m-1$). Сигнал, що надходить у лінію зв'язку позначений $u(t)$,

прийняте коливання — $z(t)$, відновлена послідовність кодових символів — $\hat{d}_y(t)$ (ii)

компоненти $\hat{d}_l^{(i)}$ і декодоване (відновлене) повідомлення — $\hat{a}_v$. Позначення прийнятих сигналів, кодових символів і відновленого повідомлення обрані іншими, ніж переданих. Цим підкреслюється та обставина, що через дію завад прийнятий сигнал відрізняється від переданого, а відновлене повідомлення може не збігатися з вихідним.

У сучасних системах передачі дискретних повідомлень прийнято розрізняти дві груп щодо самостійних пристроїв: кодеки і модеми. Кодеком називаються пристрої, що перетворюють повідомлення в код (кодер) і код у повідомлення (декодер), а модемом — пристрій, що перетворюють код у сигнал (модулятор) і сигнал у код (демодулятор). Канальні пристрої (смугові підсилювачі передавача і приймача, коректори і т.п.) разом з лінією зв'язку утворюють неперервний канал, а останній разом з модемом - дискретний канал.

Варто мати на увазі, що в системах радіозв'язку після передавача за допомогою передавальних антен утвориться просторово-часовий сигнал $u(t, r)$ (електромагнітна хвиля), що залежить не тільки від часу t , але й просторових координат точки спостереження $r = X, Y, Z$. Крім декартових, можна ввести інші просторові координати, наприклад, полярні координати.

Сигнал, що залежить від багатьох координат, називають полем. У місці прийому (виході радіоканалу) для аналізу надходить поле чи просторово-часовий сигнал $z(t, r) = s(t, r) + n(t, r)$. Найчастіше воно спочатку за допомогою прийомної антени перетворюється в чисто часовий сигнал $z(t)$, який надалі піддається чисто тимчасовій обробці. Питання формування й обробки просторово-тимчасових сигналів у дійсному підручнику не розглядаються, тобто будемо вважати, що пристрої перетворення часовий сигнал-поле на передачі і поле-часовий сигнал на прийомі включені всередині заданої "лінії зв'язку". Ці питання розглядаються в спеціальних курсах.

При передачі неперервного повідомлення (рис. В.6) а його спочатку перетворюють у неперервний первинний електричний сигнал $b(t)$, а потім, як правило, за допомогою модулятора формують канальний сигнал $u(t)$, який і посиляють у лінію зв'язку. Прийняте коливання $z(t)$ піддається зворотним перетворенням, у результаті яких

виділяється первинний сигнал $\hat{b}$. По ньому потім відновлюється з тією чи іншою точністю повідомлення $\hat{a}$.

Загальний принцип модуляції полягає в зміні одного чи декількох параметрів несучого коливання (переносника) $f(t, \alpha, \beta, \dots)$ відповідно до переданого повідомлення. Так якщо в якості переносника обране гармонійне коливання

$$f(t) = U \cos(\omega_0 t + \varphi)$$

, то можна створити три види модуляції: амплітудну (АМ), частотну (ЧМ) і фазову (ФМ).

Якщо переносником є періодична послідовність імпульсів

$$f(t) = U_0 \sum_{l=-\infty}^{\infty} v(t - lT - t_0)$$

, то при заданій формі імпульсів $v(t)$ можна утворити чотири основних види імпульсної модуляції: амплітудно-імпульсну (АІМ), широтно-імпульсну (ШІМ), часово-імпульсну (ЧІМ, ФІМ) і частотно-імпульсну (ЧІМ). Застосування радіоімпульсів дозволяє одержати ще два види модуляції: за частотою і за фазою високочастотного заповнення.

При дискретній (цифровій) модуляції закодоване повідомлення а, яке представляє собою послідовність кодових символів $\{b_l\}$, перетвориться в

послідовність елементів (посилок) сигналу $\{u_l(t)\}$ шляхом впливу кодових символів

на переносник $f(t)$. За допомогою модуляції один з параметрів переносника змінюється за законом, що визначається кодом. При безпосередній передачі переносником може бути постійний струм, параметрами, що змінюються, якого є величина і напрямок струму. Зазвичай в якості переносника як і в неперервній модуляції, використовують змінний струм (гармонійне коливання). У цьому випадку можна одержати АМ, ЧМ і ФМ.

На рис. 31.7 приведені форми сигналу при двійковому коді для різних видів дискретної чи цифрової модуляції (маніпуляції). При АМ символу 1 відповідає передача несучого коливання протягом часу T (посилка), символу 0 — відсутність коливання (пауза). При ЧМ передача несучого коливання з частотою f_1 відповідає

символу 1, а передача коливання з частотою f_0 відповідає 0. При двійковій ФМ змінюється фаза несучої на π при кожному переході від 1 до 0 і від 0 до 1.

Нарешті, на практиці застосовують систему відносної фазової модуляції (ВФМ), На відміну від ФМ, при ВФМ фазу сигналів відраховують не від деякого еталона, а від фази попереднього елемента сигналу. Наприклад, символ 0 передається відрізком синусоїди з початковою фазою попереднього елемента сигналу, а символ 1 — таким же відрізком з початковою фазою, що відрізняється від початкової фази попереднього елемента сигналу на π . При ВФМ передача починається з посилки однієї не несучого інформації елемента, який служить опорним сигналом для порівняння фази наступного елемента. Докладніше про прийом таких сигналів і особливості відносного методу модуляції буде сказано в гл. 5.

У більш загальному випадку дискретну модуляцію варто розглядати як

перетворення кодових символів 0, 1, ..., $m - 1$ у визначені відрізки сигналу $u_i(t)$, де i

= 0, 1, ..., $m - 1$ — переданий символ. При цьому вид сигналу $u_i(t)$, у принципі, може бути довільний. У дійсності його вибирають так, щоб задовольнити вимоги, висунуті

до системи зв'язку (зокрема, за швидкістю передачі і за займаною смугою частот), і щоб сигнали добре розрізнялися з урахуванням діянь завад.

Тривалість посилки первинного сигналу $b_u(t)$ при дискретній передачі визначає швидкість передачі посилок (технічну швидкість чи швидкість модуляції). Ця швидкість ν виражається числом посилок, переданих за одиницю часу. Вимірюється технічна швидкість у Бодах. Один Бод — це швидкість, при якій за 1с передається одна посилка. Якщо тривалість посилки T виражена в секундах, то швидкість модуляції $\nu = 1/T$ в Бодах. Якщо смугу частот обмежити третьою гармонікою, то ширина спектра первинного сигналу $F = 1.5\nu$, Гц.

Демодуляція і декодування

Передане повідомлення в приймачі зазвичай відновлюється в такій послідовності. Спочатку сигнал демодулюється. У системах передачі неперервних повідомлень у результаті демодуляції відновлюється первинний сигнал, що відображає передане повідомлення. Цей сигнал потім надходить на відтворюючий чи записуючий пристрій. У радіомовленні таким пристроєм може бути гучномовець чи магнітофон. У системах передачі дискретних повідомлень, зазвичай, в результаті демодуляції послідовність елементів сигналу перетворюється в послідовність кодових символів. Потім за ними відновлюються повідомлення, що видаються отримувачу. Останнє перетворення називається декодуванням.

Не слід думати, що демодуляція і декодування — це прості операції, зворотні модуляції і кодуванню, виконувани над сигналом, який прийшов з каналу. У результаті різних спотворень і впливу завад сигнал, що прийшов, може істотно відрізнятись від переданого. Тому завжди можна висловити ряд припущень (гіпотез) про те, яке повідомлення передавалося. Задачею прийомного пристрою є ухвалення рішення про те, яке з можливих повідомлень дійсно передавалося джерелом. Для цього прийнятий сигнал піддається аналізу з урахуванням усіх відомостей про джерело (наприклад, про імовірності, з якими джерело посилає те чи інше повідомлення), про застосовуваний код і метод модуляції, а також про властивості каналу. У результаті аналізу, як правило, можна визначити умовні (апостеріорні) імовірності можливих гіпотез, і на підставі цих імовірностей прийняти рішення, яке й надходить до одержувача. Та частина прийомного пристрою, що здійснює аналіз сигналу, що приходить, і приймає рішення про передане повідомлення, називається вирішальною схемою.

У системах передачі неперервних повідомлень при аналоговій модуляції вирішальна схема визначає по спотвореному каналному (вторинному) сигналу, о прийшов, найбільш ймовірний переданий первинний сигнал і відновлює його. Тут вирішальною схемою є демодулятор. У системах передачі дискретних повідомлень вирішальна схема найчастіше складається з двох частин: першої вирішальної схеми — демодулятора і другої вирішальної схеми — декодера.

Іноді при передачі дискретних повідомлень операції демодуляції і декодування виконує один пристрій, що послідовність елементів сигналу, яка прийшла, перетворює відразу в послідовність символів (букв) повідомлення. Такий метод прийому називають спільною демодуляцією-декодуванням чи прийомом в цілому, на відміну від поелементного прийому з двома вирішальними схемами. У першому випадку аналізується цілком відрізок сигналу, що відповідає кодової комбінації, і на підставі

того чи іншого критерію відновлюється переданий елемент повідомлення (буква). В другому випадку спочатку аналізуються окремі елементи сигналу, що відповідають кодовим символам, а потім відновлена кодова комбінація декодується, тобто перетворюється в елемент (букву) повідомлення.

У деяких випадках роль вирішальної схеми виконує повністю чи частково людина. Так при прийомі телеграфних сигналів на слух оператор вирішує, який сигнал ("точка" чи "тире") був переданий. Він же виконує й операцію декодування. У приймачах дискретних повідомлень, призначених для запису інформації, усі зазначені операції виконуються автоматично. У найпростішому випадку перша вирішальна схема являє собою граничний пристрій у формі реле, тригера, що працюють за принципом "так" чи "ні". Якщо прийнятий елемент сигналу вище порога, видається один символ коду (наприклад, 1), якщо нижче - іншої (0). У деяких випадках застосовують вирішальні схеми з двома порогами. При влученні рівня сигналу між двома порогами рішення не приймається — замість сумнівного елемента сигналу видається спеціальний символ стирання. Уведення такого символу, що стирає, полегшує можливість правильного декодування прийнятої кодової комбінації.

Для ухвалення рішення про те, яке повідомлення передавалося, необхідно проаналізувати сигнал, що прийшов. Для цього він піддається різним перетворенням, яке називають обробкою сигналу. Однією з задач теорії зв'язку є відшукування правил оптимальної обробки сигналу, при яких рішення про передане повідомлення виявляється найбільш достовірним. Ці правила залежать від властивостей каналу і методів передачі (кодування і модуляції). Іноді оптимальні правила обробки виявляються складними, і для спрощення апаратури використовують іншу, не оптимальну обробку. Нарешті, якість прийому й обробки сигналів істотно залежить від точності синхронізації переданих і прийнятих сигналів. Розрізняють синхронізацію тактову (визначення границь одиничних елементів сигналу), циклову синхронізацію (правильний поділ кодових комбінацій), синхронізацію несучих частот і ін. Похибки синхронізації приводять до зниження вірогідності прийому повідомлень, а в ряді випадків — до неправильного прийому всього повідомлення чи його частини.

Найпростішим методом, що дозволяє на прийомі відокремити одну кодову комбінацію від іншої, є стартопний режим передачі, коли на початку і кінці кожної комбінації передається спеціальний сигнал ("старт" і "стоп"). Такий метод передачі відноситься до асинхронних, тому що передачу будь-якої кодової комбінації можна починати в будь-який момент часу після закінчення попередньої комбінації. При синхронних способах передачі елементи сигналу передаються неперервно через однакові проміжки часу. Поділ кодових комбінацій здійснюється в цьому випадку за допомогою циклової синхронізації. Питання синхронізації не розглядаються в даному підручнику. Вони вивчаються в спеціальних курсах.

Цифрове кодування неперервних повідомлень

На даний час широке застосування знаходять цифрові системи передачі (ДСП), в яких неперервні повідомлення передаються дискретними сигналами. Перетворення неперервного повідомлення в цифрову форму здійснюється шляхом операцій дискретизації і квантування.

Дискретизація за часом виконується шляхом взяття відліків первинного сигналу $b(t)$ у визначені дискретні моменти t . У результаті неперервну функцію $b(t)$ заміняють сукупністю миттєвих значень (відліків) $\{b(k)\}$ чи $\{b(t_k)\}$. Відліків

Операція квантування зводиться до того, що замість даного миттєвого значення $b(t_k)$ (рівня) переданого повідомлення передають найближчі значення по встановленій цифровій шкалі дискретних рівнів $b_{KB}(t_k)$. Дискретні значення по шкалі рівнів найчастіше вибираються рівномірно: $\{b_{KB}^{(l)} = l\Delta b\}$, де Δb - крок квантування, $l = 0, 1, \dots, L-1$. Саме собою зрозуміло, що при квантуванні вноситься похибка, тому що справжнє значення $b(t_k)$ заміняють округленим значенням $b_{KB}(t_k)$. Величина цієї похибки $\varepsilon = b(t_k) - b_{KB}(t_k)$ не перевершує половини кроку квантування Δb , і може бути зведена до припустимого рівня. Похибка ε є випадковою функцією і проявляється на виході в якості додаткового шуму (шуму квантування), накладеного на передане повідомлення. Дискретизація за часом дозволяє перетворити неперервні повідомлення в дискретний (у часі) сигнал, який після квантування перетворюється в цифровий. Достоїнством цифрових способів передачі є можливість застосування кодів як для підвищення завадостійкості, так і для скорочення надмірності джерела. На даний час найбільше застосування знаходить система з імпульсно-кодовою модуляцією (ІКМ). В цій системі неперервне повідомлення спочатку підлягають дискретизації за часом і квантуванню за рівнем, а потім отримана послідовність L рівнів (цифр) кодується (як правило, двійковим кодом) При цьому кожному рівню привласнюється кодова комбінація, що складається із n символів 1 і 0. Отримана послідовність двійкових символів передається по каналу зв'язку одним з методів дискретної модуляції. Звичайно використовується частотна (ІКМ-ЧМ) чи фазова (ІКМ-ФМ) модуляція.

Завади та спотворення в каналі

У реальному каналі сигнал при передачі спотворюється, і повідомлення відтворюється з деякою помилкою. Причиною таких помилок є як спотворення, внесені самим каналом, так і завади, що впливають на сигнал. Частотні й тимчасові характеристики каналу визначають так звані лінійні спотворення. Крім того, канал може вносити і нелінійні спотворення, зумовлені нелінійністю тих чи інших ланок каналу. Якщо лінійні і нелінійні спотворення зумовлені відомими характеристиками каналу, то вони, принаймні в принципі, можуть бути усунуті належною корекцією. Варто відрізнити спотворення від завад, що мають випадковий характер. Завади заздалегідь не відомі, і тому не можуть бути цілком усунуті.

Завадою називається будь-який випадковий вплив на сигнал, що погіршує вірність відтворення переданих повідомлень. Завади дуже різноманітні як за своїм походженням, так і за фізичними властивостями. У радіоканалах часто зустрічаються атмосферні завади, зумовлені електричними процесами в атмосфері, і насамперед грозовими розрядами. Енергія цих завад зосереджена головним чином в області довгих і середніх хвиль. Сильні завади створюються також промисловими установками. Це так називані індустриальні завади, що виникають через різкі зміни струму в електричних ланцюгах усіляких електроприладів. Сюди відносяться завади від електротранспорту, електричних двигунів, медичних установок, систем запалювання двигунів і т.п. Розповсюдженим видом завад є завади від сторонніх радіостанцій і каналів. Вони зумовлені порушенням регламенту розподілу робочих частот, недостатньою стабільністю частот і поганою фільтрацією гармонік сигналу, а також нелінійними процесами в каналах, що ведуть до перехресних спотворень.

У провідних каналах зв'язку основним видом завад є імпульсні шуми та переривання зв'язку. Поява імпульсних завад часто зв'язана з автоматичною комутацією і перехресними наведеннями. Переривання зв'язку - це явище, при якому сигнал у лінії різко чи загасає зникає.

Практично в будь-якому діапазоні частот мають місце внутрішні шуми апаратури, зумовлені хаотичним рухом носіїв заряду в підсилювальних приладах, резисторах й інших елементах апаратури. Ці завади особливо позначаються при радіозв'язку в діапазоні ультракоротких хвиль, де інші завади невеликі. У цьому діапазоні мають значення і космічні завади, зв'язані з електромагнітними процесами, що відбуваються на Сонці, зірках й інших неземних об'єктах. У загальному вигляді

вплив завади $n(t)$ на корисний сигнал $u(t)$ можна виразити через оператор:

$$z(t) = L[s(u(t)), n(t)]. \quad (B.9)$$

В окремому випадку, коли оператор вироджується в суму:

$$z(t) = s(t) + n(t), \quad (B.10)$$

завада називається адитивною. Якщо ж оператор може бути представлений у виді добутку:

$$z(t) = k(t)u(t), \quad (B.11)$$

то заваду називають мультиплікативною. Тут $k(t)$ - випадковий процес. У реальних каналах як правило, мають місце й адитивні, і мультиплікативні завади, і тому

$$z(t) = k(t)u(t) + n(t). \quad (B.12)$$

Серед адитивних завад різного походження виділяють зосереджені по спектру (вузькосмугові) завади, зосереджені в часі (імпульсні) завади і так називану флуктуаційну заваду, не обмежену в часі і спектрі. Флуктуаційна завада (флуктуаційний шум) являє собою випадковий процес з нормальним розподілом (гаусовський процес). Така завада найбільш вивчена і становить найбільший інтерес як у теоретичному, так і в практичному відношенні. Цей вид завад практично має місце у всіх реальних каналах. У діапазоні оптичних частот істотне значення має квантовий шум, викликаний дискретною природою сигналу. Мультиплікативні завади зумовлені

випадковими змінами параметрів каналу зв'язку. Зокрема, ці завади виявляються в зміні рівня сигналу.

Варто відмітити, що між сигналом і завадою відсутнє принципове розходження. Більш того, вони існують у єдності, хоча і протилежні за своєю дією. Так випромінювання радіопередавача є корисним сигналом для приймача, якому призначене це випромінювання і завадою для всіх інших приймачів. Електромагнітне випромінювання зірок є однією з причин космічного шуму в діапазоні надвисоких частот і, тому є завадою для систем радіозв'язку. З іншого боку, це випромінювання є корисним сигналом, за яким визначають деякі фізико-хімічні властивості зірок.

Основні характеристики системи зв'язку

При оцінці роботи системи зв'язку необхідно насамперед врахувати, яку точність передачі повідомлення забезпечує система і з якою швидкістю передається інформація. Перше визначає якість передачі, друге - кількість. У реальній системі зв'язку якість передачі залежить від ступеня спотворень прийнятого повідомлення. Ці спотворення залежать від властивостей і технічного стану системи, а також від інтенсивності і характеру завад. У правильно спроектованій і технічно справній системі зв'язку незворотні спотворення повідомлень зумовлені лише впливом завад. У цьому випадку якість передачі цілком визначається завадостійкістю системи. Під завадостійкістю звичайно розуміють здатність системи протистояти шкідливому впливу завад на передачу повідомлень. Оскільки дія завад виявляється в тому, що прийняте повідомлення відрізняється від переданого, то кількісно завадостійкість при заданій заваді можна охарактеризувати ступенем відповідності прийнятого повідомлення переданому. Назвемо цю величину загальним терміном - вірність. Кількісну міру вірності доводиться вибирати по-різному, у залежності від характеру повідомлення і вимог одержувача. Нехай повідомлення являє собою дискретну послідовність елементів з деякої закінченої множини. Вплив завади на передачу такого повідомлення виявляється в тому, що замість фактично переданого елемента може бути прийнятий який-небудь інший, така подія називається помилкою. Як кількісну міру вірності можна взяти імовірність помилки P або будь-яку монотонну функцію цієї імовірності.

При передачі неперервних повідомлень ступенем відповідності прийнятого повідомлення $\hat{B}(t)$ (з реалізацією $\hat{b}(t)$) переданому $B(t)$ (з реалізацією $b(t)$) може служити деяка величина, що представляє собою "відстань" між $b(t)$ і $\hat{b}(t)$. Часто приймають критерій квадратичного відхилення:

$$E^2(t) = [\hat{B}(t) - B(t)]^2, \quad (B.13)$$

де пряма риска зверху означає усереднення за ансамблем випадкових величин.

Кількісну міру вірності можна також визначити як імовірність того, що відхилення E

не перевершить деякої заздалегідь заданої величини ε_0 :

$$Q = P[|E| \leq \varepsilon_0] \quad (\text{В.14})$$

Як буде показано в наступних главах, вірність передачі залежить від відношення середніх потужностей сигналу і завади (ОСП). При даній інтенсивності завади імовірність помилки тим менша, чим сильніше розрізняються між собою сигнали, які відповідають різним повідомленням. Завдання полягає в тому, щоб вибрати для передачі сигнали з великим розходженням. Нарешті, вірність передачі залежить і від способу прийому. Необхідно вибрати такий спосіб прийому, який щонайкраще реалізує розходження між сигналами при даному відношенні сигналу до завади. Звернемо увагу на істотне розходження між аналоговими і дискретними системами передачі повідомлень. В аналогових системах усякий, навіть як завгодно малий заважаючий вплив на сигнал, що викликає спотворення параметра, який модулюється, завжди спричинює внесення відповідної погіршеності в повідомлення. Тому абсолютно точне відновлення переданого повідомлення неможливе. У дискретних системах помилка при передачі повідомлень виникає тільки тоді, коли сигнал орієнтується неправильно, а це відбувається лише при спотвореннях, що перевищують деякий поріг.

У теорії завадостійкості, розробленої В.А. Котельниковим, показано, що при обраному критерії та заданій множини сигналів, прийнятих при адитивному білому гаусовському шумі (БГШ), існує гранична (потенційна) завадостійкість, що ні при якому способі прийому не може бути перевищена. Приймний пристрій, що реалізує потенційну завадостійкість, називається оптимальним (найкращим) за даним критерієм.

Поряд з вірністю найважливішим показником роботи системи зв'язку є швидкість передачі. У системах передачі дискретних повідомлень швидкість вимірюється числом переданих символів за одиницю часу V , Бод. Кількість переданої інформації прийнято вимірювати в бітах (двійкових одиницях). Як буде показано в гл. 6, максимальна кількість інформації, яку можна передати двійковим символом, дорівнює 1 біту. Там же буде показано, що при використанні не двійкових, а m -кових

символів максимальна кількість інформації, яку можна передати, дорівнює $\log_2 m$ біт. Тому дискретне джерело може забезпечити максимальну швидкість видачі інформації (максимальну продуктивність):

$$R_u = \frac{\log_2 m}{T} \text{ бум/с}, \quad (\text{В.15})$$

де T — тривалість посліжки; m — основа коду. При $m = 2$, $R_u = 1/T$ біт/с, швидкість передачі інформації R_u чисельно дорівнює технічній швидкості V . При $m > 2$ можливо, що швидкість передачі інформації $R_u > V$. Однак нерідко в дискретних системах зв'язку швидкість передачі інформації $R_u < V$. Це буває, коли не всі посліжки використовуються для передачі інформації, наприклад якщо частина з них служить для синхронізації чи для виявлення і виправлення помилок (при використанні коригувального коду).

Вводиться також характеристика середньої швидкості передачі інформації з заданого каналу в одиницю часу з заданою вірністю. Існує максимально можлива

(гранична) швидкість передачі, що називається пропускнуою здатністю каналу C' . Це фундаментальне поняття визначає потенційні можливості системи зв'язку, що використовує даний канал. У реальній системі середня швидкість передачі інформації завжди менша пропускнуої здатності каналу C' . У теорії інформації К.Шенноном доведена теорема, відповідно до якої для джерела без надмірності (див. гл. 6) при $R_{\text{и}} < C'$ можна знайти такий спосіб кодування-декодування, при якому можлива передача повідомлень по каналі з завадами з як завгодно малою помилкою.

Універсальним показником ефективності системи зв'язку є коефіцієнт η , що характеризує використання системою пропускнуої здатності каналу $\eta = R_{\text{и}} / C'$ (інформаційна ефективність).

Своєчасність передачі повідомлень визначається припустимою затримкою, зумовленою перетворенням повідомлень і сигналів, а також кінцевим часом поширення сигналу по каналу зв'язку. Вона залежить, по-перше, від характеру і довжини каналу, по-друге, від тривалості обробки сигналу в передавальному і прийомному пристроях. Швидкість передачі і затримка є незалежними характеристиками, практично не зв'язаними один з одним.

Існують і багато інших параметрів, що характеризують з різних точок зору якості системи зв'язку. До них, зокрема, відносяться скритність зв'язку, надійність системи, габаритні розміри і маса апаратури, вартість устаткування, експлуатаційні витрати і т.п. Ці характеристики в курсі "Теорія електричного зв'язку" не розглядаються. Їм присвячені окремі розділи інших спеціальних курсів.

Питання, задачі і вправи

1. Задано таблицю трьох кодів:

повідомлення № коду	а	б	у	г	д	є	ж	з
1	000	001	010	011	100	101	110	111
2	0	1	00	01	10	11	110	111
3	00	01	100	101	1100	1101	1110	1111

Необхідно: 1) визначити вид коду; 2) визначити, які з цих кодів не вимагають розділових знаків; чому? 3) закодувати слова: зірка, вега, база; 4) розшифрувати послідовності кодових комбінацій: а) коду 1: 111000001000010000; 001101100000; б) коду 3; 0100101001110; 01110110100; 10000111000.

2. Стартстопний телеграфний апарат передає одну літеру 7 послілками: однією стартовою (20 мс), п'ятьма кодовими (20 мс кожна) і однією стоповою (30 мс). Визначити швидкість модуляції в у Бодах і максимальну швидкість передачі інформації $R_{\text{и}}$ в бітах за секунду.

3. У чому відмінність між дискретною і неперервною модуляцією несучої?
4. Що розуміють під вирішальною схемою при прийомі дискретних повідомлень?
5. Практично вірність передачі дискретних повідомлень визначають коефіцієнтом помилок - відносним числом, помилково прийнятих елементів повідомлення. Обчисліть, цей коефіцієнт для випадку передачі літер кодом МТК-2 зі швидкістю 50 Бод, якщо за 2 ч неперервної передачі було зафіксовано 10 помилково прийнятих літер.
6. Запишіть кодові слова (комбінації), що відповідають числу 45, при основі коду $m = 2, 3, 4, 8, 10$. Як змінюється число розрядів і число використовуваних цифр у слові зі зміною основи коду?
7. Яке число розрядів повинне мати рівномірний код, призначений для первинного кодування 32-літереного алфавіту, при основі коду $m = 2, 8, 16, 32$?
8. За даними, приведеним у § 1.1, обчисліть і порівняйте обсяги первинних сигналів: мовного, радіомовного і телевізійного (при однаковій їхній тривалості).
9. Текст із 180 літер передається телефонним каналом протягом 30 с. Той же текст за той ж час передається телеграфом кодом МТК-2. Порівняйте обсяги телефонного і телеграфного сигналів (при однакових динамічних діапазонах).
10. Покажіть, як за рахунок кодування можна здійснити обмін смуги частот і відношення сигнал-завада в каналі при незмінному обсязі сигналу.

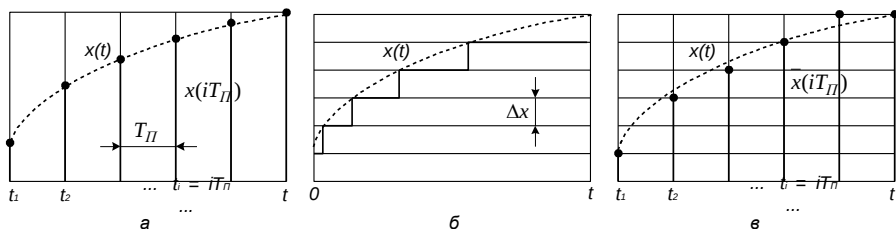


Рис.13.1. Види сигналів: а - квантування за часом; б - квантування за рівнем; в - квантування за часом та рівнем

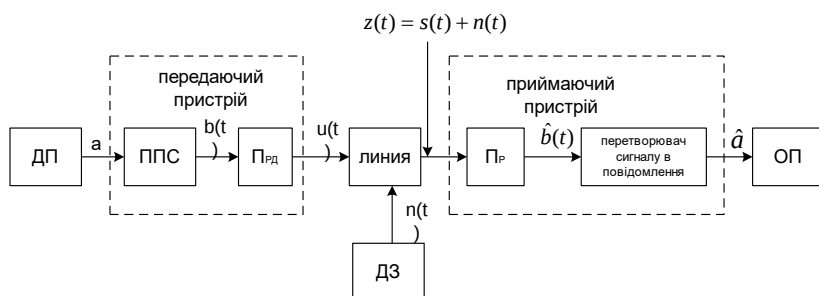


Рис.13.2. Структурна схема одноканальної системи зв'язку. ДП - джерело повідомлення; ППС - перетворювач повідомлення в сигнал; Прд - передавач; Пр - приймач; ОП - отримувач повідомлення; ДЗ - джерело завад.

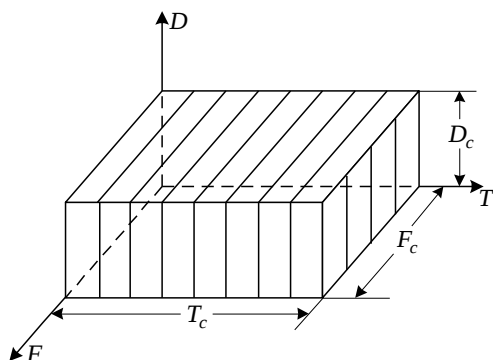


Рис.13.3. Геометричне представлення обсягу сигналу

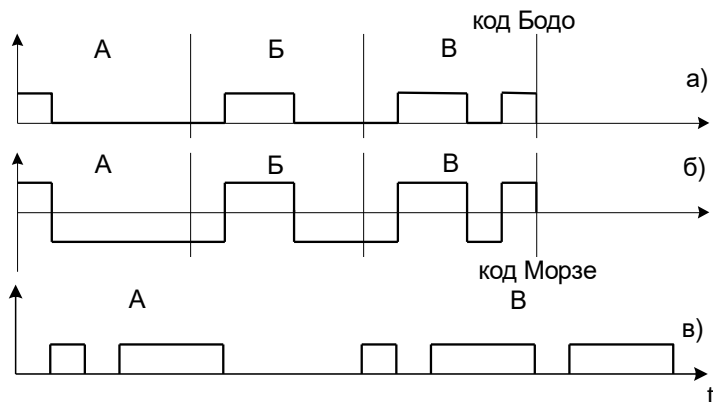


Рис.13.4. Варіанти побудови кода Бодо (а, б) та Кода Морзе (в)

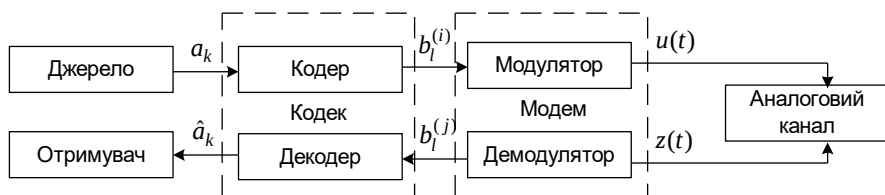


Рис.13.5. Структурна схема системи передачі дискретних повідомлень

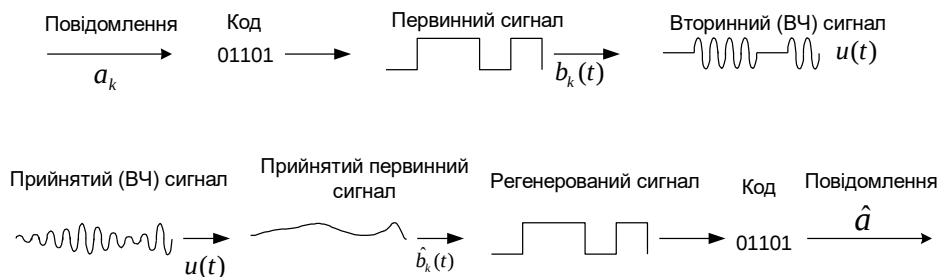


Рис.13.7. Форми сигналів при двійковому коді для різних видів дискретної модуляції

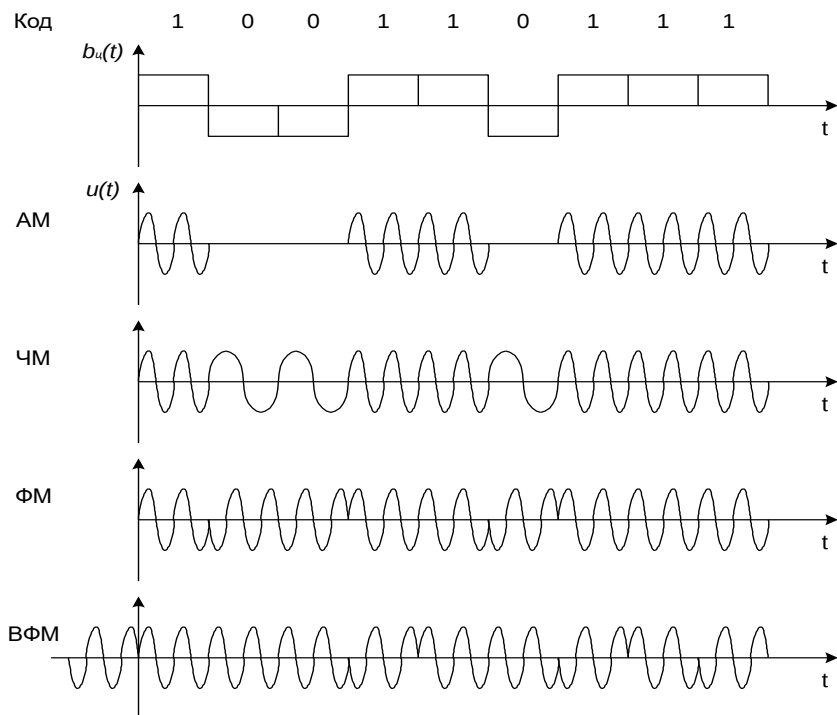


Рис.13.7. Форми сигналів при двійковому коді для різних видів дискретної модуляції

РОЗДІЛ 1. МАТЕМАТИЧНІ МОДЕЛІ ДЕТЕРМІНОВАНИХ СИГНАЛІВ

1.1 Детерміновані сигнали

З точки зору вирішення задач електричного зв'язку математичні моделі, що відображають основні властивості повідомлень і завад, є фундаментом теорії електрозв'язку. Оскільки сигнали і завади перш за все представляються електричними комбінаціями, що змінюються в часі, то їх базовою математичною моделлю являється деяка функція часу $S(t)$. За своїми фізичними та математичними властивостями процеси $x(t)$ діляться на детерміновані та випадкові. Нижче в даному розділі розглядаються детерміновані сигнали.

Детермінованими чи регулярними називають такі процеси $S(t)$, протікання яких в часі можна повністю визначити наперед ($S(t) = at$, $S(t) = at^2$, $S(t) = A\cos(\omega t + \varphi)$). Інакше кажучи, для будь якого наперед заданого моменту часу t може бути однозначно визначене значення функції $S(t)$. Розрізняють наступні види детермінованих сигналів: прості; складні; еталонні чи пробні, періодичні і неперіодичні.

Простим сигналом називають сигнал, що відповідає одній елементарній посліці. Наприклад, при передачі дискретної інформації простому сигналу відповідає сигнал кодового символу в комбінації (рис.1.1,а).

Сигнал, що представляє собою сукупність елементарних посилок, називають складеним чи складним (рис.1.1, б).

В теорії зв'язку вводиться поняття бази сигналу V :

$$V = 2TF,$$

де T - тривалість сигналу, F - смуга частот, що займається сигналом.

Для простих сигналів $V \approx 1$, а для складних $V \gg 1$. Враховуючи цю обставину, прості сигнали називають вузькосмуговими, а складні – широкосмуговими.

Еталонні або пробні сигнали:

1. Гармонічний сигнал (рис.1.2, а):

$$S(t) = A\sin(\omega t + \varphi), \quad (-\infty < t < \infty),$$

де A, ω, φ - амплітуда, частота і фаза сигналу відповідно.

2. Одинична функція (функція вмикання (рис.1.2, б)):

$$1(t) = \begin{cases} 0, & \text{нпу } t < 0; \\ f, & \text{нпу } t > 0. \end{cases}$$

3. Одиничний імпульс (дельта функція (рис.1.2, в)):

$$\delta(t) = \begin{cases} 0, & \text{нпу } t < 0; \\ \infty, & \text{нпу } t = 0; \\ 0, & \text{нпу } t > 0. \end{cases}$$

До основних типів детермінованих сигналів відносяться періодичні та неперіодичні.

1.1.1 Періодичні сигнали

Періодичним сигналом називається сигнал, який повторюється через визначені проміжки часу:

$$S(t) = S(t + nT), \quad (-\infty < t < \infty),$$

де T - період повтору, n - будь-яке число.

Будь-яка періодична функція, що задовольняє умовам Дирихле, може бути представлена у вигляді ряду Фур'є. Умови визначаються наступним чином. Функція $S(t)$ повинна бути скрізь однозначною, кінцевою, кусково-неперервною (тобто інтервал, на якому функція визначена, може бути розбитий на кінцеве число інтервалів, в кожному із яких $S(t)$ неперервна і монотонна), повинна мати обмежене число максимумів і мінімумів, у будь-якій точці розриву існує $S(t-0)$ і $S(t+0)$;

значення $S(t)$ в точці розриву рівне $\frac{S(t-0) + S(t+0)}{2}$. При цьому ряд Фур'є

записується наступним чином:

$$S(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} A_k \cos(k\omega_1 t - \varphi_k)$$

або

$$S(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos k\omega_1 t + b_k \sin k\omega_1 t$$

де

$$a_0 = \frac{2}{T} \int_{-T/2}^{T/2} S(t) dt, \quad a_k = A_k \cos \varphi_k, \quad b_k = A_k \sin \varphi_k, \\ A_k = \sqrt{a_k^2 + b_k^2}, \quad \varphi_k = \arctg b_k / a_k.$$

Комплексна форма запису ряду Фур'є:

$$S(t) = \frac{1}{2} \sum_{k=-\infty}^{\infty} \dot{A}_k e^{jk\omega_1 t},$$

де

$$\dot{A}_k = A_k e^{-j\varphi_k} = a_k + jb_k, \quad A_k = |\dot{A}_k|, \\ \dot{A}_k = \frac{2}{T} \int_{-T/2}^{T/2} S(t) e^{-jk\omega_1 t} dt.$$

Якщо $S(t) = S(-t)$ (функція парна), то

$$b_k = 0, \quad A_k = a_k,$$

якщо $S(t) = -S(t)$ (функція непарна), то

$$a_k = 0, \quad A_k = b_k.$$

Спектр періодичного сигналу зображається графічно у вигляді спектральних ліній, довжини яких пропорційні амплітудам (фазам) відповідних частотних складових.

Спектр періодичного сигналу – дискретний (лінійчатий). При цьому відстань між сусідніми лініями постійне і рівне частоті першої гармоніки (рис.1.5).

Неперервна крива, що з'єднує кінці ліній спектру, називається огибаючою спектра.

$$S(j\omega) = \frac{2}{T} \int_{-T/2}^{T/2} S(t) e^{-j\omega t} dt = S(t) e^{i\varphi(\omega)},$$

де $S(\omega), \varphi(\omega)$ - огибаючі спектрів амплітуд і фаз відповідно.

Середня потужність сигналу, що виділяється на навантаженні з опором 1 Ом, рівна:

$$P_{cp} = \overline{S^2(t)} = \frac{1}{T} \int_{-T/2}^{T/2} S^2(t) dt = \frac{a_0^2}{2} + \frac{1}{2} \sum_{k=1}^{\infty} A_k^2.$$

Отримана формула для P_{cp} називається рівністю Парсеваля, згідно якій середня потужність періодичного сигналу рівна сумі середніх потужностей всіх частотних складових його спектра.

Реальні канали зв'язку, через які проходить сигнал, мають обмежену смугу перепуску, тому сигнал $S(t)$ можна представити у вигляді:

$$S(t) \approx \sum_{k=l_2}^{l_1} (a_k \cos k\omega_1 t + b_k \sin k\omega_1 t). \quad (1.1)$$

Число членів ряду (1.1) рівне n , де $n = 2TF$, $F = (e_2 - e_1) / \tau$ - смуга частот, що займається каналом. Апроксимація сигналу $S(t)$ рядом (1.1) дає найменше середньоквадратичне відхилення від точного значення $S(t)$, якщо a_k і b_k являються коефіцієнтами ряду Фур'є.

1.1.2 Неперіодичні сигнали

Неперіодичними називають сигнали, які задовольняють умові $S(t) \neq S(t + nT)$ на інтервалі $(-\infty < t < \infty)$. Такий сигнал представляється функцією на кінцевому чи напівбезкінечному інтервалах:

$$t_1 < t < t_2, \quad t_1 < t < \infty.$$

Для неперіодичного сигналу існує пряме та зворотнє перетворення Фур'є:

$$S(j\omega) = \int_{-\infty}^{\infty} S(t)e^{-j\omega t} dt$$

$$S(t) = \frac{1}{2T} \int_{-\infty}^{\infty} S(j\omega)e^{j\omega t} d\omega.$$

Величина $S(j\omega)$ виражає спектральну щільність сигналу і називається спектральною характеристикою або комплексним спектром, а модуль $S(j\omega)$ - спектром сигналу.

$$S(j\omega) = \pi \frac{dA}{d\omega}; \quad |S(j\omega)| = S(\omega) \quad - \text{спектр сигналу. Спектральну}$$

характеристику як комплексну величину можна представити у вигляді:

$$S(j\omega) = A(\omega) + jB(\omega) = S(\omega)e^{-j\varphi(\omega)},$$

де

$$A(\omega) = \int_{-\infty}^{\infty} S(t) \cos \omega t dt, \quad B(\omega) = \int_{-\infty}^{\infty} S(t) \sin \omega t dt$$

$$S(\omega) = \sqrt{A^2(\omega) + B^2(\omega)}, \quad \varphi(\omega) = \arctg \frac{B(\omega)}{A(\omega)}.$$

Енергія неперіодичного сигналу рівна:

$$E = \int_{-\infty}^{\infty} S^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega)S(-j\omega) d\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} [S(\omega)]^2 d\omega.$$

Якщо спектр обмежений смугою частот $F = f_2 - f_1$, можна наближено записати:

$$\frac{1}{\pi} \int_{f_1}^{f_2} [S(\omega)]^2 d\omega \approx \frac{1}{\pi} \int_{-\infty}^{\infty} [S(\omega)]^2 d\omega.$$

(1.2)

Практично ширину спектра сигналу необхідно вибирати так, щоб забезпечити виконання цієї рівності з заданою точністю. Це однак не являється єдиним критерієм вибору ширини спектру. Другим критерієм є допустимий ступінь спотворення форми сигналу.

Практично під шириною спектра розуміється область частот, у межах якої зосереджено 90% або 99% енергії сигналу.

Добуток ширини спектра імпульсного сигналу і тривалості імпульсу є величина постійна:

$$F \tau = const.$$

$F \tau = 1$ - для прямокутного імпульсу,

$F \tau = 1,5$ - для косинусоїдального імпульсу,

$F \tau = 2$ - для трикутного імпульсу.

На рис.1.10 приведені спектри деяких неперіодичних сигналів.

1.2 Часове представлення сигналів. Дискретизація інформації

Під дискретизацією розуміється перетворення неперервних повідомлень в дискретні. При цьому використовується дискретизація за часом та за рівнем. Перехід від аналогового представлення до дискретного у багатьох випадках дає значні переваги при передачі, обробці та зберіганні інформації.

Заміна неперервної функції дискретною дає можливість всі результати, отримані для дискретних повідомлень, застосувати до неперервних повідомлень.

Дискретизація за часом. Суть дискретизації за часом полягає в тому, що деякій неперервній функції ставиться у відповідність друга функція, створена шляхом переривання вихідної функції (рис.1.6.). Природно, що така заміна допустима лиш в тих випадках, коли дискретизована функція повністю представляє вихідну функцію. Інтуїція підказує, що заміна неперервної функції дискретною можлива лише при певних допущеннях. Яким повинен бути інтервал Δt між окремими відліками? При малому інтервалі між відліками їх кількість буде більшою і точність наступного відновлення функції також буде високою. Якщо ж інтервал між відліками взяти великим, то кількість відліків зменшиться, однак похибка відновлення неперервного повідомлення може виявитися більшою від допустимої. Оптимальним слід вважати такий інтервал між відліками, при якому вихідна функція із заданою точністю представляється мінімальним числом відлікових значень. У цьому випадку всі відліки

будуть істотними для відновлення вихідної функції. При більшому числі відліків буде мати місце надлишковість інформації

Спосіб дискретизації, згідно до якого вибираються відліки або коефіцієнти розкладу, доцільно оцінювати за величиною похибки відновлення вихідної функції. Розрізняють три види критеріїв відліків [1]:

1. Частотний критерій Котельникова, згідно до якого інтервали між відліками вибираються виходячи із ширини спектра повідомлення, що дискретизується [8].
2. Кореляційний критерій відліків Желєзнова, згідно до якого інтервал дискретизації вибирається рівним часу кореляції повідомлення, що передається [9].
3. Квантовий критерій відліків Темникові, запропонований для детермінованих функцій, який встановлює залежність інтервалів між відліками від величини ступеня квантування за рівнем і крутизною функції [10].

Оцінку точності відновлення неперервного повідомлення можна проводити за величиною найбільшого, середньоквадратичного або інтегрального відхилення.

Теорема В.А.Котельникова. Реальні сигнали завжди мають кінцеву тривалість і обмежену смугу частот. Граничні частоти спектра визначаються головним чином властивостями системи передачі і самим отримувачем. Так, наприклад, при передачі дискретних повідомлень смуга частот визначається швидкістю передачі, при передачі телевізійного зображення – прийнятим стандартом чіткості (числом строчок і т.п.).

Для функцій з обмеженим спектром В.А. Котельников довів чудову теорему, що лежить в основі дискретизації неперервних сигналів.

Згідно до цієї теореми функція $S(t)$, яка не містить частот вищих від F , повністю визначається послідовністю своїх значень в моменти, які відстають один від одного на $\Delta t = 1/2F = \pi/\omega$.

Фізичним підтвердженням цього є відомий факт про те, що сигнал $S(t)$ не може істотно змінитися за час, менший, ніж половина періоду його найвищої частоти, тобто $1/2F$.

Якщо неперервний сигнал обмежений за спектром, не обов'язково передавати всі миттєві значення цієї функції, а тільки окремі її значення.

Доведення теореми В.А.Котельникова. Нехай функція $S(t)$ має обмежений спектр. Тоді

$$S(j\omega) = \begin{cases} \int_{-\infty}^{\infty} S(t)e^{-j\omega t} dt & \text{при } |\omega| \leq 2\pi F; \\ 0 & \text{при } |\omega| > 2\pi F. \end{cases}$$

На кінцевому інтервалі $(-2\pi F, 2\pi F)$ функцію $S(j\omega)$ можна представити у вигляді ряду Фур'є за аргументом ω :

$$S(j\omega) = \frac{1}{2} \sum_{k=-\infty}^{\infty} A_k e^{jk \frac{\omega}{2F}}.$$

Коефіцієнти Фур'є у цьому випадку запишуться у вигляді:

$$A_k = \frac{2}{4\pi F} \int_{-2\pi F}^{2\pi F} S(j\omega) e^{-j \frac{k\omega}{2F}} d\omega.$$

Так як $S(j\omega)$ є перетворенням Фур'є для $S(t)$, то $S(t)$ також є перетворенням Фур'є для $S(j\omega)$:

$$S(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(j\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-2\pi F}^{2\pi F} S(j\omega) e^{j\omega t} d\omega.$$

Вважаючи $t = k/2F$, отримуємо:

$$S\left(\frac{k}{2F}\right) = \frac{1}{2\pi} \int_{-2\pi F}^{2\pi F} S(j\omega) e^{j \frac{k\omega}{2F}} d\omega.$$

Порівнюючи отриманий вираз із виразом для A_k бачимо, що

$$A_k = \frac{1}{F} S\left(-\frac{k}{2F}\right) = 2\Delta t f(-k\Delta t).$$

Отже

$$S(j\omega) = \Delta t \sum_{k=-\infty}^{\infty} S(k\Delta t) e^{-j \frac{k\omega}{2F}}.$$

Так як $S(t)$ однозначно визначається своїм спектром $S(j\omega)$, то із отриманого співвідношення випливає, що $S(t)$, як $S(j\omega)$, однозначно визначаються відліками $S(k\Delta t)$. Це і доводить теорему В.А.Котельникова.

Виражаючи функцію $S(t)$ через її спектр $S(j\omega)$ отримаємо:

$$S(t) = \frac{1}{2\pi} \int_{-2\pi F}^{2\pi F} S(j\omega) e^{j\omega t} d\omega = \frac{\Delta t}{2\pi} \int_{-2\pi F}^{2\pi F} e^{j\omega t} d\omega \sum_{-\infty}^{\infty} S(k\Delta t) e^{j\omega k\Delta t}$$

або, змінюючи порядок дії й інтегруючи, отримаємо:

$$S(t) = \frac{\Delta t}{2\pi} \sum_{k=-\infty}^{\infty} S(k\Delta t) \int_{-2\pi F}^{2\pi F} e^{ij\omega(t-k\Delta t)} dt = \sum_{k=-\infty}^{\infty} S(k\Delta t) \frac{\sin 2\pi F(t-k\Delta t)}{2\pi F(t-k\Delta t)}. \quad (1.3)$$

Вираз (1.3) представляє собою розклад неперервної функції $S(t)$ в ряд за ортогональними функціями виду $\frac{\sin x}{x}$. Величини $S(k\Delta t)$ називаються відліками функції $S(t)$. Вони визначають значення $S(k\Delta t)$ у дискретні моменти часу $k\Delta t$.

Множник $\frac{\sin 2\pi F(t-k\Delta t)}{2\pi F(t-k\Delta t)}$ називається функцією відліків.

Позначивши $\tau = t - k\Delta t$, отримаємо:

$$\psi(\tau) = \frac{\sin 2\pi F \tau}{2\pi F \tau}, \quad \lim_{\tau \rightarrow 0} \psi(\tau) = 1.$$

Графік функції $\psi(\tau)$ зображений на рис.1.5, а. Функція відліків приймає найбільше значення, рівне одиниці в моменти $t = k\Delta t$ ($\tau = 0$), і перетворюється в нуль у моменти часу $t = (k \pm m)\Delta t$, де $m = 1, 2, \dots$. Ширина головної пелюстки функції відліків на нульовому рівні рівна $1/F$. Спектр функції відліків є рівномірним у смузі $(-F, F)$ і рівний нулю поза цією смугою. Модуль спектра $S(\omega) = 1/(2F)$. Енергія сигналу через відлікові значення визначається наступним чином:

$$E = \int_{-\infty}^{\infty} S^2(t) dt = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} S^2(k\Delta t).$$

Якщо смуга частот сигналу $S(t)$ необмежено розширюється, то Δt буде необмежено зменшуватися, і в межах при $F \rightarrow \infty$ функція відліків прямує до дельта-функції $\delta(t - t_k)$, а ряд Котельникова перетворюється в інтеграл:

$$S(t) = \int_{-\infty}^{\infty} S(t) \delta(t - t_k) dt. \quad (1.4)$$

Згортка дельта-функції з будь-якою функцією $S(t)$ дає нерівність:

$$\int_{-\infty}^{\infty} f(t) \delta(t - t_k) dt = f(t_k) \int_{-\infty}^{\infty} \delta(t - t_k) dt = f(t_k). \quad (1.5)$$

Із порівняння виразів (1.4) і (1.5) випливає, що інтеграл, який визначається виразом (1.4) не зміниться при заміні $S(t)$ на $S(t_k)$. Тоді інтеграл (1.4) можна записати у вигляді інтегралу Дюамеля:

$$S(t) = \int_0^T S(\tau) \delta(t - \tau) d\tau$$

На рис.1.5, б показаний неперервний сигнал, що дискретизується за часом, а на рис.1.5, в – функції відліків для деяких сусідніх відліків.

Характеристики виглядом $\frac{\sin x}{x}$ має ідеальний фільтр нижніх частот

(рис.1.6, а) при передачі на вхід одиничного імпульсу. Реальні характеристики $k(\omega)$ і $\varphi(\omega)$ відрізняються від ідеальних, що приводить до відхилення форми реальної функції від ідеальної, і, як наслідок, до появи додаткових похибок відновлення функції $S(t)$.

Теорема Котельникова дозволяє єдиним чином розглядати передачу любого сигналу (дискретного чи неперервного) як передачу чисел. Ця теорема лежить в основі всіх видів імпульсної модуляції. Згідно до цієї теореми необхідна частота слідування імпульсів F_i , яку називають тактовою частотою, повинна визначатися із умови $F_i \geq 2F_m$, де F_m - верхня гранична частота спектру повідомлення, яке передається.

Використовуючи теорему Котельникова можна здійснювати часове ущільнення каналів, тобто передавати декілька каналів одночасно. На рис.1.5, 6 зображено два сигнали: $S_1(t)$ і $S_2(t)$, які передаються одночасно.

Теорема Котельникова може бути узагальнена і на випадок функцій, спектр яких не починається з нульової частоти. У формулюванні автора ця теорема проголошує: будь-яку функцію $S(t)$, що складається із частот від f_1 до f_2 можна передавати з любою точністю за допомогою чисел, які передаються один за одним через інтервали $\Delta t = \frac{1}{F}$, де $F = f_2 - f_1$. При цьому в кожній точці відраховуються дві величини: амплітуда і фаза.

Коли тривалість сигналу рівна T , то кількість відліків на осі часу рівна $T/\Delta t = TF$, а загальна кількість відліків (амплітуд і фаз) рівна $N = 2TF$.

Функцію $S(t)$, обмежену частотами f_1, f_2 , можна представити наступним рядом:

$$S(t) = \sum_{k=-\infty}^{\infty} S\left(\frac{k}{F}\right) \frac{\sin \pi F(t - \frac{k}{F})}{\pi F(t - \frac{k}{F})} \cos \left[\omega_0(t - \frac{k}{F}) + \varphi\left(\frac{k}{F}\right) \right],$$

Де $S(\frac{k}{F})$ і $\varphi(\frac{k}{F})$ - відлікові значення відповідно амплітуди і фази:

$$\omega_0 = 2\pi \frac{f_1 + f_2}{2} - \text{середнє значення кругової частоти спектру сигналу.}$$

Функція відліків в цьому випадку представляє собою модульоване коливання із несучою, рівною ω_0 і огибаючою, яка визначається функцією виду $\frac{\sin x}{x}$ (рис.1.7).

До цих пір розглядалися функції з обмеженим спектром, і питання про їх тривалість не ставилося. Однак відомо, що реальні повідомлення мають кінцеву тривалість. Із теорії перетворення Фур'є відомо, що часове представлення функції з обмеженим спектром розтягується на безкінечно великий інтервал часу. В той же час кінцева тривалість повідомлення визначає безкінечно протяжний спектр. Це свідчить про неможливість існування функції кінцевої тривалості із обмеженим спектром. Таким чином, принципіально теорема Котельникова несправедлива для реальних повідомлень. Однак, враховуючи відмічені раніше особливості реальних повідомлень (зосередженість майже всієї енергії в кінцевих інтервалах часу і смуги частот (рис.1.3.)), можна з достатнім ступенем точності використати її для представлення реальних повідомлень. В цьому випадку при тривалості повідомлення T_C число відліків m буде кінцевим:

$$m = \frac{T_C}{\Delta t} = 2F_B T_C, \quad (1.6)$$

а (1.3) прийме вигляд:

$$S(t) \approx \sum_{k=-\frac{m}{2}}^{\frac{m}{2}} S(k\Delta t) \frac{\sin \omega_B(t - k\Delta t)}{\omega_B(t - k\Delta t)} \quad (1.7)$$

При кінцевому числі відліків сума (1.7) буде співпадати з миттєвими значеннями $S(t)$ не на всьому інтервалі існування повідомлення T_C , а тільки у відлікових точках. В проміжках між цими значеннями $S(t)$ відрізняється від суми кінцевого числа членів ряду, в результаті чого виникає похибка. Ця похибка мінімальна в середині інтервалу T_C і буде зростати до його країв (рис.1.8).

Визначити середньоквадратичну похибку відновлення повідомлення кінцевої тривалості рядом Котельникова із кінцевим числом членів важко. Однак в роботі [1]

визначені границі похибки, що виникає при відновленні реального повідомлення із необмеженим спектром рядом Котельникова (1.3). Позначаючи відновлене повідомлення $S'(t)$, запишемо вираз для середньоквадратичної похибки δ [1]:

$$\sqrt{\frac{\Delta E}{E}} \leq \delta = \frac{\sqrt{\int_{-\infty}^{\infty} [S'(t) - S(t)]^2 dt}}{\sqrt{\int_{-\infty}^{\infty} S^2(t) dt}} \leq \sqrt{\frac{\Delta E}{E}} (3 + Q), \quad (1.8)$$

де E - повна енергія повідомлення, рівна $\int_{-\infty}^{\infty} S^2(t) dt$;

ΔE - енергія повідомлення, заключена в смузі частот вище, ніж $F_B = \frac{1}{2\Delta t}$;

Q - коефіцієнт, що враховує швидкість зменшення модуля спектральної щільності на частотах, більших, ніж F_B .

Згідно до [1]:

$$Q = 4 \sum_{r=1}^{\infty} \left[\frac{\int_{(2r+1)\omega_B}^{\infty} [S(\omega)]^2 d\omega}{\int_{\omega_B}^{\infty} [S(\omega)]^2 d\omega} \right]^{1/2}, \quad (1.9)$$

де $r = 1, 2, \dots$ і т. д.

Якщо спадання модуля спектральної щільності відбувається достатньо швидко, користуються більш простою оцінкою:

$$\sqrt{\frac{\Delta E}{E}} \leq \delta \leq \sqrt{\frac{3\Delta E}{E}}, \quad (1.10)$$

за допомогою якої можна визначити F_B таким чином, щоб похибка відновлення неперервного повідомлення не перевищувала заданої величини.

В практичних розрахунках для визначення відносної похибки відтворення повідомлень може бути використане наступний вираз:

$$\delta^2 \approx \frac{\omega_B}{\int_0^\infty [S(\omega)]^2 d\omega}, \quad (1.11)$$

який дозволяє за заданою величиною δ при відомих спектральних характеристиках повідомлення знайти верхню граничну частоту спектра ω_0 і інтервал дискретизації:

$$\Delta t = \frac{\pi}{\omega_B}. \quad (1.12)$$

Принцип дискретизації Желєзнова. Н.А. Желєзновим вирішена задача визначення максимального інтервалу між відліками випадкового повідомлення. У літературі, що використовується, моделі повідомлення характеризуються: кінцевою тривалістю T_C ; суцільним спектром, що відрізняється від нуля по всій частотній осі, тобто $S(\omega) \neq 0$ при $-\infty \leq \omega \leq \infty$.

Вводиться допущення обмеженості інтервалу кореляції (тобто вважається, що функція кореляції рівна нулю поза інтегралом τ_0) і малому розміру інтервалу кореляції в порівнянні з тривалістю повідомлення ($T_C \gg \tau_0$). Розглядаються повідомлення, що є стаціонарними й нестаціонарними функціями часу. Введені допущення не вступають у протиріччя з природою реальних повідомлень, так як

внаслідок кінцевої тривалості їх значення в будь-який момент залежить тільки від деякого відрізка минулого з обмеженою тривалістю. Тому інтервал кореляції реальних повідомлень являється обмеженою величиною. Це обмеження, що записується у вигляді

$$B_s(\tau) = \begin{cases} B_s(\tau) & \text{при } \tau \leq \tau_0; \\ 0 & \text{при } \tau > \tau_0, \end{cases} \quad (1.13)$$

представлено на рис.1.9, а.

Н.А. Желєзновим доведено, що такі повідомлення можуть бути передвіщені системою лінійного прогнозування із середньоквадратичною похибкою, що як завгодно мало відрізняється від нуля, тільки на проміжку часу, рівному інтервалу кореляції τ_0 . Таким чином, згідно [9] дискретизацію слід проводити з інтервалом, який не перевищує τ_0 , оскільки лише в цьому випадку можливе безпомилкове відновлення вихідного повідомлення. Число некорельованих відліків N при цьому визначиться:

$$N = \frac{T_c}{\tau_0} \quad (1.14)$$

Визначення інтервалу кореляції τ_0 проводиться з використанням поняття ефективної смуги частот повідомлення:

$$\Delta\omega_{eff} = 2\pi\Delta f_{eff},$$

$$\tau_0 = \frac{1}{2\Delta f_{eff}}, \quad (1.15)$$

де $\Delta\omega_{eff}$ визначається як основа прямокутника з висотою, яка рівна максимальному значенню спектральної щільності $S_X(\omega)$ і площею, рівною площі під кривою спектральною щільності повідомлення (рис.1.9, б):

$$\Delta\omega_{eff} = \frac{1}{S_S(\omega)_{\max}} \int_0^{\infty} S_S(\omega) d\omega. \quad (1.16)$$

Для нестационарних функцій використовується поняття поточного інтервалу кореляції, який являється функцією часу $\tau_0 = \tau_0(t)$. У цьому випадку відлікові значення неперервного повідомлення будуть розташовуватись на осі часу нерівномірно. Метод дискретизації Н.А Желєзнова менш розроблений, ніж метод В.А. Котельникова, і тому не отримав ще широкого поширення.

Дискретизація неперервних повідомлень із використанням критерія найбільшого відхилення. Вибір кроку дискретизації з використанням даного критерія проводиться у припущенні, що вихідне повідомлення відновлюється за допомогою полінома степеня n . На деякому відрізку часу $[t_0, t_n]$ для рівновіддалених відліків відновлене повідомлення $S'(t)$ може бути представлено виразом [1]:

$$S'(t) = \sum_{k=0}^n S(t_k) \frac{(t-t_0)(t-t_1)\cdots(t-t_{k-1})(t-t_{k+1})\cdots(t-t_n)}{(t_k-t_0)(t_k-t_1)\cdots(t_k-t_{k-1})(t_k-t_{k+1})\cdots(t_k-t_n)}. \quad (1.17)$$

Вводячи скорочений запис, отримаємо:

$$S'(t) = \sum_{k=0}^n \left[S(t_k) \prod_{\substack{i=0 \\ i \neq k}}^n \frac{t-t_i}{t_k-t_i} \right]. \quad (1.18)$$

Для відновлення функції $S'(t)$ за допомогою полінома степеня n необхідно мати $n+1$ відліків.

Похибка відновлення вихідного повідомлення у цьому випадку визначиться залишковим членом:

$$R(t) = S(t) - S'(t) = \frac{\chi^{(n+1)}(\xi)}{(n+1)!} \prod_{k=0}^n (t-t_k), \quad (1.19)$$

де $\chi^{(n+1)}(\xi)$ - значення $(n+1)$ -ї похідної повідомлення $\chi(t)$, взятої в деякій точці ξ , що лежить всередині інтервалу $t_n - t_0$. Оскільки розташування точки ξ невідоме, для оцінки використовують модуль максимального значення похідної M_{n+1} на заданому інтервалі. Тоді:

$$R(t) = \frac{M_{n+1}}{(n+1)!} \prod_{k=0}^n (t - t_k), \quad (1.20)$$

Вводячи обмеження

$$R(t) \leq \varepsilon_0, \quad (1.21)$$

де ε_0 - допустима похибка дискретизації за часом, можна найти крок дискретизації або довжину інтервалу $t_n - t_0$, на якому можна визначити $n + 1$ відлікових значень неперервного повідомлення. При цьому інтервал $t_n - t_0$ є деякою функцією похибки ε_0 , степеня відтворюючого повідомлення і т.д.

Цікавість представляє визначення кроку дискретизації при використанні відтворюючих поліномів нульового та першого степенів.

Нульовому ступеню відтворюючого полінома відповідає ступінчата апроксимація неперервного повідомлення. У цьому випадку (1.21) прийме вигляд:

$$R(t) = \frac{M_1}{1!} (t_0 - t_1) = M_1 \Delta t \leq \varepsilon_0,$$

звідки

$$\Delta t \leq \frac{\varepsilon_0}{M_1}. \quad (1.22)$$

При $n = 1$ (лінійна апроксимація) залишковий член визначиться:

$$R(t) = \frac{M_2}{2!} \prod_{k=0}^1 (t - t_k) \leq \varepsilon_0,$$

Максимізуючи добуток $\prod_{k=0}^1 (t - t_k) = (t - t_0)(t - t_1)$ отримаємо

$$\frac{M_2 \Delta t^2}{8} \leq \varepsilon_0, \text{ звідки}$$

$$\Delta t \leq \sqrt{\frac{8\varepsilon_0}{M_2}}. \quad (1.23)$$

Відновлення неперервного повідомлення з використанням поліномів нульової і першої степені представлено на рис.1.10, а і б.

Значення кроку дискретизації для поліномів більш високих степенів такі:

Степінь полінома	0	1	2	3	4
Крок дискретизації	$\frac{\varepsilon_0}{M_1}$	$\sqrt{\frac{8\varepsilon_0}{M_2}}$	$\sqrt[3]{\frac{15,53\varepsilon_0}{M_3}}$	$\sqrt[4]{\frac{24\varepsilon_0}{M_4}}$	$\sqrt[5]{\frac{33\varepsilon_0}{M_5}}$

Слід відмітити, що для деяких видів повідомлень частота дискретизації, визначена по розглянутому методу, виявляється значно меншою від частоти, знайденої за теоремою Котельникова. Наприклад, для відновлення повідомлення, що описується

експоненціальною функцією $e^{-\frac{t}{\tau}}$ при $\tau = 1$ з похибкою, рівною 0,5%, за теоремою Котельникова необхідна частота дискретизації більша від 8 кГц, тоді як за критерієм найбільшого відхилення при $n = 1$ частота дискретизації буде складати 5 Гц. При похибці 0,05% значення частоти будуть відповідно більші від 700 кГц і 15,9 Гц. Різниця, зумовлена використанням різних методів відновлення повідомлень, є дуже істотною.

Розглянуті способи не вичерпують всіх можливих способів відновлення неперервних повідомлень і на основі їх дискретних значень. З цією метою можуть використовуватись різноманітні фільтри, характеристики яких здійснюють істотний вплив на вибір частоти дискретизації. Основними вихідними даними і в цьому випадку є ширина спектру повідомлення і припустима похибка відновлення неперервного повідомлення. Значення частоти дискретизації при використанні деяких видів фільтрів (фільтр Баттерворта, оптимальний фільтр, RC -фільтр) можна знайти в монографії [12].

Поняття про адаптивну дискретизацію. Для розглянутих раніше способів дискретизації характерна постійність інтервалів між відліковими значеннями функції, що дискретизується. В той же час із (1.22) і (1.23) випливає, що якщо яким-небудь способом фіксувати швидкість зміни або прискорення неперервного повідомлення, то з'являється можливість такої зміни кроку дискретизації у відповідності з ходом повідомлення, яке дискретизується, що похибка ніколи не буде перевищувати

встановленої величини ε_0 . При адаптивній дискретизації (для $n = 0$) відлікові значення $S(t_k)$ неперервного повідомлення беруться в моменти часу, коли

$$|S(t) - S(t_{k-1})| = \varepsilon_0. \quad (1.24)$$

В цей момент здійснюється також відлік величини інтервалу дискретизації τ_k (рис.1.10, в).

У випадку використання відтворюючого полінома першого степеня у пристрої дискретизації на кожному інтервалі генерується наближаюча пряма, і безперервно визначається різниця між поточним значенням неперервного повідомлення і наближаючою функцією. Коли різниця досягає величини ε_0 , береться черговий відлік, і визначається величина інтервалу дискретизації τ_k .

Вказані способи дискретизації відносяться до способів адаптивної дискретизації із не однократними інтервалами, оскільки величини інтервалів дискретизації можуть приймати довільні значення. Це зумовлює необхідність фіксації всіх значень τ_k

При адаптивній дискретизації з кратними інтервалами відлікові значення беруться у моменти часу, що віддалені один від одного на ціле число Δt , тобто $t_k = t_{k-1} + i\Delta t$, де $i = 1, 2, 3, \dots$, з використанням співвідношень, аналогічних (1.24).

При адаптивній дискретизації зменшується число відлікових значень неперервного повідомлення без збільшення похибки його відновлення, тобто знижується надлишковість інформації. Подальше зниження надлишковості може бути досягнуто шляхом зміни степені відтворюючого полінома.

Слід відмітити, що реалізація алгоритмів адаптивної дискретизації складніший, ніж алгоритмів рівномірної дискретизації.

1.3. Квантування за рівнем

Відліки неперервних повідомлень, взяті в дискретні моменти, характеризуються тим, що рівні повідомлень можуть приймати безкінечну множину значень, якщо точність їх визначення не обмежена. Оскільки любе повідомлення приходить до споживача сумісно з завадою (помилку вимірювання також можна

тракувати як заваду), то безглуздо передавати повідомлення точніше, ніж це дозволяє рівень завад. Таким чином, доцільно в залежності від рівня супроводжуючих завад і необхідної точності представлення повідомлення обмежити число рівнів повідомлення. Заміна неперервної шкали рівнів повідомлення дискретною шкалою рівнів називається квантуванням за рівнем. Якщо миттєве значення рівня повідомлення знаходиться всередині інтервалу, розташованого між двома дозволеними дискретними рівнями, то замість нього передається значення, що відповідає ближчому дозволеному рівню. При кількості рівнів квантування, рівній m , будь-яке повідомлення буде містити не більш, ніж m значень, які можна розрізнити.

Квантування за рівнем може здійснюватись з рівномірним чи нерівномірним кроком. При рівномірному квантуванні динамічний діапазон зміни повідомлення ділиться на $m - 1$ рівних частин (тобто інтервали між всіма рівнями квантування однакові). При нерівномірному квантуванні ділення проводиться на $m - 1$ нерівних частин. Таким чином отримуємо шкалу рівнів квантування.

Щоб мати можливість чітко визначити належність істинних миттєвих значень повідомлення певному рівню, динамічний діапазон повідомлення $S(t)$ ділиться на інтервали, які називають кроками квантування. Рівні квантування при цьому розташовуються всередині кроків квантування. Крок квантування зазвичай позначають ΔS , і для i -го рівня крок квантування буде ΔS_i . Якщо значення повідомлення в деякий момент часу (зазвичай цей момент співпадає з моментом взяття відліку за часом $S(t_k)$) знаходиться всередині i -го кроку квантування ΔS_i , то значення повідомлення замінюється значенням i -го рівня $S_i(t_k)$. Оптимальним стосовно точності відтворення квантованого сигналу є розташування рівня квантування в середині кроку квантування. Процедура квантування за рівнем представлена на рис.1.11, а.

Квантова не повідомлення $S_i(t_k)$ представляє собою суму істинного повідомлення $S(t_k)$ і помилки δ_{S_i} . Таким чином, квантування можна розглядати як проходження повідомлення через систему, що підлягає дії завади δ_{S_i} . Зазвичай цю заваду називають шумом квантування. Так як саме повідомлення являється випадковим процесом, шум квантування також буде випадковою функцією часу. Оцінімо статистичні характеристики шуму квантування. Виникаюча в момент t_k помилка визначається виразом:

$$\delta_{S_i} = S(t_k) - S_i(t_k). \quad (1.25)$$

Повідомлення $S(t)$ в момент часу t_k може приймати одне із можливих значень в інтервалі від $S_{\min}$ до $S_{\max}$ із деякою імовірністю, що визначається законом розподілу випадкової величини $S(t)$. Імовірність появи рівня S_i визначиться як імовірність попадання випадкової величини $S(t)$ в інтервал від S_a до S_B в окрузі рівня S_i . При цьому похибка δ_{S_i} може змінюватися в межах від $S_i - S_a$ до $S_B - S_i$ змінюючи знак (рис.1.11, б). Імовірність появи похибки δ_{S_i} визначається імовірністю появи відповідного значення повідомлення $S(t)$ в момент t_k .

Математичне очікування помилки:

$$M(\delta_{S_i}) = \int_{S_a}^{S_B} (S - S_i) p(S) dS, \quad (1.26)$$

а середньоквадратичне значення:

$$M(\delta_{S_i}^2) = \int_{S_a}^{S_B} (S - S_i)^2 p(S) dS, \quad (1.27)$$

Вважаючи динамічний діапазон зміни повідомлення набагато більшим від величини кроку квантування, прийнемо функцію щільності імовірності $p(S)$ на інтервалі інтегрування постійною і рівною своєму значенню для деякого середнього рівня $p(S_c)$. Тоді:

$$M(\delta_{S_i}) = p(S_c) \int_{S_a}^{S_B} (S - S_i) dS = \frac{1}{2} p(S_c) [(S_B - S_i)^2 - (S_a - S_i)^2] \quad (1.28)$$

а

$$M(\delta_{S_i^2}) = p(S_c) \int_{S_a}^{S_B} (S - S_i)^2 dS = \frac{1}{3} p(S_c) [(S_B - S_i)^3 - (S_a - S_i)^3], \quad (1.29)$$

$$\text{де } S_c = \frac{S_B - S_a}{2}.$$

Визначимо оптимальне положення рівня S_i відносно S_a і S_B , при якому середньоквадратична помилка буде мінімальною. Прирівнюючи до нуля похідну $\frac{\partial [M(\delta_{S_i^2})]}{\partial S_i}$, отримаємо:

$$(S_B - S_i)^2 = (S_a - S_i)^2$$

і визначимо положення рівня S_i , яке відповідає мінімуму помилки:

$$S_i = \frac{S_B + S_a}{2} = S_c. \quad (1.30)$$

Неважко помітити, що при цьому (1.28) перетворюється в нуль, тобто $M(S_i) = 0$.

Отриманий результат свідчить про те, що для забезпечення мінімальної дисперсії рівень квантування S_i слід розташовувати посередині кроку квантування. В цьому випадку $S_a = S_i - \frac{\Delta S_i}{2}$, $S_B = S_i + \frac{\Delta S_i}{2}$, а дисперсія знаходиться із (1.29) і дорівнює:

$$M(\delta_{S_i^2}) = \frac{1}{12} p(S_i) \Delta S_i^3. \quad (1.31)$$

Дисперсію шуму квантування D отримаємо шляхом сумування значень $M(\delta_{S_i,2})$ за всіма рівнями:

$$D = \sum_{i=1}^m M(\delta_{S_i^2}) = \frac{1}{12} \sum_{i=1}^m p(S_i) \Delta S_i^3. \quad (1.32)$$

У випадку рівномірного квантування, коли $\Delta S_1 = \Delta S_2 = \dots = \Delta S_m = \Delta S$, отримуємо:

$$D = \frac{\Delta S^2}{12} \sum_{i=1}^m p(S_i) \Delta S. \quad (1.33)$$

Вважаючи, що значення повідомлення обмежуються динамічним діапазоном, можна записати: $\sum_{i=1}^m p(S_i) \Delta S = 1$. Остаточно для рівномірного квантування:

$$D = \frac{\Delta S^2}{12}. \quad (1.34)$$

При виборі рівня, згідно до (1.30), максимальна різниця (при рівномірному квантування) між істинним значенням повідомлення і квантованим значенням не перевищує $\frac{\Delta S}{2}$. Середньоквадратична помилка квантування при цьому складе

$$\sigma_c = \sqrt{D} = \frac{\Delta S}{2\sqrt{3}}, \text{ тобто буде меншою від максимальної помилки в } \sqrt{3} \text{ раз.}$$

Із (1.31), (1.33) випливає, що значення дисперсії шуму квантування і середньоквадратичної помилки залежить як від виду функції розподілу щільності імовірності, так і від характеру зміни шуму квантування в межах динамічного діапазону повідомлення (1.32). Із аналізу (1.32) випливає, що для зменшення середньоквадратичної похибки доцільно для менш імовірних значень вибирати більший крок квантування, зменшуючи його до більш імовірних значень. Квантування за рівнем, яке забезпечує мінімум дисперсії шуму квантування, називають оптимальним квантуванням за рівнем. Умовою оптимальності квантування являється постійність добутку $\Delta S_i \sqrt[3]{p(S_i)}$ в межах динамічного діапазону. Функціональний зв'язок між функцією повідомлення Y і змінною S , що утворює в інтервалі від 0 до $S_{\max}$ оптимальну шкалу квантування, визначається наступним чином [13]:

$$S(y) = \frac{S_{\max}}{S_1} \int_0^y \frac{dy}{\sqrt[3]{p(y)}}, \quad (1.35)$$

де $S_1 = \int_0^{S_{\max}} \frac{dy}{\sqrt[3]{p(y)}}$; $p(y)$ - функція щільності ймовірності повідомлення $y(t)$.

Розрахунки, проведені з використанням (1.35), показали, що при використанні оптимального квантування неперервних функцій з нормальним законом розподілу є можливість зменшити дисперсію шуму квантування в порівнянні з рівномірним законом в окремих випадках більш, ніж в 3 – 4 рази. Аналогічно показана можливість зменшення дисперсії шуму квантування і для других законів розподілу. Можливі два способи оптимального квантування повідомлень: безпосередній і з попереднім перетворенням динамічного діапазону. При безпосередньому способі пристрій квантування на основі апріорно відомої функції розподілу $p(S)$ вибирає величину кроку квантування ΔS_i у відповідності із значеннями повідомлення, яке квантується.

Такий пристрій є достатньо складним за своєю структурою і операціями, які він виконує. При квантуванні з перетворенням динамічного діапазону повідомлення подається на нелінійний амплітудний перетворювач, що забезпечує перетворення закону розподілу в рівномірний (або близький до нього). Пристрій квантування оперує з однаковою величиною кроку квантування для всіх рівнів. Амплітудна характеристика перетворювача визначається виглядом функції $p(y)$. Часто на практиці використовується логарифмічна характеристика нелінійного перетворення. Як показано в роботі [12], використання такого перетворення дозволяє зменшити шум квантування без збільшення числа рівнів. Відмітимо, що при стискуванні динамічного діапазону повідомлення виявляються спотвореними. Для усунення спотворення повідомлення піддають зворотному нелінійному перетворенню у відповідності з прийнятним раніше законом стискування діапазону.

1.4 Геометричні представлення сигналів

Аналогія між сигналами і векторами. Будь-яка задача легше сприймається, якщо її можна зв'язати з яким-небудь відомим явищем. При математичному описі сигнали зручно розглядати як вектори чи точки в деякому функціональному просторі - *просторі сигналів*.

Електричні сигнали складної форми за своєю фізичною природою далеко не завжди схожі зі звичними нам представленнями векторів як спрямованих відрізків. Проте практичний інтерес має узагальнення операцій над векторами на сигнали (функції), які описують різні коливання. Справа в тому, що серед різних математичних прийомів, що використовуються при дослідженні електричних ланцюгів і сигналів,

найбільш широко застосовується представлення довільної функції у вигляді суми більш простих ("елементарних") функцій. Такий підхід лежить в основі принципу незалежності дії (суперпозиції) при вивченні перетворень сигналів у лінійних електричних ланцюгах. Наочні геометричні представлення, зв'язані з відображенням функцій в якості векторів простору сигналів, допомагають часто усвідомити фізичну сутність процесів формування, передачі і розподілу сигналів, синтезу оптимальних сигналів і пристроїв обробки сигналів при наявності завад.

Задача розкладу сигналу складної форми на найпростіші складові схожа з розкладанням звичайного вектора $\mathbf{S}$ *тривимірного простору* на його складові по *координатному базисі одиничних ортогональних векторів $\mathbf{i}, \mathbf{j}, \mathbf{k}$* (рис. 1.12, а). Таке представлення можна записати як

$$\mathbf{s} = s_1 \mathbf{i} + s_2 \mathbf{j} + s_3 \mathbf{k}. \quad (1.36)$$

Складовими вектора $\mathbf{s}$ за базисом $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ будуть вектори $s_1 \mathbf{i}, s_2 \mathbf{j}, s_3 \mathbf{k}$.

Коефіцієнти s_1, s_2, s_3 являють собою проекції вектора $\mathbf{s}$ на координатні осі $\mathbf{i}, \mathbf{j}, \mathbf{k}$ і називаються *координатами* вектора $\mathbf{s}$. Інакше кажучи, вектор $\mathbf{s}$ у тривимірному просторі цілком визначається сукупністю його координат

$$\mathbf{s} = (s_1, s_2, s_3).$$

Щоб перейти до узагальнення поняття вектора тривимірного простору для випадку *n-вимірного простору*, звернемося до прикладу. Деяке наближене представлення про функцію (сигнал) $s(t)$ можна скласти за її відображенням послідовністю прямокутних імпульсів, що мають на інтервалах $i\Delta$ значення $s(i\Delta)$ (мал. 1.12, б). Якщо тепер умовно представити функцію $s(t)$ на інтервалі $(0; T)$ "вектором", то для його визначення буде потрібно $n = T/\Delta$ координат $s_i = s(i\Delta)$. Це означає, що функцію $s(t)$ за аналогією з (1.36) можна представити у вигляді суми:

$$s(t) = \sum_{i=0}^{n-1} s_i \Psi_i(t), \quad (1.37)$$

де $\Psi_i(t) = \Psi(t - i\Delta)$ - елементарні базисні функції;

$$\Psi(t) = \begin{cases} 1, & t \in [0; \Delta]; \\ 0, & t \notin [0; \Delta]; \end{cases} \quad \frac{1}{\Delta} \int_0^{\Delta} \Psi^2(t) dt = 1.$$

Вектор $\mathbf{s}$, що відповідає функції $s(t)$, у n -вимірному просторі одиничних ортів Ψ_i буде цілком визначатися його координатами $\mathbf{S} = (s_0, s_1, \dots, s_{n-1})$. Таким чином, сигнал $s(t)$ довільної форми представляється сумою n найпростіших елементарних сигналів, у даному випадку у вигляді імпульсів прямокутної форми. Слово *простір* використовується тут, щоб придати множині сигналів геометричний зміст і тим самим наочність. Найбільш простим і в той же час фізично досить змістовним є трактування сигналів як елементів *нормованого лінійного метричного простору*.

Основні визначення, що відносяться до функціональних просторів Евкліда, Гільберта, Хеммінга. *Лінійним* чи *векторним* називається простір сигналів, для елементів якого виконуються правила додавання і множення на будь-яке число з деякої множини $\{\lambda\}$, що називають множиною скалярів. Додавання векторів проводиться по координатах, тобто сумою векторів $\mathbf{s}$ (функції $s(t)$) і $\mathbf{y}$ (функції $y(t)$)

називається вектор $\mathbf{s} + \mathbf{y} = (s_0 + y_0, s_1 + y_1, \dots, s_n + y_n)$, що належить даному простору, а добуток $\lambda \mathbf{s}$ вектора $\mathbf{s}$ на число λ дає вектор

$\lambda \mathbf{s} = (\lambda x_0, \lambda x_1, \dots, \lambda x_{n-1})$, що також належить даному простору. У лінійному просторі існує нульовий елемент 0 , такий, що $\mathbf{s} + 0 = \mathbf{s}$ і кожному елементу $\mathbf{s}$ відповідає протилежний елемент $-\mathbf{s}$, так що $\mathbf{s} + (-\mathbf{s}) = 0$. Вектор, утворений підсумовуванням n лінійно незалежних (базисних) векторів Ψ_i зі скалярними коефіцієнтами S_i

називається їх лінійною комбінацією $\mathbf{s} = \sum_{i=0}^{n-1} s_i \Psi_i$.

Множина векторів $\{\Psi_i\}$ називається *лінійно незалежною (базисом)*, якщо умова

$$\sum_{i=0}^{n-1} s_i \Psi_i = 0 \quad (1.38)$$

виконується лише тоді, коли всі $S_i = 0$. Інакше кажучи, лінійно незалежним

називається множина $\{\Psi_i\}$, для якої жодна з його компонент не може бути утворена лінійною комбінацією інших. Розмірність лінійного простору визначається числом

будь-яких лінійно незалежних базисних векторів $\{\Psi_i\}$, що утворюють цей простір.

Лінійно незалежні вектори $\{\Psi_i\}$ можна розглядати як координатні осі простору.

Метричним називається лінійний простір, у якому визначена *відстань* між елементами (векторами) простору (*метрика*), тобто кожній парі елементів, скажемо, s і y може бути поставлене у відповідність деяке дійсне не негативне число $d(s, y)$ і спосіб, відповідно до якого знаходиться це число. Відстань задовольняє наступним правилам:

$$1. d(s, y) = 0, \text{ якщо } s = y$$

$$2. d(s, y) = d(y, s);$$

$$3. d(s, y) \leq d(s, z) + d(z, y),$$

де s, y, z - елементи (точки) простору. Зміст перших двох умов очевидний. Третю умову називають *нерівністю трикутника*: довжина сторони трикутника менше (чи дорівнює) суми довжин двох інших сторін.

Нормовані простори. Серед лінійних метричних просторів важливе місце займають нормовані простори. Цей вид простору визначається завданням норми $\|s\|$, що задовольняє наступним аксіомам:

$$1. \|s\| \geq 0;$$

$$2. \|\lambda s\| = |\lambda| \cdot \|s\|;$$

$$3. \|s + y\| \leq \|s\| + \|y\|.$$

Перша аксіома встановлює, що норма є позитивне речовинне число, рівне нулю тільки для нульового вектора, у другій аксіомі λ — будь-яке число (скаляр), третя аксіома — аксіома трикутника.

Почнемо з перерахування термінів і визначень, що відносяться до n -вимірного дійсного евклідового простору R_n . Любий вектор s у цьому просторі визначається сукупністю його координат: $s = (s_0, s_0, \dots, s_{n-1})$.

Сукупність n лінійно незалежних векторів утворює n -вимірний евклідовий простір, що позначається R_n . Простір R_n можна визначити як множину точок, представлених кінцями векторів, для яких норма:

$$\|s\| = \sqrt{\sum_{i=0}^{n-1} s_i^2}.$$

Як бачимо, норма є узагальненням довжини вектора в двовірному і тривірному просторі. Відстань між двома векторами s і y визначається як норма різниці векторів:

$$d(s, y) = \|s - y\| = \sqrt{\sum_{i=0}^{n-1} (s_i - y_i)^2}.$$

Для простору Евкліда R_n можна ввести поняття скалярного добутку двох векторів s і y :

$$(s, y) = \sum_{i=0}^{n-1} s_i y_i = \|s\| \cdot \|y\| \cos \varphi, \quad (1.39)$$

де φ — кут між двома векторами. Для проєкцій s на y і назад, y на s , маємо:

$$\|s\| \cos \varphi = \frac{(s, y)}{\|y\|}, \quad \|y\| \cos \varphi = \frac{(s, y)}{\|s\|}$$

Координати вектора являють собою проєкції вектора на координатні осі, аналогічно (1.36)- Зі співвідношення (1.39) випливає очевидна нерівність:

$$\|(s, y)\| \leq \|s\| \cdot \|y\|, \quad (1.40)$$

відома в літературі як нерівність Буняковського-Шварца. Знак рівності має місце лише тоді, коли

$$y = \kappa s \quad (1.41)$$

(κ — скаляр), тобто коли вектори s і y колінеарні. Для відповідних сигналів $s(t)$ і $y(t)$ це означає, що вони співпадають за формою $y(t) = \kappa s(t)$. Квадрат вище визначеної норми вектора s можна знайти як скалярний добуток вектора на самого себе:

$$\|s\|^2 = (s, s).$$

При $n \rightarrow \infty$ простір R_n переходить у безкінечно вимірний простір Гільберта, що позначається L_2 . Гільбертовим простором є, зокрема простір усіх неперервних комплексних функцій аргументу t , що задані на інтервалі $(-\frac{T}{2}; \frac{T}{2})$, у якому скалярний добуток визначений співвідношенням:

$$(s, y) = \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t) y^*(t) dt, \quad (1.42)$$

а квадрат норми:

$$\|s\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t) s^*(t) dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} |s(t)|^2 dt. \quad (1.43)$$

Норма (1.43) має не тільки геометричний, але і виразний фізичний зміст. Так, якщо сигнал $s(t)$ — дійсний електричний струм в одиничному опорі 1 Ом, то квадрат норми:

$$\|s\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} s^2(t) dt = E_x$$

визначає енергію сигналу. Елементи гільбертового простору L_2 характеризуються квадратом, що інтегрується, тобто якщо елементи цього простору - дійсні сигнали $s(t)$, визначені на інтервалі $(-\frac{T}{2}; \frac{T}{2})$, то виконується умова:

$$E_x = \int_{-\frac{T}{2}}^{\frac{T}{2}} s^2(t) dt < \infty. \quad (1.44)$$

Гільбертовий простір позначається при цьому $L_2(T)$. При $T \rightarrow \infty$ одержуємо простір $L_2(\infty)$. Для деяких сигналів (функцій) простору $L_2(\infty)$ умова (1.44) при $T \rightarrow \infty$ може не виконуватися, але виконується умова:

$$P_x = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} s^2(t) dt < \infty. \quad (1.45)$$

У цьому випадку можна замість (1.42) ввести скалярний добуток з розмірністю потужності (для струмів і напруг на одиничному опорі):

$$(s, y)_P = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t) s^*(t) dt. \quad (1.46)$$

Квадрат норми вектора s у цьому випадку:

$$\|s\|_P^2 = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \|s(t)\|^2 dt = P_x. \quad (1.47)$$

При виконанні умови (1.45) у просторі $L_2(\infty)$ визначені й співвідношення (1.46) і (1.47) при $T \rightarrow \infty$.

Надалі, говорячи про функції з квадратом, що інтегрується, у просторі $L_2(\infty)$ маємо у виду виконання умови (1.44) чи умови (1.45) при $T \rightarrow \infty$. Квадрат відстані між двома векторами в дійсному просторі $L_2(T)$ визначається співвідношенням:

$$d^2(s, y) = \|s - y\|^2 = \int_{-\frac{T}{2}}^{\frac{T}{2}} (s(t) - y(t))^2 dt \quad (1.48)$$

або

$$d^2(s, y) = \|s - y\|_P^2 = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (s(t) - y(t))^2 dt. \quad (1.49)$$

Форму (1.49) можна використати і при $T \rightarrow \infty$ для сигналів з кінцевою середньою потужністю.

Простір L_2 являє собою природне узагальнення простору R_n , одержане шляхом переходу от дискретизованої функції до функції неперервного аргументу. У курсі ТЕЗ простір L_2 має особливе значення, оскільки воно дозволяє застосувати загальні геометричні представлення до повідомлень, сигналів і завад, визначені як функції неперервного аргументу. Спрямовуючи в (1.37) $n \rightarrow \infty$, одержуємо представлення неперервної функції $s(t)$ в просторі Гільберта:

$$s(t) = \sum_{i=0}^{\infty} s_i \Psi_i(t). \quad (1.50)$$

Простір з нескінченним числом вимірів являє собою природне узагальнення n -вимірного простору, що одержується шляхом граничного переходу от дискретної послідовності до функції неперервного аргументу. Помітимо, що норми векторів сигналів дорівнюють кореням з їх енергії, а скалярний добуток є мірою кореляції цих сигналів.

Сигнали кінцевої тривалості T обмежені смугою F , геометрично представляються різними векторами в n -вимірному просторі. Розходження між двома якими-небудь сигналами виражається через відстань між векторами, що зображують їх (мал.1.13). Відстань залежить від довжин векторів і кута між ними, а косинус останнього є не що інше, як коефіцієнт взаємної кореляції сигналів. Повна відсутність кореляції (тобто рівність нулю коефіцієнта кореляції) виражається через ортогональність векторів ($\alpha = \pi/2$).

Завада, обмежена тією ж смугою, що й сигнал, також визначається вектором у n -вимірному просторі. Цей вектор додається до вектора сигналу (мал.1.13). На відміну від вектора сигналу, вектор завади може мати будь-яку величину і напрямок (вектор є випадковим). У результаті при накладанні завади на сигнал навколо кінця вектора сигналу утвориться «хмара», змінна щільність якої виражає ймовірність попадання результуючого вектора $\mathbf{x} = \mathbf{s} + \mathbf{w}$ (вектора прийнятого сигналу) у даний елемент обсягу. Для флуктуаційної завади ця «хмара» має сферичну форму з ефективним радіусом

$$r = \sqrt{2TFP_{II}}.$$

Повідомлення (відеосигнал) $u(t)$, яке не містить коливання з частотами вище F_m , так само як і сигнал може бути представлено вектором у m -вимірному просторі, де $m = 2TF_m$. Сукупність можливих повідомлень визначає цей простір (простір повідомлень). На мал. 1.14 представлена двовірсна модель цього простору з двома різними повідомленнями u_1 і u_2 .

При передачі повідомлення $u(t)$ перетворюється в сигнал $s(t)$ з використанням деякого переносника $f(t)$. Математично цю операцію формування сигналу можна представити у вигляді

$$s = M(u, f), \quad (1.51)$$

де M — оператор, у загальному випадку нелінійний. Геометрично формування сигналу може бути представлено як перетворення простору повідомлень U у простір сигналів S : вектори u_1 і u_2 перетворюються у вектори s_1 і s_2 . Вимірність простору повідомлень m у загальному випадку не дорівнює вимірності простору сигналів m . При односмуговій передачі $m=n$. У випадку амплітудної модуляції сигнали мають вдвічі більше число координат, ніж повідомлення: $n = 2m$, а при частотній простір сигналів має значно більше число вимірів, ніж простір повідомлень.

При накладенні завади на сигнал утворюється область невизначеності, у яку попадають прийняті сигнали $X = S + \omega$. Взаємодія сигналу і завади можна виразити через оператор:

$$x = \Psi(s, \omega). \quad (1.52)$$

Оператор Ψ перетворює простір сигналів S у простір прийнятих сигналів X : вектори s_1 і s_2 переходять в x_1 і x_2 .

Приймач по прийнятих сигналах x відтворює передане повідомлення:

$$v = \Phi(x) \quad (1.53)$$

тобто перетворює простір сигналів X у простір прийнятих повідомлень V .

Якщо завада відсутня, то прийнятий сигнал перетворюється в повідомлення зворотним оператором $v = M^{-1}(s)$. У цьому випадку прийняте повідомлення тотожне переданому $v = M^{-1}(s) = M^{-1}M = u$.

При наявності завад повідомлення на прийомі відтворюються з деякою помилкою: замість повідомлення u_1 може бути відтворене повідомлення u_2 і навпаки. Помилка відбудеться, коли результуючий вектор x виявиться ближче до кінця вектора того сигналу s , що у даний момент не передається. Можна побудувати приймач, що відтворює повідомлення $v = u_1$ кожного разу, коли кінець вектора ближчий до кінця вектора s_1 ніж до кінця вектора s_2 і навпаки. Такий приймач по Котельникову називається ідеальним чи *оптимальним*. Очевидно, помилка при оптимальному прийомі буде тим меншою, чим більша відстань d між сусідніми сигналами. Ця відстань, у свою чергу, залежить від відстані r між сусідніми повідомленнями та від виду оператора M , що визначає спосіб формування сигналу (зокрема, спосіб модуляції).

Оператори і функціонали. У задачах перетворення повідомлень і сигналів нам будуть потрібні деякі узагальнення функціональних залежностей. Величина y називається функцією незалежної змінної x , якщо кожному значенню x (з множини його можливих значень) відповідає визначене значення y . Інакше кажучи, функціональна залежність $y = f(x)$ встановлює відповідність між деякою множиною чисел x і множиною чисел або, що те ж саме, функція встановлює залежність одного числа от другого.

Більш загальним поняттям є поняття функціонала. Функціонал установлює відповідність між множиною чисел, з однієї сторони, і деякою множиною функцій — з іншої. Можна сказати, що функціонал Φ встановлює залежність числа від функції: $y = \Phi(f(x))$. Прикладом функціоналу є визначений інтеграл, величина якого (при незмінних межах) залежить від виду підінтегральної функції.

Дуже корисним є поняття функціонального оператора, який установлює відповідність між двома множинами функцій, тобто за допомогою оператора L встановлюється залежність функції від функції: $y(t) = L[x(t)]$. Так як функції можуть бути представлені векторами, і множина функцій визначається як векторний простір, то дія оператора може бути описаною у геометричних термінах як перетворення простору X векторів x у простір Y векторів y . Зворотнє перетворення Y у X позначають L^{-1} . У задачах перетворення повідомлень і сигналів використовуються поряд з лінійними операторами також нелінійні і параметричні оператори.

Векторне представлення цифрових сигналів у просторі Хеммінга. Якщо функція $x(t)$ на кожнім інтервалі $i\Delta$ може приймати одне з m можливих значень $x_i^{(k)}$ ($k = \overline{0, m-1}$), то на відрізок тривалістю T вона буде цілком визначена $n = \frac{T}{\Delta}$ чи значеннями $x_i^{(k)}$, що те ж, сукупністю коефіцієнтів $\{x_0^{(k)}, x_1^{(l)}, \dots, x_{n-1}^{(v)}\}$ ($k, l, v \in \overline{0, m-1}$), яку називають *n-набором*. Зокрема,

при $m = 2$ коефіцієнт $x_i^{(k)}$ приймає значення 0 чи 1, n -набір являє собою просто кодову комбінацію n -значного двійкового ($m=2$) коду, що відображає символ (букву, цифру) переданого повідомлення. Двійкові n -набори відображаються векторами (точками) у просторі Хеммінга 2_n . Скалярний добуток у цьому просторі зручно задати функцією:

$$(x, y) = \left[\sum_{k=0}^{m-1} x_k y_k \right]^2,$$

де $\sum$ — сума в звичайному змісті. Звідси норма двійкового вектора:

$$\|x\| = \sum_{k=0}^{n-1} x_k^2 = \sum_{k=0}^{n-1} x_k$$

Можна бачити, що норма двійкового вектора визначається кількістю одиниць, що містяться в ньому. Цю норму називають також вагою вектора (кодової комбінації) і позначають ω . Відстань у просторі Хеммінга:

$$d(x, y) = \|x - y\| = \sum_{k=0}^{n-1} |x_k - y_k| = \sum_{k=0}^{n-1} |x_k \oplus y_k|,$$

де знак $\oplus$ означає операцію підсумовування за модулем 2 (mod 2): $0 \oplus 0 = 0$, $0 \oplus 1 = 1$, $1 \oplus 0 = 1$, $1 \oplus 1 = 0$. Приведемо приклад підсумовування за mod 2 двох векторів:

$$\begin{array}{r} x = \{1001011\} \\ \oplus \\ y = \{0101101\} \\ \hline x \oplus y = \{1100110\} \end{array}$$

Підсумовування і вирахування за mod 2 еквівалентні. У просторі Хеммінга відстань між двійковими векторами визначається по числу позицій у кодовій комбінації, у яких вектори x і y мають символи, що відрізняються. У розглянутому прикладі $d(x, y) = 4$ одиницям. У більш загальному випадку, якщо число значень, які розрізняються, дорівнює m , використовується різниця за модулем m .

1.5 Розклад сигналів в узагальнений ряд Фур'є

Важливим поняттям у просторах Евкліда, Гільберта і Хеммінга є ортогональність векторів. Два вектори x і y ортогональні, якщо:

$$(x, y) = 0 \quad (1.54)$$

Легко показати, що якщо вектори Ψ_i і Ψ_j при $j \neq i$ взаємно ортогональні, то вони також лінійно незалежні. Тому сукупність ортогональних векторів можна використовувати як базис лінійних просторів.

Представимо неперервну (у часі і за рівнем) функцію $x(t)$ з квадратом, що інтегрується, у $L_2(T)$ просторі через довільну ортонормовану систему базисних $\{\Psi_i(t)\}$ функцій, для яких:

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} \Psi_i(t) \Psi_j(t) dt = \begin{cases} 0, & j \neq i, \\ 1, & j = i. \end{cases} \quad (1.55)$$

Замість (1.53) маємо представлення:

$$x(t) = \sum_{i=0}^{\infty} C_i \Psi_i(t), \quad (1.56)$$

де C_i - коефіцієнти (координати) розкладу в ортонормованому базисі $\{\Psi_i(t)\}$. Представлення (1.56) називають узагальненим рядом Фур'є. Для визначення коефіцієнтів C_i знайдемо скалярний добуток:

$$(x, \Psi_j) = \sum_{i=0}^{\infty} C_i \int_{-\frac{T}{2}}^{\frac{T}{2}} \Psi_i(t) \Psi_j(t) dt.$$

Із врахуванням (1.55) отримуємо:

$$C_j = (x, \Psi_j) = \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \Psi_j(t) dt. \quad (1.57)$$

Таким чином, коефіцієнти узагальненого ряду Фур'є C_i є проєкціями вектора x на ортогональні осі (одичні орти) Ψ_i . З врахуванням (1.55) і (1.56) можна одержати:

$$E_x = \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt = \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[\sum_{i=0}^{\infty} C_i \Psi_i(t) \right]^2 dt = \sum_{i=0}^{\infty} C_i^2. \quad (1.58)$$

(1.58) є частинним випадком рівності Парсеваля: З врахуванням ортонормованого базису $\{\Psi_i(t)\}$ легко бачити, що скалярний добуток і норму в просторі R_η можна знаходити за формулами (1.42) і (1.43).

Представимо тепер приблизно функцію $x(t)$ розкладу в усічений ряд по ортонормованих базисних функціях $\{\Psi_i(t)\}$:

$$x_\varepsilon(t) = \sum_{i=0}^{\infty} \gamma_i \Psi_i(t) \quad (1.59)$$

і визначимо коефіцієнти γ_i так, щоб мінімізувати середньоквадратичну похибку (СКП):

$$\begin{aligned}
\overline{\varepsilon^2(t)} &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} [x(t) - x_\varepsilon(t)]^2 dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \left[x(t) - \sum_{i=0}^{n-1} \gamma_i \Psi_i(t) \right]^2 dt \\
&= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt - \frac{2}{T} \sum_{i=0}^{n-1} \gamma_i \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \Psi_i(t) dt + \frac{1}{T} \sum_{i=0}^{n-1} \gamma_i^2.
\end{aligned}$$

З врахуванням (1.57) можна записати:

$$\overline{\varepsilon^2(t)} = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt + \frac{1}{T} \sum_{i=0}^{n-1} (\gamma_i - C_i)^2 - \frac{1}{T} \sum_{i=0}^{n-1} C_i^2. \quad (1.60)$$

Похибка $\overline{\varepsilon^2(t)}$ приймає мінімальне значення, якщо $\gamma_i = C_i$, тобто якщо коефіцієнти розкладу в усиченому представленні (1.59), є коефіцієнтами узагальненого ряду Фур'є. Позначивши $\overline{\varepsilon^2(t)} = \Delta_n$, можна записати виходячи з (1.60):

$$\Delta_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt - \frac{1}{T} \sum_{i=0}^{n-1} C_i^2 \geq 0$$

або

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} x^2(t) dt \geq \sum_{i=0}^{n-1} C_i^2.$$

Нерівність (1.61) називають нерівністю Бесселя: Зі зростанням n величина Δ_n зменшується. Якщо при $n \rightarrow \infty$ СКП прямує до нуля, то систему базисних функцій $\{\Psi_i(t)\}$ називають повною. Маючи на увазі, що при $n \rightarrow \infty$ справедливо (1.58), можна стверджувати, що в просторі Гільберта система базисних функцій $\{\Psi_i(t)\}$ є повною. Ця система функцій є також замкнутою, так як для будь-якої функції $x(t)$ з $L_2(T)$ нерівність (1.61) переходить при $n \rightarrow \infty$ в рівність.

1.1. Аналітичний сигнал

Реальні процеси описуються дійсними функціями часу. Однак при розгляді багатьох задач зручно виражати сигнал у вигляді суми елементарних сигналів, кожний з яких є комплексною функцією часу, або розглядати сам сигнал як комплексну функцію:

$$\hat{s}(t) = s(t) + i s(t) = U(t) e^{i\psi(t)}, \quad (1.62)$$

де $\psi(t)$ — огибающая і фаза сигналу. Дійсний сигнал $s(t)$ у цьому випадку визначається наступним виразом:

$$s(t) = \operatorname{Re} U(t) e^{i\psi(t)} = U(t) \cos \Psi(t). \quad (1.63)$$

Таке представлення особливо зручне для вузькосмугових сигналів.

Сигнал $s(t)$ називається «аналітичним», якщо $s(t)$ і $\hat{s}(t)$ складають пари перетворень Гільберта:

$$\left. \begin{aligned} \hat{s}(t) &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau \\ s(t) &= -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{s}(\tau)}{t - \tau} d\tau \end{aligned} \right\}. \quad (1.64)$$

Інтеграл в цих співвідношеннях розуміється у вигляді основного значення Коші. Функція $\hat{s}(t)$ називається спряженою з функцією $s(t)$ за Гільбертом. При такому виборі $s(t)$ і $\hat{s}(t)$ огинаюча і фаза сигналу визначаються однозначно:

$$U(t) = \sqrt{[s(t)^2 + \hat{s}(t)]^2} \quad (1.65)$$

$$\psi(t) = \arctan \frac{\hat{s}(t)}{s(t)}. \quad (1.66)$$

Якщо ефективна ширина спектра сигналу $s(t)$ мала в порівнянні з його середньою частотою ω_0 , то $U(t)$ і $\psi(t)$ змінюються повільно в порівнянні з функцією $s(t)$. Легко показати, що функції $s(t) = \cos \omega t$ відповідає спряжена

функція $\hat{s}(t) = \sin \omega t$, а функції $\hat{s}(t) = \sin \omega t$ відповідає $s(t) = -\cos \omega t$. Якщо вихідний сигнал представлений рядом Фур'є $s(t) = \sum_{\kappa} (a_{\kappa} \cos \omega_{\kappa} t + b_{\kappa} \sin \omega_{\kappa} t)$,

то спряжена до нього функція буде $\hat{s}(t) = \sum_{\kappa} (a_{\kappa} \sin \omega_{\kappa} t - b_{\kappa} \cos \omega_{\kappa} t)$. Таким чином, найпростішому сигналу у вигляді гармонійного коливання $s(t) = A \cos \omega t$

відповідає аналітичний сигнал $s(t) = A \cos \omega t + i A \sin \omega t = A e^{i \omega t}$. Якщо сигнал $s(t)$ може бути представлений у вигляді інтегралу Фур'є

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(i\omega) e^{i\omega t} d\omega, \text{ то частотний спектр сигналу буде:}$$

$$\hat{F}[s(t)] = [-i \operatorname{sgn}(\omega) S(i\omega)], \quad (1.67)$$

де

$$\operatorname{sgn}(\omega) = \begin{cases} +1, & \omega > 0, \\ 0, & \omega = 0, \\ -1, & \omega < 0. \end{cases}$$

Таким чином, перетворення Гільберта сигналу $s(t)$ можна розглядати як результат проходження $s(t)$ через ланцюг, що зсовує фазу всіх частотних і перехідних компонент спектра на кут -90° . Частотна і перехідна характеристики цього ланцюга такі:

$$\begin{aligned} K(i\omega) &= -i \operatorname{sign}(\omega), \\ h(t) &= 1/\pi t. \end{aligned}$$

Підстановка виразу (1.67) у перетворення Фур'є функції (1.62) показує, що спектр $S(i\omega)$ сигналу $s(t)$ є «однобічним»:

$$S(i\omega) = \begin{cases} 2S(i\omega), & \omega > 0, \\ S(0), & \omega = 0, \\ 0, & \omega < 0. \end{cases} \quad (1.68)$$

Це дуже корисна властивість аналітичного сигналу.

Для періодичних функцій $s(t)$ спряжена за Гільбертом функція $s^\wedge(t)$ також є періодичною з тим же періодом. Ці функції на відрізку, рівному їх періоду T , взаємно

ортогональні, тобто $\int_0^T s(t) s^\wedge(t) dt = 0$.

Сигнали $s_i(f)$ і $s_j(f)$, для яких умова ортогональності зберігається при заміні одного з них на спряжений за Гільбертом, тобто коли виконуються умови:

$$\left. \begin{aligned} (s_i, s_j) &= \frac{1}{T} \int_{-T/2}^{T/2} s_i(t) s_j(t) dt = 0 \\ (s_i, s_j) &= \frac{1}{T} \int_{-T/2}^{T/2} s_i(t) s_j^\wedge(t) dt = 0, \quad i \neq j \end{aligned} \right\}, \quad (1.69)$$

називаються ортогональними в підсиленому змісті. Можна показати, що такі сигнали також задовольняють і умови:

$$(s_i, s_j) = \frac{1}{T} \int_{-T/2}^{T/2} s_i(t) s_j^*(t) dt = 0, i \neq j, \quad (1.70)$$

де «зірочка» позначає комплексно-сполучену величину.

Введення поняття «аналітичний сигнал» дозволяє представити кожну функцію часу (при несуттєвих обмеженнях) у комплексному вигляді, й однозначно визначити огинаючу і фазу сигналу. В аналітичній формі (1.67) можуть бути надані як детерміновані так і випадкові сигнали. Корисність такого представлення полягає в можливості окремого розгляду огинаючої і фази процесу (чи сигналу завади). Так, при дослідженні випадкового процесу можна замість вивчення миттєвих значень обмежитися вивченням огинаючої і фази процесу, а в багатьох випадках — або огинаючої, або фази.

У загальному випадку спектри і кореляційні функції процесів $x(t)$ і $x(t)^\wedge$ однакові: $G_x(\omega) = G_{x^\wedge}(\omega)$, $B_x(\tau) = B_{x^\wedge}(\tau)$.

Взаємний енергетичний спектр $G_{x^\wedge}(\omega)$, дорівнює $G_{x^\wedge}(\omega) = iG_x(\omega)$, а
взаємна кореляційна функція визначається виразом

$$B_{x^\wedge}(\tau) = -B_x(\tau) = \frac{1}{\pi} \int_0^\infty G_x(\omega) \sin \omega \tau d\omega. \quad \text{Між законом розподілу}$$

випадкового процесу $x(t)$ з однієї сторони і законами розподілу огинаючої U і фази ψ , з іншої, існує жорсткий зв'язок: знаючи закон розподілу $p(x)$, можна знайти закони розподілу $p(U)$ і $p(\psi)$. Порівняно легко цей зв'язок встановлюється у випадку гаусовських процесів.

Контрольні питання

1. Дайте визначення детермінованому сигналу.
2. Які види детермінованих сигналів Ви знаєте?
3. Які види еталонних сигналів Ви знаєте?
4. Які види детермінованих сигналів можна розкласти в ряд Фур'є?
5. Який спектр має періодичний сигнал?
6. Запишіть ряд Фур'є в звичайній та комплексній формі і визначте значення коефіцієнтів ряду.

7. Напишіть пряме та зворотнє перетворення Фур'є для неперіодичного сигналу.
8. Як визначається середня потужність періодичного сигналу?
9. Що розуміється під практичною шириною спектра сигналу?
10. Які дискретизації сигналів за часом Ви знаєте?
11. Сформулюйте теорему В.А. Котельникова.
12. Запишіть ряд В.А. Котельникова.
13. В чому сутність теореми В.А. Котельникова для сигналів, спектр яких не розпочинається з нульової частоти?
14. У чому особливості кореляційного критерію відліків Н.А. Желєзнова?
15. Як оцінюється похибка дискретизації, коли не враховуються частоти, що лежать вище граничної частоти спектра сигналу?
16. У чому особливості квантування критерію відліків Темникова?
17. У чому полягають особливості адаптивної дискретизації сигналів за часом?
18. У чому полягає суть квантування за рівнем сигналів?
19. Як визначається помилка квантування?
20. Дайте визначення функціональних просторів за Евклідом, Гільбертом і Хеммінгом.
21. Дайте визначення нормованому простору.
22. Як визначається відстань між двома векторами?
23. Дайте визначення «хмарі» завад при геометричному представленні сигналу?
24. У чому полягають особливості представлення цифрових сигналів у просторі Хеммінга?
25. Запишіть узагальнений ряд Фур'є.
26. Дайте визначення аналітичного сигналу.

Запитання, задачі і вправи

- 2.1. Задані на інтервалі $(-O/2; O/2)$ сигнали $x(t) = \sqrt{2} \cos(2\pi t/T)$ і $y(t) = \sqrt{2} \cos(4\pi t/T)$ трактуються як елементи простору Гільберта $L_2(T)$. Доведіть, що: а) норми цих сигналів рівні 1; б) їх скалярний добуток дорівнює 0; в) відстань між ними дорівнює $\sqrt{2}$.
- 2.2. Доведіть, що в n -вимірному просторі Хеммінга 2_n існує лише один ортонормований базис, утворений n векторами:

$$(1,0,0,\dots,0), (0,1,0,\dots,0), \dots, (0,0,0,\dots,1).$$
- 2.3. Доведіть, що функції Уолша $wal(3, \theta)$ і $wal(5, \theta)$ ортогональні на інтервалі $(-1; +1)$.

- 2.4. Як визначається похибка представлення сигналу $x(t)$ як елемента простору $L_2(T)$ усіченим узагальненим рядом Фур'є?
- 2.5. Знайдіть спектральні компоненти періодичної послідовності прямокутних відеоімпульсів тривалістю τ з періодом проходження $T = 2\tau$ по спектральній щільності Фур'є одиночного імпульсу.
- 2.6. Покажіть, що якщо спектральна щільність за Фур'є $\dot{S}_y(f) = \dot{S}_x(f)\dot{S}_g(f)$ сигналу $y(t)$ дорівнює добутку спектральних щільностей сигналів $x(t)$ і $g(t)$, то сам сигнал $y(t)$ визначається часовою згорткою сигналів $x(t)$ і $g(t)$: $y(t) = x(t) \otimes g(t)$. Аналогічно якщо $y(t) = x(t)g(t)$, то $\dot{S}_y(f) = \dot{S}_x(f) \otimes \dot{S}_g(f)$
- 2.7. Користуючись визначенням $\mathcal{D}$ -функції, знайдіть спектральну щільність за Фур'є постійної складової $x(t) = a$.
- 2.8. Даний прямокутний імпульс $x(t) = A, t \in [-\tau_c/2; \tau_c/2]$. Знайдіть його ФК $B_E(\tau)$ (з розмірністю енергії) і $B(\tau)$ (з розмірністю потужності), а також СПЕ $W_x(f)$ і СПМ $G_x(f)$.
- 2.9. Чому дорівнює максимальний інтервал дискретизації мовного і телевізійного сигналу, якщо верхню частоту в спектрі сигналу прийняти рівною відповідно 4 кГц і 6,5 МГц?
- 2.10. З яких розумінь інтервал дискретизації Δ неперервної функції часу з фінітним спектром на відрізку $[-F_B; +F_B]$ часто обирають меншим, ніж $1/2F_B$?
- 2.11. Є експоненціальний імпульс $x(t) = 1(t)e^{-\alpha t}$. Яка відносна похибка при відновленні цього сигналу за його дискретними відліками при кроці дискретизації $\Delta = 1/2F_B$?
- 2.12. Дискретне стаціонарне двійкове джерело описується простим ланцюгом Маркова з матрицею перехідних імовірностей $\begin{bmatrix} P(1|1') & P(1|0') \\ P(0|1') & P(0|0') \end{bmatrix}$, де $P(a_i|a'_j)$ -

імовірність появи символу a_i за умови, що йому передувє символ a'_j . Нехай $P(1|1') = 0.9$; $P(1|0') = 0.7$. Визначити: ймовірності $P(0|1')$, $P(0|0')$; безумовні ймовірності передачі символів $P(0)$ і $P(1)$; ймовірності передачі ланцюжків із трьох символів $(1,1,0)$ і $(0,1,0)$.

- 2.13. Дискретний випадковий процес приймає три значення з ймовірностями $p(a_1) = 0,3$; $p(a_2) = 0,2$ $p(a_3) = 0,5$. Написати вираз для щільності ймовірності процесу, побудувати графік ІФР.

- 2.14. Покажіть, що випадковий вузько смуговий процес $X(t) = A_c(t) \cos \omega_0 t - A_s(t) \sin \omega_0 t$ з незалежними квадратурними компонентами $A_c(t)$ і $A_s(t)$ стаціонарний лише тоді, коли $\overline{A_c(t)} = \overline{A_s(t)} = 0$ і $\overline{A_c^2} = \overline{A_s^2} = \sigma^2 = \text{const}$

- 2.15. СП $X(t) = A \cos(\omega_0 t + \Phi)$ з рівномірно розподіленою фазою ергодичний. Знайдіть його МО. Напишіть аналітичний вираз для ФК і СПМ.

- 2.16. СП, визначений на відріжку $[-T/2; +T/2]$, має експоненціальну функцію

кореляції $B(\tau) = \exp\left\{-\frac{|\tau|}{\tau_{kor}}\right\}$. Знайти канонічний розклад цієї функції

по ортонормованій системі базисних функцій $\left\{\sqrt{2} \cos(2\pi i t / T)\right\}$, $i = 0, 1, 2, \dots$

- 2.17. Напишіть двовимірну щільність ймовірності $w(x_1, x_2, \tau)$ стаціонарного гаусовського випадкового процесу з кореляційною матрицею

$$R = \begin{vmatrix} 1 & e^{-\alpha|\tau|} \\ e^{-\alpha|\tau|} & 1 \end{vmatrix} \text{ МО } m_1 = 0 \text{ і дисперсією } \sigma^2. \text{ Визначите умовну ПВ}$$

цього процесу $w(x_2, t_2 | x_1, t_1)$.

- 2.18. Сигнал $x(t)$ є сумою двох гармонійних компонент: $x(t) = U_1 \cos \omega_1 t + U_2 \cos \omega_2 t$ Знайти: спряжений по Гільберту сигнал $\hat{x}(t)$, аналітичний сигнал $\dot{x}(t)$, огибающую $A(t) = \sqrt{x^2(t) + \hat{x}^2(t)}$, повну

фазу $\Psi(t) = \arctg(\hat{x}(t)/x(t)),$

миттєву

частоту

$$\omega(t) = \frac{d\Psi(t)}{dt} = \frac{\hat{x}'(t)x(t) - x'(t)\hat{x}(t)}{\hat{x}^2(t) + x^2(t)}.$$

2.19. Знайти єдине диференціальне рівняння коливань у контурі за заданою системою рівнянь стану (2.151).

2.20. Знайти інтервал кореляції $\tau_{\text{кор}}$ за методом рівновеликого прямокутника для випадкового стаціонарного синхронного двійкового сигналу $X_{\hat{A}}(t)$. Чому дорівнює ефективна ширина спектра $F_{\dot{y}}$

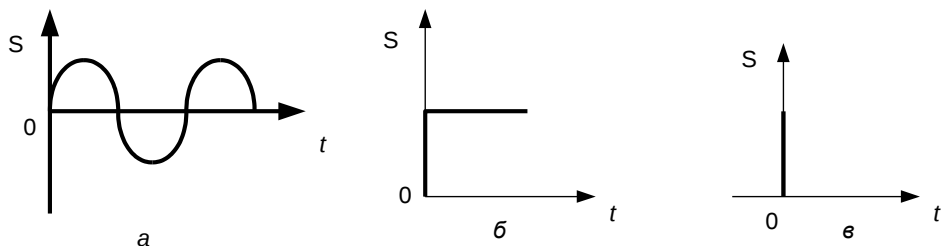


Рис. 1.2. Еталонні сигнали: а - гармонічний, б - одинична функція, в - дельта функція

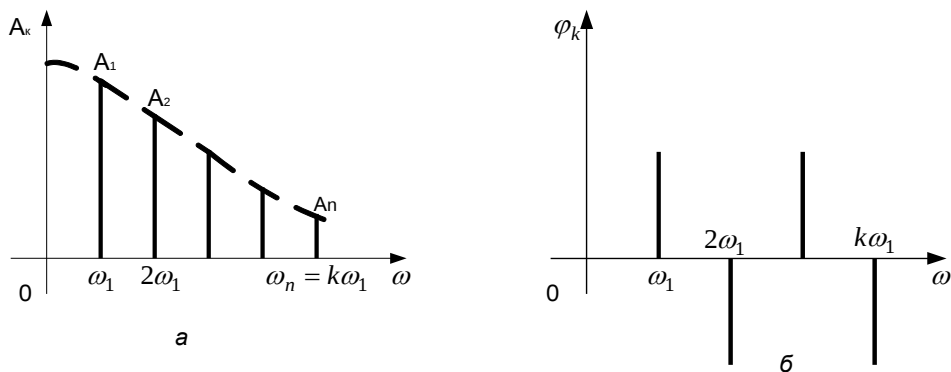
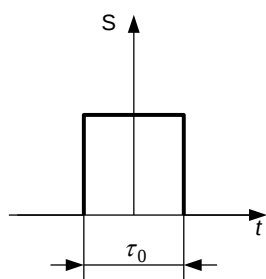
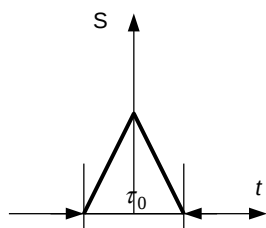
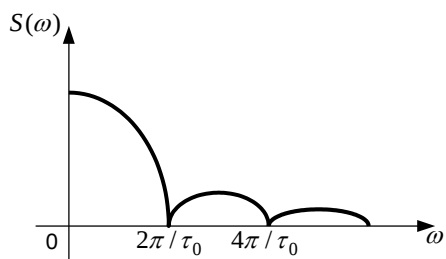


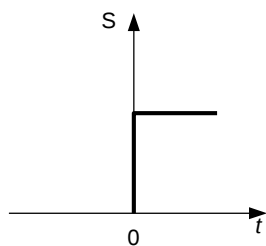
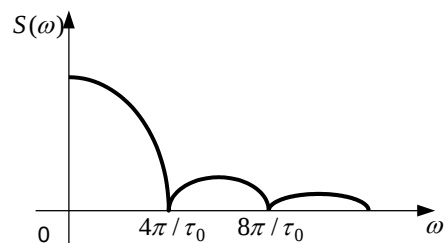
Рис. 1.3 спектр амплітуд (а) і фаз (б) періодичного сигналу.



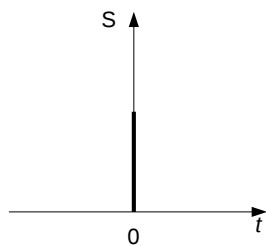
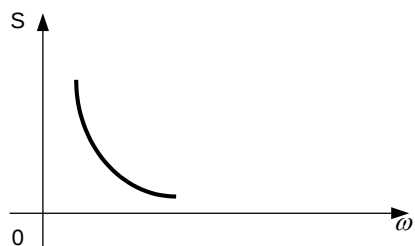
a



б



в



г

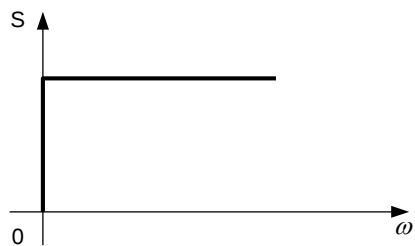


Рис. 1.4 Спектры деяких неперіодичних функцій.

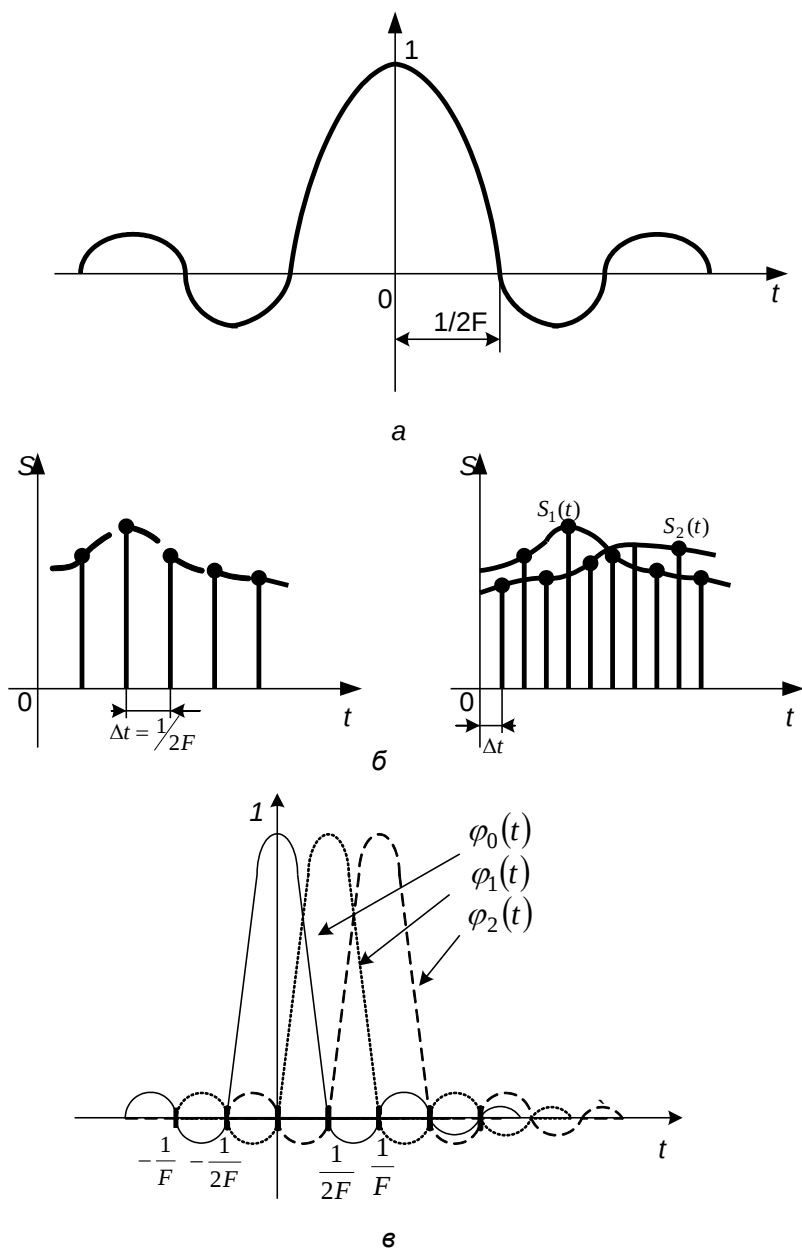


Рис.1.5. Функції відліків і дискретизація сигналу за часом: а, в - функції відліків; б - дискретизація

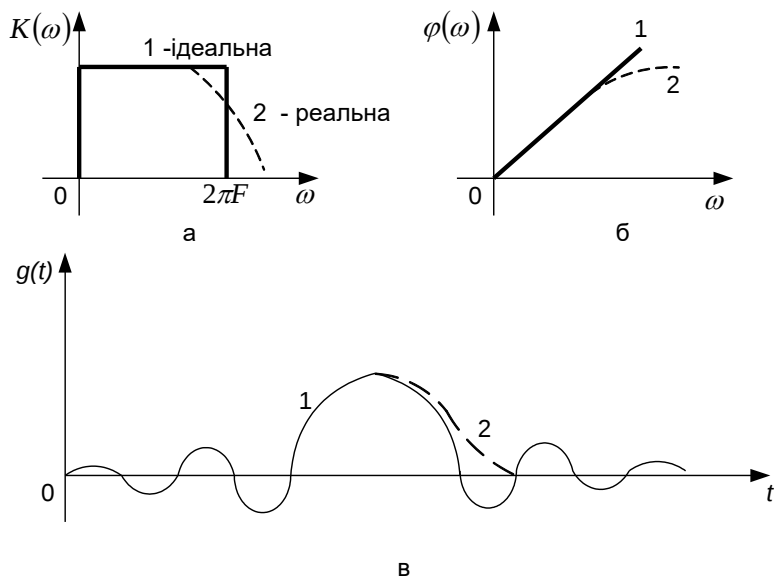


Рис. 1.6 Характеристика фільтру нижніх частот: а - частотна, б - фазова, в - імпульсна; 1 - ідеальний фільтр, 2 - реальний фільтр.

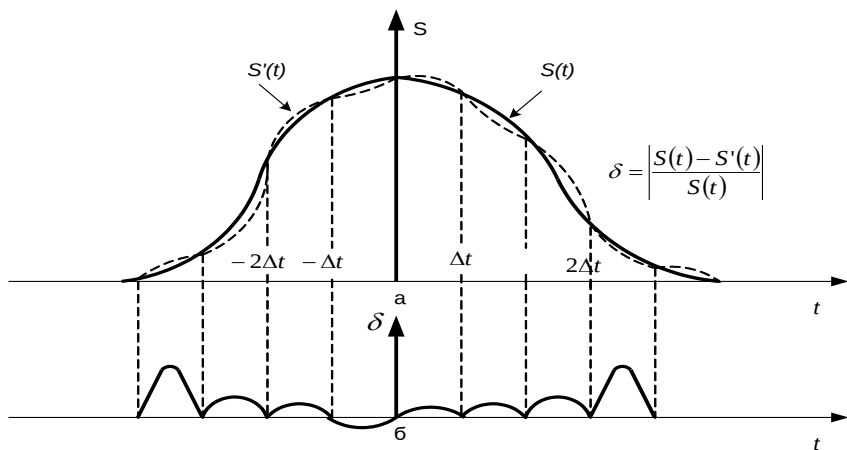


Рис. 1.8 До оцінки похибки відновлення неперервної функції.

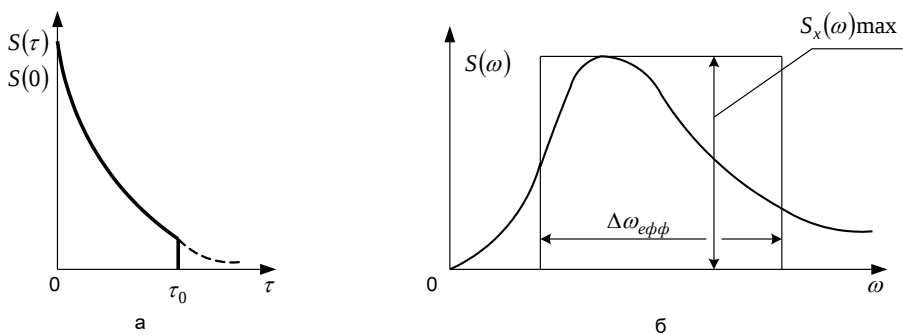


Рис. 1.9. До обмеження інтервалу кореляції (а) та визначення ефективної полоси частот (б) .

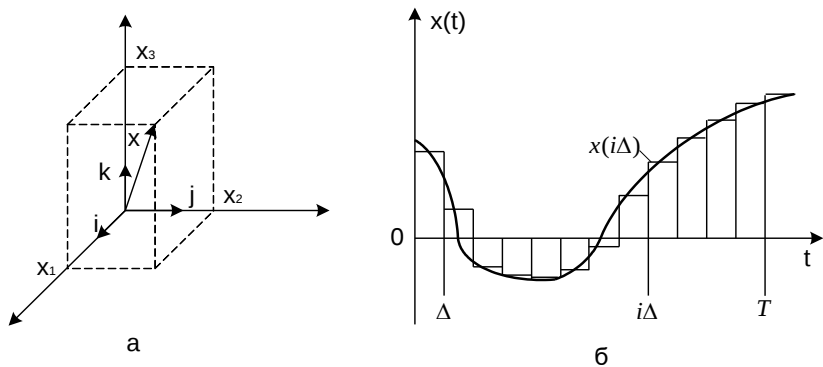


Рис. 1.12. Представлення вектора $\mathbf{x}$ в тривимірній ортогональній системі координат (а) та відображення неперервної функції послідовністю прямокутних імпульсів (б)

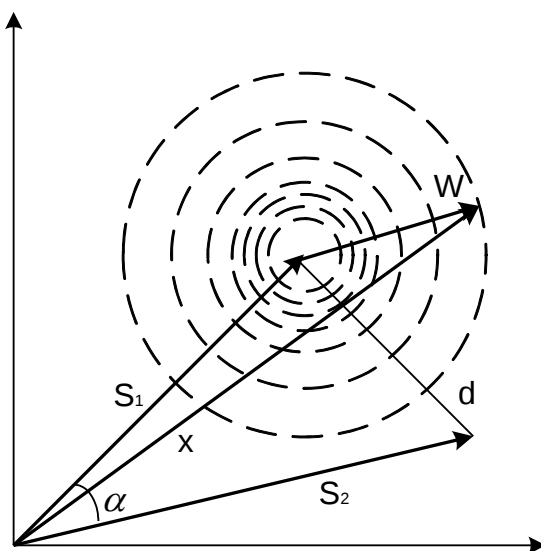


Рис.1.13. Векторна діаграма сигналу та завади

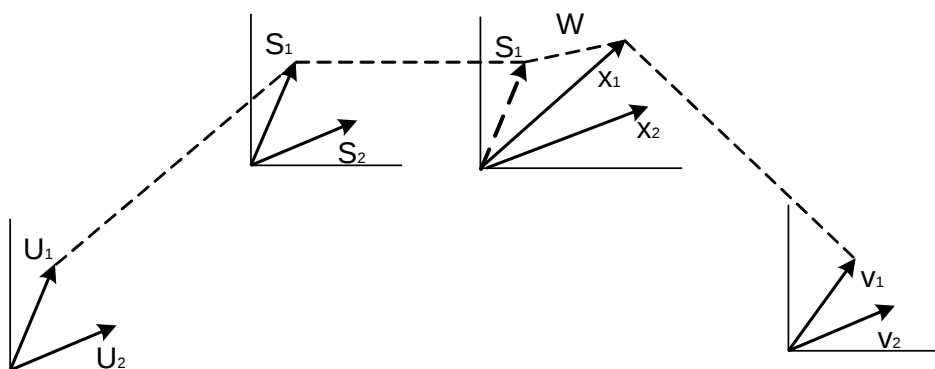


Рис.1.14. Геометрична модель системи зв'язку

РОЗДІЛ 2. ВИКОРИСТАННЯ ШУМОПОДІБНИХ СИГНАЛІВ У СИСТЕМАХ ЗВ'ЯЗКУ

2.1. Основні поняття визначення

Системи зв'язку із шумоподібними сигналами (ШПС) відомі понад 40 років. За цей час їх переваги стали очевидними, а багато з недоліків усунуто. В даний час системи зв'язку з ШПС набувають все більш широкого поширення. Процес розширення галузей використання систем зв'язку з ШПС незворотний, і найближчим часом увагу до них посилюватиметься.

Основу теорії систем зв'язку з ШПС заклали роботи В.О. Котельникова [1] та К.Шеннон [3], а основи кодового поділу – робота Д. В. Агєєва.

Системи зв'язку з ШПС займають особливе місце серед різних систем зв'язку, що зумовлено їх властивостями. По-перше, вони мають високу завадозахищеність при дії потужних завад. По-друге, забезпечують кодову адресацію великої кількості абонентів та їх кодовий поділ під час роботи у загальній смузі частот. По-третє, вони забезпечують сумісність прийому інформації з високою достовірністю та вимірювання параметрів руху об'єкта з високими точностями та здібностями. Всі ці властивості систем зв'язку з ШПС були відомі давно, але, оскільки потужності завад були відносно невисокі, а елементна база не дозволяла реалізувати пристрої формування та обробки в прийнятних габаритах, то довгий час системи зв'язку з ШПС широкого розвитку не набували. На даний момент становище різко змінилося. Потужність завад на вході приймача може на кілька порядків перевищувати потужність корисного сигналу. Для забезпечення високої завадозахищеності необхідно використовувати системи з ШПС із надвеликими базами (десятки-сотні тисяч), а ансамблі (системи) сигналів повинні складатися з десятків — сотень таких ШПС. Слід зазначити, що основи теорії ШПС із надвеликими базами сформувалися лише останнім часом. У свою чергу реалізація пристроїв формування та обробки ШПС стає можливою в найближчому майбутньому завдяки бурхливому розвитку надвеликих інтегральних схем, спеціалізованих мікропроцесорів, приладів з поверхневими акустичними хвилями, приладів із

зарядовим зв'язком. Всі ці причини викликали новий період розквіту систем зв'язку з ШПС.

Крім застосування в умовах впливу потужних завад, системи зв'язку з ШПС починають успішно конкурувати з широко розповсюдженими системами зв'язку з частотною модуляцією і частотним поділом каналів, що застосовуються для зв'язку з рухомими об'єктами. Створення малогабаритних пристроїв формування та обробки ШПС з базами $10^2 \dots 10^3$ дозволить широко впровадити ШПС у системи зв'язку і керування рухомими об'єктами у великих містах.

З цих причин до систем зв'язку з ШПС привернуто серйозну увагу великого кола науково-технічних працівників, розробників заводо захищених систем зв'язку, про що свідчить зростання публікацій у періодичній пресі з систем зв'язку з ШПС та суміжних питань.

Визначення шумоподібних сигналів широкосмугових систем зв'язку. Шумоподібними сигналами (ШПС) називають такі сигнали, добуток ширини спектра F яких на тривалість T набагато більше за одиницю. Цей добуток називається *базою сигналу* і позначається B :

$$B=FT. \quad (2.1)$$

Шумоподібні сигнали, база яких $B \gg 1$, іноді називають *складними* на відміну від *простих* сигналів з базою $B \approx 1$. Оскільки спектр сигналів з обмеженою тривалістю має необмежену довжину, то для визначення ширини спектра використовують різні методи і прийоми. Для реальних ШПС, що складаються з кінцевого числа елементів, завжди можна однозначно визначити F і B .

У системах зв'язку з ШПС ширина спектра ШПС (F) завжди набагато більша за ширину спектра повідомлення, що передається. У цифрових системах зв'язку, що передають інформацію у вигляді двійкових символів, тривалість ШПС і швидкість передавання інформації R пов'язані співвідношенням $T=1/R$.

$$B=2F/R \quad (2.2)$$

Тому база ШПС характеризує розширення спектра сигналу щодо спектра повідомлення. У аналогових системах зв'язку, верхня частота повідомлення яких дорівнює W , а частота відліку дорівнює $2W$,

$$B = F(2W) \quad (2.3)$$

І якщо $B \gg 1$, то $F \gg R$ і $F \gg 2W$. Саме тому системи зв'язку з ШПС у зарубіжній літературі одержали назву *системи зв'язку з розширеним (або розподіленим) спектром*, а у вітчизняній літературі — *широкосмугові системи зв'язку*. Надалі термін «широкосмугові системи зв'язку» (ШСЗ) буде відноситься тільки до систем зв'язку з ШПС.

Шумоподібні сигнали отримали застосування в широкосмугових системах зв'язку, так як: забезпечують високу завадозахищеність систем зв'язку; дозволяють організувати одночасно роботу багатьох абонентів у загальній смузі частот за асинхронно-адресним принципом роботи системи зв'язку, заснованому на кодовому поділі абонентів; дозволяють успішно долати багатопроменеве поширення радіохвиль шляхом поділу променів; забезпечують сумісність передавання інформації з вимірюванням параметрів руху об'єкта в системах рухомого зв'язку; забезпечують електромагнітну сумісність (ЕМС) широкосмугових систем з вузькосмуговими системами радіозв'язку та радіомовлення, системами телебачення, забезпечують найкраще використання спектра частот на обмеженій території в порівнянні з вузькосмуговими системами зв'язку.

Завадозахищеність - це здатність системи зв'язку протидіяти впливу потужних завад. Завадозахищеність містить у собі скритність системи зв'язку та її завадостійкість, так як для створення потужних завад треба спочатку виявити систему зв'язку та виміряти основні параметри її сигналів, а потім організувати потужну, найбільш сильну діючу заваду. Чим вище скритність і завадостійкість, тим вище завадозахищеність системи зв'язку.

Завадостійкість ШСЗ визначається виразом, який пов'язує відношення сигнал-завада на виході приймача (на виході погодженого фільтра або корелятора) q^2 відношенням сигнал-завада на вході приймача ρ^2 :

$$q^2 = 2B\rho^2 \quad (2.4)$$

де B - база ШПС;

$$\rho^2 = P_c/P_3 \quad (2.5)$$

де P_c, P_3 – потужності ШПС та завади;

$$q^2 = 2E/N_3 \quad (2.6)$$

E - Енергія ШПС; N_3 -спектральна щільність потужності завади в смузі ШПС;

$$E = P_c T; \quad (2.7)$$

$$N_3 = P_3 F; \quad (2.8)$$

Відношення сигнал-завада на *виході* q^2 визначає робочі характеристики приймання ШПС, а відношення сигнал-завада на *вході* ρ^2 - енергетику сигналу і завади. Величина q^2 може бути отримана згідно з вимогами до системи (10...30 дБ), навіть якщо $\rho^2 \ll 1$. Для цього достатньо вибрати ШПС із базою B , що задовольняє (2.4). Як видно зі співвідношення (2.4) приймання ШПС погодженим фільтром або корелятором супроводжується посиленням сигналу (або придушенням завади) у 2В рази.

$$K_{\text{ШПС}} = q^2/\rho^2 \quad (2.9)$$

Саме тому величину називають *коефіцієнтом посилення ШПС при обробці* чи *посиленням обробки*. Із (2.4), (2.9) випливає, що посилення обробки $K_{\text{ШПС}} = 2B$. У ШСЗ приймання інформації характеризується відношенням сигнал-завада $h^2 = \frac{q^2}{2}$, або

$$h^2 = B + \rho^2 \quad (2.10)$$

На рис (2.1 а) представлені залежності посилення обробки і бази ШПС від відношення сигнал-завада на вході ρ^2 дБ при значеннях q^2 (суцільні лінії) і h^2 (штрихові лінії), що дорівнюють 10, 20 і 30 дБ. Ці залежності побудовані згідно (2.4), (2.6). Наприклад, якщо необхідно отримати $h^2 = 20$ дБ, а на вході приймача $\rho^2 = -40$ дБ, то потрібна база має дорівнювати 60 дБ, т. є. $B = 106$.

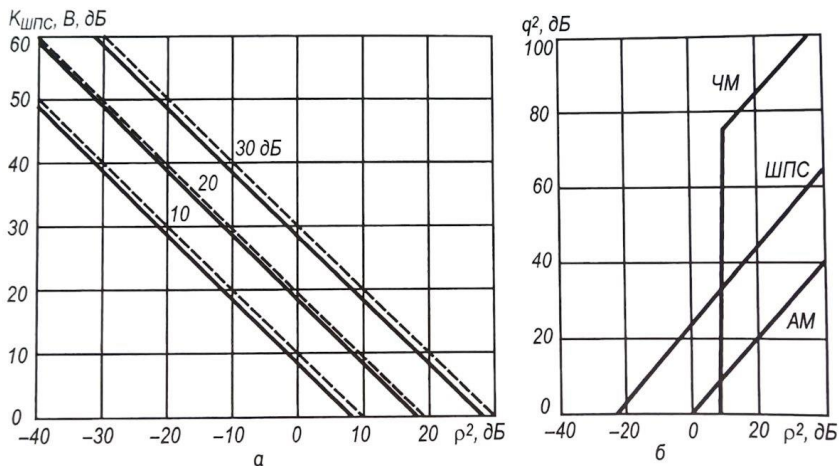


Рис. 2.1. Залежності посилення обробки та бази ШПС від відношення сигнал-завада на виході приймача (а) та завадостійкості систем зв'язку з ШПС (б)

Співвідношення (2.4), (2.6) є фундаментальними в теорії систем зв'язку з ШПС. Вони отримані для завад у вигляді білого шуму з рівномірною спектральною щільністю потужності в межах смуги частот, ширина якої дорівнює ширині спектра ШПС. Разом з тим ці співвідношення слушні для широкого кола завад (вузькосмугових, імпульсних, структурних), що визначає їх фундаментальне значення.

$$K_{шпс} \approx 2B \quad (2.11)$$

Загалом, посилення обробки ШПС для довільних завад де ступінь наближення залежить як від вигляду завад, так і від бази ШПС. У табл. 2.1 наведено значення посилення обробки деяких систем зв'язку і навігації.

У табл. 2.1 введені позначення: *ФМ – фазоманіпульований сигнал, ЧС – частотно-маніпульований сигнал.

Таблиця 2.1.

Параметри систем зв'язку з ШПС

Система зв'язку	Тип ШПС*	Ширина спектру, МГц	Посилення обробки, дБ
RAKE	ФМ	0,01	33
CHEROKEE	ФМ	1	16
RACEP	ЧС	4	-
MAGNAVOX	ФМ	1	35
GPS	ФМ	35	47

Наведені в перших чотирьох рядках таблиці параметри відповідають системам зв'язку 1960-х років і лише в п'ятому рядку наведено параметри сучасної системи GPS (Global Position System) - багатосупутникової радіонавігаційної системи.

На рис. 2.1 б наведені графіки завадостійкості систем зв'язку з ШПС і з частотною модуляцією (ЧМ) і з амплітудною модуляцією (АМ). Для порівняння ЧС та ШПС узято однакові смуги частот, що відповідає $B = 100$. Завадостійкість системи зв'язку з ШПС розрахована згідно з (2.4), причому покладено, що інформація передається за допомогою широтно-імпульсної модуляції (ШІМ). Відомо, ЧС має високу завадостійкість і забезпечує високу якість відтворення інформації за умови, що відношення сигнал-завада на вході вище за граничні значення $\rho^2 = 10 \dots 15$ дБ. Якщо ρ^2 нижче за граничні значення, завадостійкість системи зв'язку з ЧМ різко падає. Система з АМ та еквівалентною базою $B=1$ працює лише за $\rho^2 > 0$, залежність q^2 від ρ^2 лінійна. Система зв'язку з ШПС забезпечує надійне приймання інформації і при $\rho^2 < 0$. Наприклад, якщо взяти $q^2 = 10$ дБ, то система зв'язку буде працювати при відношенні сигнал-завада на вході -13 дБ, тобто $\rho^2 = 0,05$. Таким чином, одним з основних призначень систем зв'язку з ШПС є забезпечення надійного приймання інформації при впливі потужних завад, коли відношення сигнал-завада на вході приймача ρ^2 може бути набагато менше за одиницю.

Необхідно ще раз відзначити, що наведені співвідношення суворо слушні для завад у вигляді випадкового процесу з рівномірною спектральною щільністю потужності («білий» шум).

Скритність системи зв'язку - це здатність протистояти виявленню та виміру параметрів. Скритність - поняття дуже об'ємне, тому що включає в себе безліч особливостей виявлення ШПС і вимірювання їх параметрів. Оскільки виявлення ШПС і вимірювання параметрів можливі при різній первісній обізнаності (за апріорної невизначеності) про систему зв'язку, то можна вказати тільки основні співвідношення, що характеризують скритність. У випадку, коли відомо, що в даному діапазоні частот може працювати система зв'язку, але параметри її невідомі, то в цьому випадку можна говорити про *енергетичну* скритність системи зв'язку, оскільки її виявлення можливе за допомогою аналізу спектра (енергетичне виявлення). Характеристика виявлення (ймовірності помилкової тривоги і пропуску сигналу) цілком визначається відношенням сигнал-завада на вході приймача-аналізатора $= P_c / P_z$, де завадою є власний шум приймача $P_z = k_0 T_0 (N_{ш} - 1) F$, а k_0 -стала Больцмана, T_0 – температура навколишнього середовища, $N_{ш}$ - коефіцієнт шуму приймача. Час виявлення ШПС, за умови $\rho^2 \ll 1$ наближено визначається співвідношенням:

$$T_B \approx aF$$

$$(2.12)$$

де розмірна стала a залежить як від шумових властивостей приймача, потужності сигналу на вході, так і від необхідного відношення сигнал-завада на виході q^2 . Таким чином, чим ширший спектр ШПС, тим більший час виявлення та вища енергетична скритність системи зв'язку.

Якщо ШПС системи зв'язку відтворюються приймачем-аналізатором з високою завадозахищеністю, то час аналізу приблизно визначається співвідношенням, аналогічним за виглядом співвідношення (2.12) але $a = Tb \left[k_0 T_0 \frac{N_m - 1}{P_c} \right]$, b – стала величина, T – тривалість сигналу. Чим ширше спектр ШПС, тим більше база та час аналізу, тим вище параметрична скритність системи зв'язку.

Таким чином, чим ширший спектр ШПС і чим більша його база, тим вищою є як енергетична, так і параметрична скритність. Для боротьби з радіорозвідкою у завадозахищених системах зв'язку застосовують також зміну ШПС. Частота зміни ШПС, їх вибір із деякого ансамблю (системи сигналів) визначається багатьма вимогами до системи зв'язку і не можуть бути однозначними. Проте вважають, що число сигналів у системі (або об'єм системи сигналів) має бути набагато перевищувати базу ШПС. Можна припустити, що для завадозахищених систем зв'язку обсяг системи сигналів L визначається статичним законом:

$$L \approx B^m \quad (2.13)$$

де m — деяке число, що принаймні задовольняє умові $m \geq 2$, хоча для роботи може використовуватися набагато менше ШПС.

Отже, використання ШПС підвищує завадостійкість і скритність системи зв'язку, тобто її завадозахищеність. Застосування ШПС у завадозахищених системах зв'язку буде розширюватися.

2.2. Кодовий поділ абонентів

Завадозахищені системи зв'язку є спеціальними, некомерційними. Тому на ранньому етапі розвитку систем зв'язку із ШПС не передбачалось, що вони знайдуть широке застосування у комерційних системах зв'язку. Однак з розвитком асинхронних адресних систем зв'язку впровадження ШПС у системи масового радіозв'язку стало можливим. Основу для цього представляє кодовий поділ абонентів за допомогою ШПС, що відрізняються за формою.

При великих базах можна побудувати велику кількість різних ШПС. Наприклад, нехай ШПС є фазоманіпульованим сигналом, що складається з радіоімпульсів, фази яких 0 або π , а число їх дорівнює B . Можна побудувати множину сигналів (так званий повний код), число сигналів в якому дорівнює 2^B , і сигнали між собою відрізняються хоча б в одному імпульсі. Якщо взяти $B = 100$, то маємо $2^{100} \approx 10^{30}$ різних сигналів. З такої великої множини можна

відібрати систему сигналів так, щоб кожному абоненту в системі зв'язку відповідали свої власні сигнали. При цьому всі абоненти можуть працювати у загальній смузі частот, а поділ їх можливий за рахунок відмінності ШПС за формою. Такий поділ абонентів називається *кодовим*. При цьому ШПС є по суті адресою абонента і в цьому випадку принципово немає необхідності в примусовій часовій синхронізації абонентів. Тому подібні системи зв'язку одержали назву *асинхронних адресних систем зв'язку* (ААСЗ). Вони засновані на застосуванні ШПС і кодовому поділі абонентів.

В ААСЗ всі абоненти працюють у загальній смузі частот, тому при передаванні інформації ШПС різних абонентів перекриваються за часом і частотою, створюючи взаємні завади. Однак, при використанні ШПС з великими базами, можна звести рівень взаємних завад до необхідного, щоб забезпечити необхідну якість приймання інформації. Якщо припустити, що на вході одного з приймачів системи зв'язку діє l ШПС з однаковими потужностями, то відношення сигнал-завада на виході приймача

$$h^2 = B/l \quad (2.14)$$

Таким чином, збільшуючи базу ШПС, завжди можна досягти необхідної якості приймання інформації.

На рис. 2.2 подано залежності бази ШПС від числа активних абонентів, побудовані згідно (2.14). Графіки рис. 2.2 дозволяють визначити завадостійкість ААСЗ.

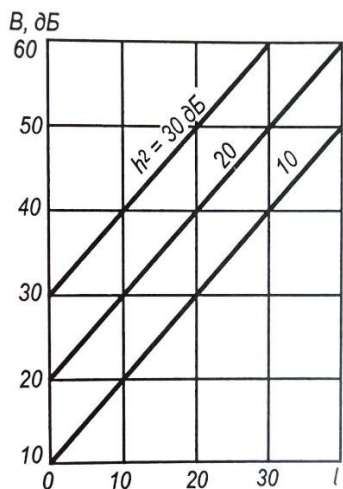


Рис. 2.2. Перешкодостійкість AACCC

Ефективність AAC3 Як випливає з (2.14), підвищення завадостійкості AAC3 при заданому числі активних абонентів можливе тільки за рахунок збільшення бази ШПС. При заданій швидкості передавання інформації збільшення бази потребує пропорційного розширення спектра ШПС відповідно до визначення (2.2). Виникає питання про ефективність використання радіоспектру у системах зв'язку із ШПС.

У системах зв'язку ефективність використання радіоспектру характеризується *питомою щільністю активних абонентів* γ , що дорівнює кількості активних абонентів, що припадають на 1 МГц смуги частот, тобто.

$$\gamma = l/F$$

$$(2.15)$$

де l — число активних абонентів, що одночасно працюють у смузі частот шириною F . Заміняючи l у (2.15) згідно з (2.14), маємо

$$\gamma = T/h^2 \quad (2.16)$$

Часто питома щільність активних абонентів називається просто *ефективністю системи зв'язку*. З (2.16) випливає, що ефективність ААСЗ тим менша, чим більше відношення сигнал-завада на виході приймача. Таким чином, ААСЗ більш перспективні у тих випадках, коли не потрібна висока якість передавання інформації, що характерно для систем масового радіозв'язку. Наприклад, у радіотелефонній системі рухомого зв'язку з частотним поділом каналів максимальна ефективність $\gamma_{max} = 250$ аб./МГц, оскільки мінімальна ширина кожного каналу дорівнює 4 кГц, і на 1 МГц можна розмістити 250 частотних каналів, тобто активних абонентів. Однак, для підвищення завадостійкості використовується частотна модуляція (ЧМ), і сусідні частотні канали рознесені на 25 кГц. При цьому на 1 МГц можна розмістити 40 частотних каналів, тобто $\gamma_{чм} = 40$ аб./МГц. Перші системи зв'язку з ШПС («RADA», «RACER» та ін.) мали низьку ефективність, з $\gamma_{шпс} \approx 7.9$ аб./МГц. Із (2.16) випливає, що високу ефективність систем зв'язку із ШПС безпосередньо отримати важко. Наприклад, якщо взяти $T=1/4$ кГц = 250 мс, а $h^2 = 10$, то $\gamma_{шпс} = 25$ аб./МГц, тобто нижче за ефективність систем зв'язку з ЧМ.

Останніми роками запропоновано інший принцип побудови радіотелефонних систем рухомого зв'язку. Вся територія, що обслуговується, розбивається на велику кількість зон у вигляді стільників. У кожній зоні радіозв'язок ведеться на частотах, спеціально виділених цій зоні. За рахунок територіального рознесення зон з однаковими частотними каналами можливе багаторазове використання тих самих частотних каналів. Такі системи зв'язку одержали назву *стільникових систем рухомого зв'язку* (ССРЗ). Приймання сигналів у таких системах принципово супроводжується взаємними завадами так само, як і в ААСЗ. Тому застосування ШПС у ССРЗ є перспективним, оскільки дозволяє успішно долати взаємні завади. Ефективність ССРЗ

$$\gamma_{ССРЗ} \approx 3,63(R_0/D)^2/F_k \quad (2.17)$$

де R_0 – радіус зони обслуговування, D – захисний інтервал, F_k – ширина частотного каналу. Якщо взяти $R_0 = 30$ км, $D \approx 4,4$ км, а $F_k = 50$ кГц, то $\gamma_{ССРЗ} \approx 3\,333$ аб./МГц, тобто ефективність ССРЗ набагато вище за ефективність звичайних систем рухомого зв'язку. Якщо ширину частотного каналу збільшити до

100...200 кГц, то ефективність ССРЗ дорівнюватиме 1 666...833 аб./МГц відповідно, проте все одно набагато перевищуватиме ефективність систем із ЧМ. При цьому можливе застосування ШПС з відносно невеликими базами (25...250), що дозволить використовувати просту апаратуру формування та обробки ШПС з невисокою стабільністю частоти. У свою чергу, застосування ШПС дозволить успішно вирішити проблему адресації великої кількості абонентів. Стільникові системи рухомого зв'язку із ШПС дозволять забезпечити зв'язком від 60 до 340 тисяч рухомих абонентів у великих містах. Крім того, в таких системах можна поєднати передавання телефонних повідомлень з визначенням розташування рухомих об'єктів і забезпеченням систем захисту.

Боротьба з багатопроменевістю. Застосування ШПС у системах зв'язку дозволяє боротися з багатопроменевістю поширення радіохвиль. Багатопроменевість виникає в тому випадку, якщо радіохвилі приходять в точку прийому, відбившись від різних завад на шляху розповсюдження (шари іоносфери, будівлі, пагорби тощо). Через розбіжності в довжині шляху ці радіохвилі приходять з різним запізненням, внаслідок чого сигнали, що прийшли різними шляхами, перекриваються в часі і між ними виникає інтерференція, яка в свою чергу спричинює глибокі завмирання результуючого сигналу. Зазвичай компенсації замирань передбачають збільшення потужності сигналу на 20 дБ. Інша справа- використання ШПС, оскільки при обробці ШПС погодженим фільтром, відбувається стиснення ШПС за часом, що ілюструється рис.2.3. На рис. (2.3, а) зображено ШПС з частотною модуляцією тривалістю T , на рис. (2.3, б) зображена залежність між напруга на виході погодженого фільтра та відкликом фільтра на ШПС.

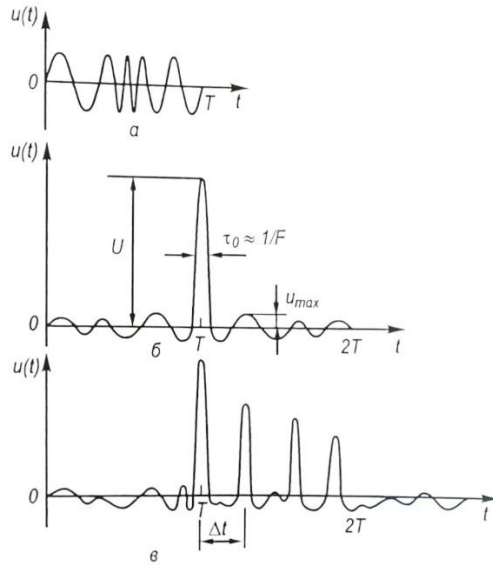


Рис. 2.3. шумоподібний сигнал (а), автокореляційна (б) та поділ променів (в)

Цей відклик називається *автокореляційною функцією* (АКФ) ШПС. Хоча АКФ має тривалість $2T$, в ній можна виділити дві області, які різко відрізняються одна від одної. Перша область знаходиться в центрі АКФ; різкий викид у вигляді вузького імпульсу, званого *центральною піком*. Його амплітуда дорівнює U , а тривалість

$$\tau_0 \approx 1/F \quad (2.18)$$

Чим ширший спектр ШПС, тим коротший центральний пік. Другу область складають бічні піки з максимальним значенням $u_{\max}$. Шумоподібні сигнали з великими базами мають властивості, які записуються двома співвідношеннями:

$$T/\tau_0 \approx B$$

$$(2.19)$$

$$U/(\gamma u_{\max}) \approx \sqrt{\alpha B}$$

(2.20)

де α - деяка стала, в загальному випадку залежна від бази B .

Співвідношення (2.19) визначає стиск ШПС - відношення тривалості ШПС T до тривалості центрального піку. Стиск ШПС дорівнює, приблизно, базі. Тому при $T = \text{const}$ із збільшення F зменшується тривалість центрального піку τ_0 і зростає стиск.

Співвідношення (2.20) характеризує придушення бічних піків. Воно дорівнює відношенню амплітуди центрального піку U до амплітуди максимального бічного піка $u_{\max}$. Чим більша база, тим більше придушення бічних піків. Границя АКФ ШПС із зростанням бази прямує до вузького дельта-імпульсу. Таку АКФ має широкосмуговий шум, що і спричинило назву — «шумоподібні сигнали».

На рис. (2.3 в), зображено відклик погодженого фільтра на три ШПС, що надійшли по різних шляхах. Якщо затримка між променями Δt більша за тривалість центрального піку τ_0 , то промені поділяються і центральні піки різних променів можна розділити один від одного, а потім об'єднати, усунувши затримку між ними. Такий принцип боротьби з багатопроменеві був використаний в одній з перших систем зв'язку з ШПС «RAKE». Таким чином, умова $\Delta t > \tau_0$ забезпечує поділ променів. Оскільки τ_0 і F пов'язані співвідношенням (2.18), то умова поділу променів записується наступним чином:

$$F \Delta t > 1$$

(2.21)

Наприклад, якщо при поширенні радіохвиль існують два промені - прямий і відбитий від деякого об'єкта, то затримка $\Delta t \approx 2d^2/(Rc)$, де c - швидкість світла, R - відстань між передавачем і приймачем, d - відстань між відбиваючим об'єктом, і прямим променем. У цьому випадку необхідно використовувати ШПС із шириною спектра

$$F \geq R_c/(2d^2) \quad (3.22)$$

Чим більше d , тим менше F , а тому при малих d можуть знадобитися ШПС з дуже широкими спектрами, що не завжди можна реалізувати на практиці.

Вимір координат рухомих об'єктів і електромагнітна сумісність.

Застосування ШПС дозволяє поєднати системи передавання інформації з системами траскторних вимірів. Серед параметрів руху об'єкта найбільший інтерес становлять відстань між приймачем та передавачем та їх відносна швидкість. Відстань вимірюється за затримкою сигналу в часі, а швидкість – за доплерівськи зсувом частоти. Точність виміру і роздільна здатність за затримкою визначаються відношенням сигнал-завада q^2 (2.4) і шириною спектра сигналу і характеризуються похибкою

$$G(t) \approx 1/(qF)$$

(2.23)

Чим більше q і F , тим менша похибка виміру затримки, тим вища точність виміру і роздільна здатність за відстанню. Точність виміру доплерівського зсуву частоти визначається відношенням сигнал-завада q^2 та тривалістю сигналу та характеризується похибкою

$$\sigma_f \approx 1/(qT)$$

(2.24)

Чим більше q і F , тим менша похибка виміру доплерівського зсуву частоти, тим вища точність виміру і роздільна здатність за швидкістю. Із (2.23), (2.24) випливає, що при сумісному вимірі відстані та швидкості необхідно використовувати ШПС, так як тільки для ШПС можна незалежно змінювати ширину спектра F та тривалість T . У системах зв'язку тривалість T зазвичай визначається швидкістю передавання інформації. Тому підвищення точності виміру відстані можна досягти розширенням спектра F , тобто використовуючи ШПС.

Шумоподібні сигнали забезпечують добру електромагнітну сумісність (ЕМС) ширококугових систем з вузькосмуговими системами радіозв'язку та мовлення. На рис. (2.4, а) зображені спектри ширококугової і вузькосмугової системи зв'язку з ШПС.

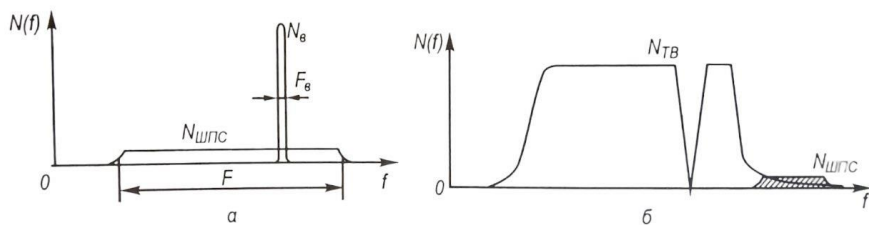


Рис. 2.4. Спектри широкосмугової та вузькосмугової систем зв'язку (а)
та телевізійного сигналу і ШПС (б)

Відповідно для ШПС спектральна щільність потужності $N_{\text{шпс}} = P_{\text{шпс}} / F$, для вузькосмугового сигналу $N_{\text{в}} = P_{\text{в}} / F_{\text{в}}$. Завадостійкість системи зв'язку з ШПС визначається фундаментальним співвідношенням (2.4), в якому $\rho^2 = P_{\text{шпс}} / P_{\text{в}}$. Посилення обробки дорівнює $2B$. Якщо вузькосмугова система зв'язку постійно займає визначений інтервал, то можна її спектр ціліком придушити, використовуючи режекторний фільтр, налаштований на частоту вузькосмугової системи зв'язку. Таким чином, вплив вузькосмугової системи зв'язку на широкосмуговий незначний. У свою чергу широкосмугова система зв'язку також слабо впливає на вузькосмугову систему зв'язку. Потужність ШПС, що проходить на вихід приймача, $N_{\text{шпс}} F_{\text{в}} = P_{\text{шпс}} F_{\text{в}} / F$. Тому відношення сигнал-завада на виході вузькосмугового приймача визначатиметься співвідношенням (2.4), в якому $\rho^2 = P_{\text{в}} / P_{\text{шпс}}$, $B = F / F_{\text{в}}$. Тому чим більше відношення $F / F_{\text{в}}$, тим краще фільтрація ШПС у вузькосмуговій системі зв'язку. Отже, що більше база ШПС, тим вище ЕМС широкосмугової і вузькосмугової систем зв'язку.

Системи зв'язку з ШПС можна інтегрувати і з радіотелевізійними системами. На рис. (2.3 б) зображений спектр телевізійного сигналу $N_{\text{ТВ}}$. Програми телебачення в одній територіальній зоні передаються по кількох каналах з великими захисними частотними інтервалами. Зазвичай в цих частотних захисних інтервалах заборонена робота будь-яких радіотехнічних систем, щоб не створювати завад телевізійним передачам. Однак в цих частотних інтервалах можна розмістити системи зв'язку з ШПС так, як це показано на рис. (2.4 а). Спектр ШПС розташований поблизу спектру телевізійного сигналу, там, де спектральна щільність останнього різко зменшується.

При цьому взаємні завади обох систем будуть малі. Слід зазначити, що у разі використання сигналів з частотної модуляцією (замість ШПС) рівень взаємних завад зростає, оскільки сигнали системи зв'язку й телебачення належать до однакового класу і демодулюються частотним детектором.

Таким чином, системи зв'язку з ШПС мають добру ЕМС із системами радіозв'язку, мовлення та телебачення. Раніше було згадано, що ШПС забезпечують високу ефективність використання радіоспектрів стільникових системах рухомого зв'язку. Якщо розглядати дію систем зв'язку в деякому замкнутому просторі, то виявляється, що найкращу ЕМС в обмеженому діапазоні частот забезпечують ШПС, хоча самі по собі вони вимагають ширшої смуги, ніж традиційні вузькосмугові системи. Водночас, загальна смуга частот при використанні ШПС буде меншою.

З розгляду основних властивостей ШПС випливає, що застосування ШПС в системах зв'язку дозволяє забезпечувати високу завадостійкість, скритність, адресність, працездатність у загальній смузі частот, подолання багатопроменевості, високі точності вимірів і роздільні здатності, добру ЕМС з багатьма радіотехнічними системами. Ці переваги надіються застосування ШПС з великими базами, що призводить до різкого ускладнення пристроїв формування та обробки, збільшення їх маси, об'єму споживаної потужності. Найчастіше перехід до ШПС з великими базами вимагає різкого розширення смуги частот, що зумовлює певні труднощі у створенні широкосмугової елементної бази. Однак ці труднощі переборні. І тому ШПС знаходять вже зараз застосування у різних системах зв'язку, а в майбутньому отримають ще ширше застосування, особливо у системах масового радіозв'язку.

2.3. Основні структурні схеми широкосмугових систем зв'язку ШПС

Широкасмугові системи зв'язку з ШПС залежно від призначення, технічних характеристик, бази ШПС, елементної бази можуть бути побудовані за різними схемами, перерахувати які в даний час неможливо через численні варіанти. Для

якісного уявлення про те, з яких основних пристроїв складаються ШСС, на рис. 2.5 – 2.8 наведено структурні схеми деяких систем зв'язку.

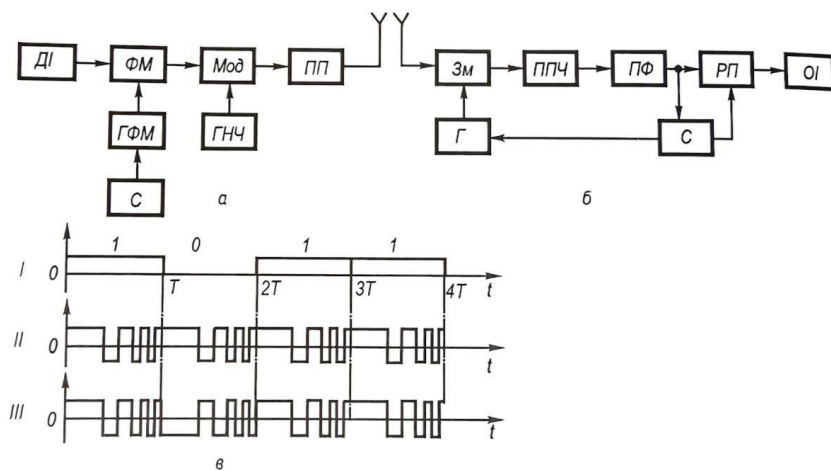


Рис. 2.5. Структурна схеми передавача (а) і приймача (б) цифрової систем зв'язку з фазоманіпульованим ШПС та модуляція цифрової інформації (в)

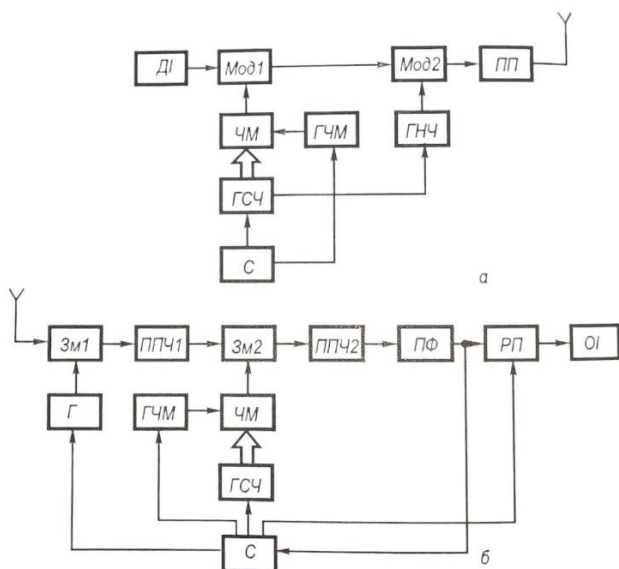


Рис. 2.7. Структурні схеми передавача (а) та приймача (б) цифрової системи зв'язку за частотно-маніюльованим ШПС

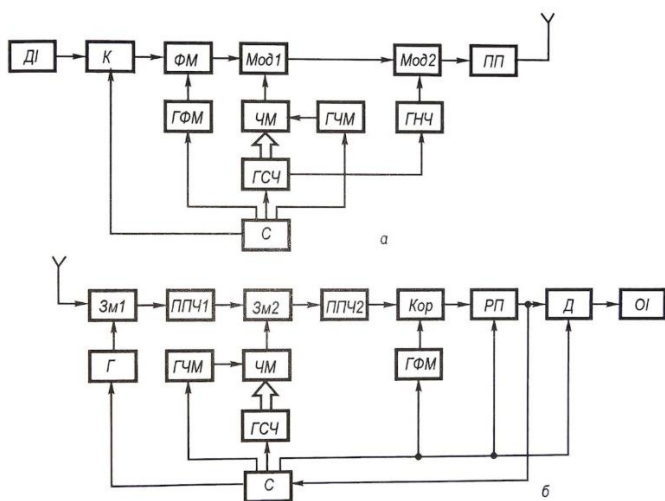


Рис. 2.8. Структурна схеми передавача (а) та приймача (б) цифрової системи зв'язку за фазочастотно-маніюльованим ШПС та коректувальними кодами

На рис. 3.5 представлені структурні схеми передавача і приймача цифрової системи зв'язку з фазоманіпульованим (ФМ) сигналом, призначена для передавання дискретних повідомлень. У передавачі (рис. 2.5 а) від джерела інформації *ДІ* послідовність двійкових одиниць 1 і 0 зі швидкістю $R=1/T$ (рис. 2.5, в, *I*) надходить на перший вхід фазового модулятора ФМ. На другий вхід ФМ надходить фазоманіпульований сигнал (рис. 2.5, в, *II*) від генератора фазоманіпульованого сигналу *ГФМ*. Фазоманіпульований сигнал (ФМ-сигнал) має тривалість T і являє собою послідовністю відеоімпульсів 1 і 0 тривалістю $\tau_0 = T/N$ де N - число імпульсів. На рис. 2.5, в $N = 13$. Зазвичай вважають, що база ФМ-сигналу дорівнює числу імпульсів, тобто $B \approx N$. Ширина спектра ФМ-сигналу $F \approx 1/\tau_0$. Роботою ГФМ керує синхронізатор *С*, що формує необхідні сигнали керування і частоти. Послідовність ШПС у вигляді ФМ-сигналів, які переносять інформаційні символи (рис. 2.5, в, *III*), надходить у модулятор *Мод*, де здійснюється балансова модуляція коливання з несучою частотою ФМ-сигналом. Коливання з несучою частотою створюються генератором низької частоти *ГНЧ*. Підсилювач потужності *ПП* посилює ФМ-сигналом, який потім через антену випромінюється в простір.

У приймачі (рис.2.5,б) сигнал проходить через змішувач *Зм*, переноситься за допомогою гетеродина *Г* на проміжну частоту, посилюється в підсилювачі проміжної частоти *ППЧ* і обробляється погодженим фільтром *ПФ*. Сигнал з виходу *ПФ* надходить на синхронізатор *С* і розв'язувальний пристрій *РП*. Синхронізатор здійснює пошук ФМ-сигналу за частотою і часом, накопичує сигнал збільшення надійності синхронізації, керує режимом роботи розв'язувального пристрою. Для пошуку ФМ-сигналу за частотою синхронізатор перебудовує гетеродин. Після закінчення пошуку і набуття синхронізму на виході *РП* з'являється інформаційна послідовність у вигляді двійкових символів, що передається одержувачеві інформації *ОІ*. Приймач, зображений на (рис.2.5,б), є найпростішим. Разом з тим необхідно відзначити, що погоджений фільтр і синхронізатор, що містить блоки пошуку та синхронізації, у разі великих баз ШПС є складними пристроями. Крім того, для пошуку ШПС та підтримки синхронізму приймач охоплений петлею зворотного зв'язку. Реальний приймач ШПС може містити кілька блоків пошуку і спостереження, у тому числі блок пошуку ШПС за часом і часовою синхронізацією, блок фазового автопідстроювання частоти (ФАПЧ), які охоплені власними та взаємними зворотними зв'язками.

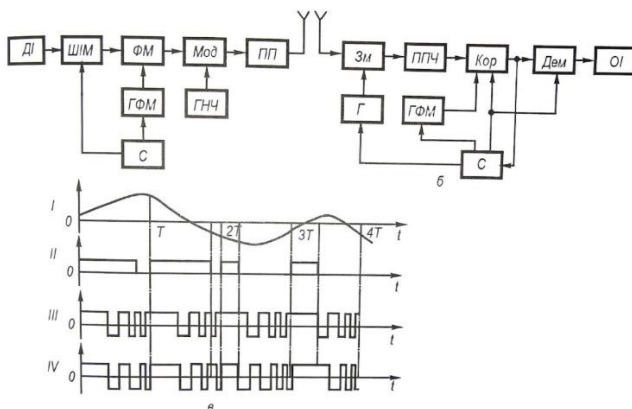


Рис. 2.6. Структурні схеми передавача (а) і приймача (б) радіотелефонної системи зв'язку з фазоманіпульованим ШПС та модуляція ФМ ШПС (в)

На рис. 2.6 представлені структурні схеми передавача та приймача радіотелефонної системи зв'язку з ФМ ШПС. У передавачі (рис. 2.6, а) телефонне повідомлення (рис. 2.6, в, I) від джерела інформації ДІ надходить на вхід широтно-імпульсного модулятора ШИМ, з виходу якого ШИМ-сигнал (рис. 2.6, в, II) подається на вхід фазового модулятора ФМ. На другий вхід ФМ подається фазоманіпульований ШПС (рис. 2.6, в, III) який формується ГФМ. Фазоманіпульований сигнал з виходу ФМ (рис. 2.6, в, IV), що містить інформацію, надходить на вхід модулятора Мод, в якому здійснюється балансна модуляція коливання з несучою частотою від генератора низької частоти ГНЧ. Потім посилений по потужності в підсилювачі потужності ПП ФМ-сигнал через антену випромінюється в простір. Роботою широтно-імпульсного модулятора ШИМ і генератора ФМ-сигналу ГФМ керує синхронізатор С, який виробляє необхідні частоти і керуючі сигнали.

У приймачі (рис. 2.6, б) прийнятий сигнал у змішувачі Зм за допомогою гетеродина Г переноситься на проміжну частоту і після підсилювача проміжної частоти ППЧ надходить на корелятор Кор. Корелятор, як і погоджений фільтр,

здійснює оптимальну обробку прийнятого сигналу. Хоча корелятор і погоджений фільтр відрізняються за принципом роботи, але забезпечують однакову завадостійкість прийому. Корелятор складається з перемножника й інтегратора. На другий вхід корелятора подається опорний сигнал у вигляді фазоманіпульованого ШПС (рис. 2.6, в, III). Напряга на виході корелятора містить телефонне повідомлення у вигляді ШІМ-сигналу, що подається на вхід демодулятора *Дем*, з виходу якого прийняте телефонне повідомлення передається одержувачу інформації *ОІ*.

Робота приймача в цілому та його окремих блоків (*Г*, *ГФМ*, *Кор*, *Дем*) керується блоком пошук ФМ ШПС за часом і частотою, а потім приймач підтримує синхронізм. Усе раніше сказане щодо синхронізатора приймача, зображеного на рис. 2.6, б, цілком стосується синхронізатора будь-якого приймача (рис. 2.7, б, 2.8, б).

На рис. 2.7 представлені структурні схеми передавача та приймача цифрової системи зв'язку з частотно-маніпульованим ШПС (іноді такий ШПС називають сигналом зі стрибуючою частотою). Відмінність передавача та приймача, зображених на рис. 2.7, від передавача і приймача на рис. 2.5 зводиться до наступного. У модуляторі *Mod1* передавача (рис. 2.7, а) проводиться модуляція частотно-маніпульованого ШПС (ЧМ ШПС) дискретним повідомленням. ЧМ ШПС являє собою сигнал, що складається з M імпульсів, несучі частоти яких набувають одне з можливих значень: від f_0 до $f_0 + (M - 1)/T$ з інтервалом між сусідніми значеннями $\Delta f = 1/T$. Усього використовується M частот, і жодна з них не застосовується двічі в одному ШПС. База такого сигналу $B \approx M^2$. ЧМ ШПС формується за допомогою частотного маніпулятора *ЧМ*, на перший вхід якого через шину (широка стрілка) подаються M частот від генератора сітки частот *ГСЧ*. На другий вхід подається кодова послідовність від генератора кодової послідовності ЧМ ШПС (*ГЧМ*), що визначає порядок зміни частот у ЧМ ШПС. У модуляторі *Mod2* ЧМ ШПС переноситься на несучу частоту. Роботою *ГСЧ*, *ГЧМ*, *ГНЧ* керує синхронізатор *С*.

У приймачі (рис. 2.7, б) ЧМ ШПС на проміжній частоті надходить на змішувач $Зм2$, де всі частоти сигналу переносяться на іншу проміжну частоту за допомогою опорного ЧМ ШПС, що надходить від частотного маніпулятора *ЧМ*. Призначення *ГСЧ* і *ГЧМ* таке ж саме, як *ГФМ* та *ГНЧ* у передавачі, що на рис. 2.5, а. З виходу ППЧ2

сигнал тривалістю T , що не має частотної маніпуляції, надходить на погоджений фільтр $ПФ$, а потім на розв'язувальний пристрій $РП$ і синхронізатор $С$. Останній шукає ЧС ШПС за часом і частотою, потім підтримує синхронізм і керує роботою Γ , ГСЧ, ГЧМ і РП.

На рис. 2.8 представлені структурні схеми передавача та приймача цифрової системи зв'язку з фазочастотно-маніпульованим ШПС. Фазочастотно-маніпульований ШПС (скорочено ФМ - ЧС ШПС) є складним. При певній подвійній маніпуляції він складається з N імпульсів, групи яких передаються на M частотах. База такого сигналу $B \approx N \cdot M^2$. Оскільки ФМ-ЧМ ШПС є об'єднанням ФМ- і ЧМ-сигналів, то і схема передавача (рис. 2.8, а), і схема приймача (2.8, б) теж, в свою чергу є об'єднанням схем передавачів, зображених на (рис. 2.5, а і 2.7, а) та приймачів (рис. 2.5, б і 2.7, б). Для додаткового підвищення завадостійкості використовуються коректувальні коди, які формуються в передавачі (рис. 2.8, а) за допомогою кодера K і декодуються в приймачі (рис. 2.8, б) за допомогою декодера D . У приймачі оптимальну фільтрацію здійснює корелятор $Кор$. Призначення інших блоків таке ж саме, як і в попередніх схемах.

Подані схеми не вичерпують усього різноманіття їх, але дозволяють виділити основні вузли широкосмугових систем зв'язку з ШПС. До таких вузлів належать генератори формування ШПС (або автомати формування ШПС з їх зміною), генератори сітки частот, погоджені фільтри, корелятори, блоки пошуку ШПС і синхронізації за часом і частотою. З викладеного матеріалу випливає, що розробник широкосмугової системи зв'язку має вміти вибрати тип ШПС і його базу, метод обробки, визначити час пошуку і синхронізації, знайти завадостійкість приймача ШПС у разі дії різного виду завад, визначити елементну базу, розробити для неї необхідні генератори ШПС, погоджені фільтри і корелятори, блоки пошуку і синхронізації. Крім цього, розробник повинен правильно спроектувати інші вузли передавача і приймача, знати, як проходить ШПС через вузли передавача та приймача і які втрати при цьому мають місце. На всі питання, що виникають у процесі проектування ШСЗ, не можна дати однозначні відповіді у більшості випадків. Тому проектування ШСЗ у даний час є інженерним мистецтвом, що ґрунтується на глибокому знанні теорії та техніки ШПС

та на інтуїції розробника. Проте з усіх питань проектування систем зв'язку з ШПС в даний час є основні (і в багатьох випадках фундаментальні) результати.

2.4. Шумоподібні сигнали та їх спектри.

Сигналом називається змінна фізична величина, що відображає повідомлення. Сигнал u , що є функцією часу t , записується як $u=u(t)$.

Множина сигналів $u_j(t)$, побудована за єдиним правилом, називається *системою сигналів*. Таким чином, система сигналів визначена, якщо відомо правило побудови сигналів. Номер сигналу позначається індексу j . Якщо в системі кількість сигналів L , то можна пронумерувати сигнали натуральними числами від 1 до L і позначити $j = 1, L$. Число L називається *об'ємом системи сигналів*.

Надалі розглядаються сигнали, які можна подати у такому вигляді:

$$u(t) = A(t)\cos[\omega_0 t + \theta(t)] \quad (2.25)$$

де $A(t)$ - обвідна сигналу, ω_0 - несуча частота, $\theta(t)$ - повільнозмінна частина фази сигналу.

Зображенню (2.25) відповідає радіочастотний сигнал. Так як розглядаються реальні сигнали (які можна сформувати та обробити), то всі функції часу та параметри правої частини (2.25) відомі.

Якщо сигнал заданий у загальному вигляді $u(t)$ і права частина (2.25) невідома, необхідно скористатися перетворенням Гільберта і знайти сполучений сигнал $\hat{u}(t)$. У цьому випадку огинаюча $A(t) = \sqrt{u^2(t) + \hat{u}^2(t)}$, фаза $\theta(t) = \omega_0 t + \theta(t) = \arctg[\hat{u}(t)/u(t)]$.

Якщо функція $\theta(t)$ безперервна і має безперервну першу похідну, то миттєва частота сигналу $\omega_0(t)$ за означенням дорівнює першій похідній фази $\theta(t)$, тобто $\omega(t) = \omega_0(t) + \dot{\theta}(t)$. Перетворення Гільберта:

$$u(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{u}(x)}{x-1} dx, \quad \hat{u}(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(x)}{x-1} dx$$

На рис. 2.9. *а* показаний фазоманіпульований сигнал (ФМ), що складається з чотирьох радіоімпульсів з однаковою несучою частотою, але з різними початковими фазами. На рис. 2.9. *б, в* зображено його обвідну $A(t)$ і нульову фазу $\theta(t)$. Обвідна є сталою величиною на інтервалі тривалістю T , а фаза дорівнює двом значенням: 0 або π .

Якщо несуча частота сигналу $\omega_0(t) = 0$, то такий сигнал є *відеочастотним*. На рис. 2.9. *г* зображений відеочастотний сигнал $u(t)$ - послідовність позитивних і негативних прямокутних імпульсів, - отриманий з ФМ-сигналу 2.9. *а* за умови, $\omega_0(t) = 0$. Оскільки знаки імпульсів відеочастотного сигналу визначаються початковими фазами імпульсів радіочастотного сигналу, то за аналогією з радіочастотним сигналом відеочастотний також називається *фазоманіпульованим* сигналом.

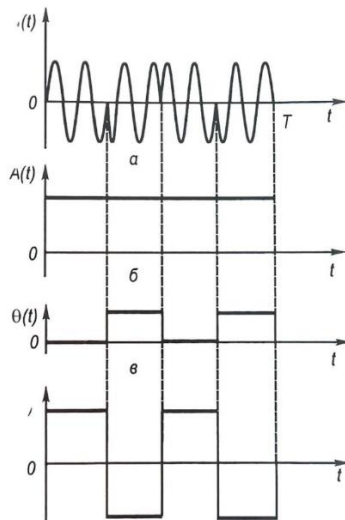


Рис. 2.9. Фазоманіпульований сигнал

Спектр сигналу визначається перетворенням Фур'є

$$g(\omega) = \int_{-\infty}^{\infty} u(t) e^{-i\omega t} dt \quad (2.26)$$

Спектр є функцією кутової частоти $\omega = 2\pi f$, де f - лінійна частота (надалі ω і f називаються просто частотою.) Нескінченні границі інтегрування відповідають загальному випадку. При визначенні спектра фінітного сигналу (з скінченною тривалістю) необхідно враховувати його розташування на осі часу t . Спектр може бути представлений у вигляді $g(\omega) = |g(\omega)| \exp[i\varphi(\omega)]$, де $|g(\omega)|$ - амплітудний, а $\varphi(\omega)$ - фазовий спектр сигналу.

Сигнал знаходиться за спектром оберненим перетворенням Фур'є

$$u(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\omega) e^{i\omega t} d\omega$$

Ширина спектра. Спектр фінітних сигналів має нескінченну довжину, тому єдиного визначення ширини спектра немає. Залежно від мети дослідження ширину спектра сигналу знаходять по-різному. Надалі ширина спектра визначатиметься так, щоб правильно відобразити суть розв'язуваної задачі. Такий підхід виправданий тим, що для сигналів, що входять в одну систему, будь-яке достатньо розумне визначення ширини спектра відбиватиме спектральні властивості кожного сигналу та системи сигналів в цілому. (Ширина спектра сигналу позначається F).

Комплексна обвідна сигналу та її спектр. Радіосигнал (2.25) містить швидко мінливий множник у вигляді косінусоїди, в аргумент якої входить несуча частота $\omega_0 = 2\pi f_0$. Відповідно спектр (2.26) цього сигналу складається з двох частотних смуг, зосереджених близько частот ω_0 і $-\omega_0$. При теоретичних дослідженнях доцільно для спрощення проміжних математичних операцій «звільнити» сигнал та його спектр від несучої частоти ω_0 . Це можна здійснити при введенні комплексної обвідної сигналу.

Комплексна обвідна радіосигналу (2.25) визначається як

$$u(t) = |u(t)| \exp[i\omega_0 t] \quad (2.27)$$

де модуль $|u(t)| = A(t)$ є обвідною сигналу $u(t)$. Перехід від комплексної обвідної до сигналу здійснюється за допомогою формули:

$$u(t) = \operatorname{Re}[U(t) \exp[i\omega_0 t]] \quad (2.28)$$

де Re – дійсна частина.

На рис. 2.9, зображена комплексна обвідна ФМ-сигналу, наведеного на рис. 2.9, *a*. Вона є послідовністю прямокутних відеоімпульсів і дійсною функцією часу. Це обумовлено тим, що початкові фази імпульсів ФМ-сигналу набувають одного з двох значень: 0 або π . У загальному випадку комплексна обвідна містить і дійсну, і уявну складові, але завжди є відеосигналом, чим і пояснюється перехід до неї від радіосигналу.

Спектр комплексної обвідної

$$G(\omega) = \int_{-\infty}^{\infty} U(t)e^{-i\omega t} dt \quad (2.29)$$

Комплексна обвідна сигналу знаходиться відповідно до оберненого перетворення Фур'є

$$U(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega)e^{i\omega t} d\omega \quad (2.30)$$

Спектр комплексної обвідної можна подати у вигляді

$$G(\omega) = |G(\omega)| \exp[i\Phi(\omega)],$$

де $|G(\omega)|$ - амплітудний, а $\Phi(\omega)$ - фазовий спектри.

Спектр сигналу $g(\omega)$ і спектр його обвідної комплексної $G(\omega)$ пов'язані співвідношенням $g(\omega) = 0,5G(\omega - \omega_0) + 0,5G^*(\omega - \omega_0)$, де $*$ - знак комплексної сполученості.

Оскільки комплексна обвідна $u(t)$ - відеосигнал, то спектр $G(\omega)$ розташований в області відеочастот.

На рис. 2.10 зображено спектр $G(\omega)$ комплексної обвідної $u(t)$ довільного сигналу.

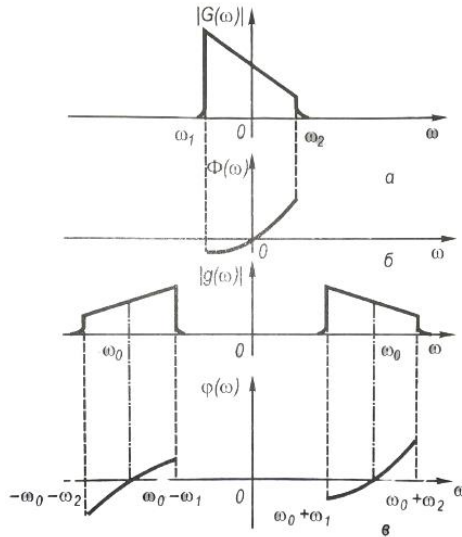


Рис. 2.10. *a* - амплітудний спектр $|G(\omega)|$ комплексної обвідної сигналу, *б* - фазовий спектр $\Phi(\omega)$, *в* - фазовий спектр сигналу.

Енергія сигналу та частотно-часова площа. За означенням, енергія сигналу

$$E = \int_{-\infty}^{\infty} u^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |g(\omega)|^2 d\omega \quad (2.31)$$

Для сигналів, у яких $|\theta'(t)|_{\max} \ll \omega_0$ енергія сигналу виражається через модулі комплексної обвідної та її спектру наступним чином:

$$E = \frac{1}{2} \int_{-\infty}^{\infty} |u(t)|^2 dt = \frac{1}{4\pi} \int_{-\infty}^{\infty} |G(\omega)|^2 d\omega \quad (2.32)$$

Зазвичай більшість енергії сигналу зосереджена у деякій смузі частот. Нехай F - ширина смуги частот, у середині якої зосереджена більша частина заданої енергії, а поза цією смугою - менша, якою можна знехтувати. Визначена в такий спосіб ширина смуги частот F вважається шириною спектра сигналу. У цьому випадку енергія сигналу зосереджена в частотно-часовому прямокутнику зі сторонами T (вздовж осі часу) і F (вздовж осі частот). Для передавання сигналу з допустимою точністю необхідно мати канал зі смугою частот, ширина котрої F і час передавання T .

На (рис. 2.11, а) наведено приклад розподілу енергії сигналу на частотно-часовій площині (f , t). Спектр сигналу зосереджений біля несучої частоти f_0 і розташовується від $f_0 - F/2$ до $f_0 + F/2$. Рисунок має якісний характер, оскільки для фінітних сигналів енергія розподілена у смугі зі $0 \leq t \leq T$, $|f| < \infty$. Оскільки комплексна обвідна є відеосигналом, то частотно-часовий прямокутник, по якому розподілена основна частина енергії сигналу, виглядатиме, як на рис. 2.11, б. Частотно-часовий прямокутник виходить з базисного прямокутника (рис. 2.11, а) при зсуві останнього вниз за частотою на f_0 , що відповідає переходу від радіосигналу з несучою частотою f_0 до його комплексної обвідної.

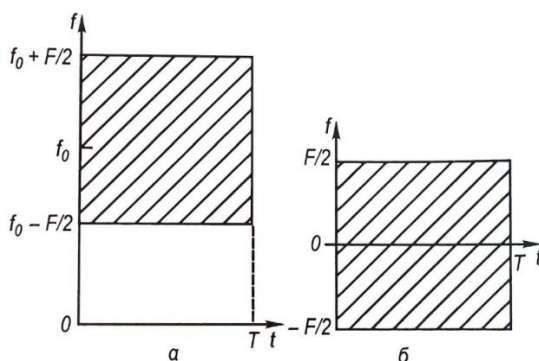


Рис. 2.11. Частотно-часова площина на радіочастоті (а) та частотно-часова площина на відеочастоті (б)

2.5. Оптимальна обробка сигналів

Якщо на вході приймача діє сигнал $x(t)$, дорівнює сумі корисного сигналу $u(t)$ і завади $n(t)$ або тільки заваді, то оптимальний приймач за відомими параметрами обчислює так званий *кореляційний інтеграл*, а потім порівнює його величину з порогом z_0 . Якщо завада є гауссовим випадковим процесом, спектральна щільність якого рівномірна (білий шум), то кореляційний інтеграл має вигляд

$$z = \int_0^T x(t)u(t)dt \quad (2.33)$$

Значення кореляційного інтеграла (2.33) визначається за допомогою корелятора (рис.2.12, а) або погодженого фільтра (рис. 2.12 б). Основними елементами корелятора, як випливає з виразу (2.33), є перемножник, генератор опорного сигналу ГОС й інтегратор. На перемножник надходять вхідний сигнал $x(t)$ та сигнал $u(t)$ від генератора сигналу ГОС. Добуток $x(t)$ і $u(t)$ інтегрується з моменту приходу ($t = 0$) і до моменту закінчення сигналу ($t = T$). Зазначимо, що корелятор є пристроєм зі змінними параметрами, оскільки режим його роботи залежить від зміни $u(t)$ у часі. Оскільки операції множення та інтегрування лінійні, то корелятор є лінійним пристроєм. Зважаючи на те, що він відфільтровує сигнал від завад і є лінійним пристроєм зі змінними параметрами, його іноді називають *активним фільтром* на відміну *пасивних* фільтрів, параметри яких є сталими в часі.

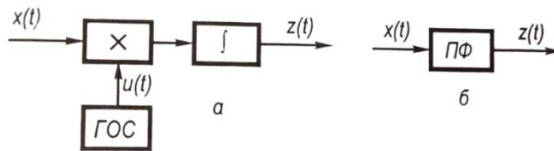


Рис. 2.12. Коректор (а) та погоджений фільтр (б)

Погоджений фільтр є пасивним. Напруга на виході погодженого фільтра в момент закінчення сигналу ($t = T$) з точністю до сталого множника a дорівнює напрузі на виході корелятора

$$U(T) = a \cdot z \quad (2.34)$$

Імпульсна характеристика погодженого фільтра

$$h(t) = au(T - t) \quad (2.35)$$

за формою є дзеркально відбитим сигналом із запізненням T .

Спільним між корелятором та погодженим фільтром є рівність (з точністю до сталої) вихідної напруги в момент часу $t=T$. Це визначає їх взаємну еквівалентність щодо виявлення сигналу. А різниця між ними полягає в наступному. Оскільки корелятор є пристроєм із змінними в часі параметрами, а погоджений фільтр -

пристроєм із сталими параметрами, то погоджений фільтр інваріантний до затримки сигналу та його початкової фази (наскільки ці величини зміняться у сигналі на вході фільтра, настільки вони зміняться і у сигналі на виході), а корелятор - не інваріантний.

Якщо сигнал має деякі невимірні або вимірні випадкові параметри, то структура оптимального приймача дещо змінюється, але його основна частина залишається такою самою, оскільки погоджений фільтр або корелятор завжди мають бути. Наприклад, при випадковій початковій фазі сигналу в приймачі з погодженим фільтром за останнім фільтром повинен слідувати детектор для виділення обвідної. У приймачі кореляційного типу мають бути ще другий (квадратурний) канал і схема виділення обвідної. Саме тому в оптимальному приймачі завжди є погоджений фільтр чи корелятор. Коефіцієнт передачі погодженого фільтра з імпульсною характеристикою (2.35)

$$k(\omega) = a^* g(\omega) \exp(-i\omega T) \quad (2.36)$$

де $g(\omega)$ - спектр сигналу, $*$ - знак комплексної спряженості. Із (2.36) випливають вирази для амплітудно-частотної характеристики (АЧХ)

$$|k(\omega)| = a|g(\omega)| \quad (2.37)$$

і фазочастотної характеристики (ФАХ) погодженого фільтра

$$\psi(\omega) = -\phi(\omega) - \omega T \quad (2.38)$$

Із (2.37) випливає, що АЧХ погодженого фільтра є вищою, чим більша спектральна щільність сигналу. На рис. 2.13, а зображені амплітудний спектр сигналу $a|g(\omega)|$, спектральна щільність завади N_0 і АЧХ фільтра $|k(\omega)|$, побудована відповідно до (2.37).

Фазовий спектр сигналу $\phi(\omega)$ і ФЧХ фільтра $\psi(\omega)$, що описуються рівнянням (2.38), показані на рис. 2.13, б. Штриховою лінією зображені складові правої частини рівняння (2.38). ФЧХ фільтра компенсує фазовий спектр сигналу, внаслідок чого на виході фільтра в якийсь момент часу всі частотні складові відгуку опиняються у фазі і, додаючись, дають максимум відклику.

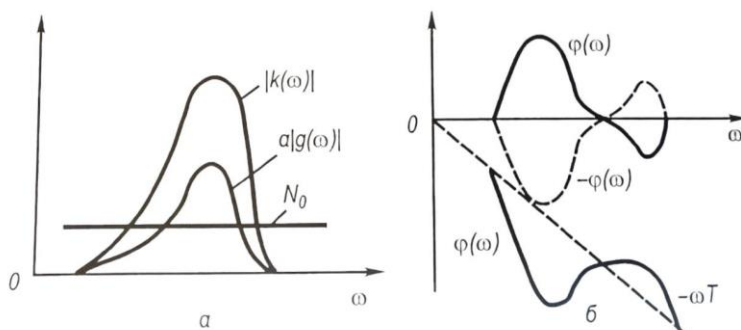


Рис. 2.13. Амплітудно-частотна (а) і фазочастотна (б) характеристики погодженого фільтра

Виняткова роль погодженого фільтра (або корелятора) в оптимальному приймачі пояснюється тим, що він максимізує відношення сигнал-завада на своєму виході. Це відношення у разі дії на вході фільтра білого шуму зі спектральною щільністю N_0 і сигналу з енергією E не залежить від форми сигналу:

$$q_0 = \sqrt{2 E / N_0} \quad (2.39)$$

При цьому максимальне значення сигналу на виході фільтра дорівнює aE , а середньоквадратичне значення завади - $a\sqrt{EN_0/2}$.

Результати, наведені в даному параграфі, суворо слушні для завади у вигляді випадкового гауссового процесу з рівномірною спектральною щільністю потужності (білий шум). Разом з тим корелятор або погоджений фільтр (чи його набір, або модифікація) завжди входять в оптимальний приймач.

2.6. Кореляційні функції сигналів

Оптимальний прийом сигналів здійснюється за допомогою погоджених фільтрів або кореляторів. Нормований до певної напруги відклик погодженого фільтра, знайдений за допомогою інтеграла згортки,

$$r_{jk}(\tau) = \frac{1}{E} \int_{-\infty}^{\infty} u_j(t) u_k(t - \tau) dt \quad (2.40)$$

де $u_j(t)$ - сигнал на вході фільтра, погодженого із сигналом $u_k(t)$; τ - зсув сигналу $u_j(t)$ щодо відклику $u_k(t)$. Енергії сигналів з номерами j і k дорівнюють E . При $j=k$ і $\tau=0$ з (2.40), відкидаючи індекси, маємо.

$$r(0) = r_{\max} = \frac{1}{E} \int_{-\infty}^{\infty} u^2(\tau) dt = 1 \quad (2.41)$$

що й визначає нормованість відклику погодженого фільтра.

Вираз у правій частині (2.40) визначає інтегральний взаємозв'язок між сигналами $u_j(t)$ та $u_k(t)$. при деякому зсуві τ . Якщо τ - змінна величина, то $r_{jk}(\tau)$ - функціонал, що залежить як від функцій $u_j(t)$ та $u_k(t)$. так і від зсуву τ . Саме тому $r_{jk}(\tau)$ називається **кореляційною функцією (КФ)** сигналів $u_j(t)$ та $u_k(t)$. Залежно від того, погоджений або не погоджений сигнал з фільтром та чи є додатково доплерівський зсув несучої частоти сигналу, кореляційні функції мають різні зображення.

Взаємна функція невизначеності (ВФН) двох сигналів з номерами j і k , за означенням, виражається через комплексні обвідні сигнали і через їх спектри таким чином:

$$R_{jk}(\tau, \Omega) = \frac{1}{2E} \int_{-\infty}^{\infty} u_j(t) u_k(t - \tau) e^{j\Omega t} dt = \frac{1}{4\pi E} \int_{-\infty}^{\infty} G_j(\omega - \Omega) G_k(\omega) e^{i\omega \tau} d\omega \quad (2.42)$$

де τ - зсув за часом між сигналами; Ω - доплерівський зсув частоти. З точністю до малих більш високого порядку нормований відклик погодженого фільтра пов'язаний із ВФН (2.42) співвідношенням

$$r_{jk}(\tau, \Omega) = \text{Re} R_{jk}(\tau, \omega) \exp(i\omega_0 \tau) \quad (2.43)$$

Взаємкореляційна функція (ВКФ) є перетином ВФН при $\Omega = 0$. Вважаючи $\Omega = 0$, з (2.43) отримуємо

$$R_{jk}(\tau) = \frac{1}{2E} \int_{-\infty}^{\infty} u_j(t) u_k^*(t - \tau) dt = \frac{1}{4\pi E} \int_{-\infty}^{\infty} G_j(\omega) G_k^*(\omega) e^{i\omega \tau} d\omega \quad (2.44)$$

Функція невизначеності (ФН). Якщо фільтр погоджений із сигналом, тобто $j = k$, то з (2.43), опустивши індекс j , отримуємо визначення ФН:

$$R(\tau, \Omega) = \frac{1}{2E} \int_{-\infty}^{\infty} u(t) u(t - \tau) e^{i\Omega t} dt = \frac{1}{4\pi E} \int_{-\infty}^{\infty} G(\omega - \Omega) G^*(\omega) e^{i\omega t} d\omega \quad (2.45)$$

Автокореляційна функція (АКФ)- перетин ФН при $\Omega = 0$. Взявши $\Omega = 0$, з (2.45) знаходимо АКФ:

$$R(\tau) = \frac{1}{2E} \int_{-\infty}^{\infty} u(t) u^*(t - \tau) dt = \frac{1}{4\pi E} \int_{-\infty}^{\infty} |G(\omega)|^2 e^{i\omega t} d\omega \quad (3.46)$$

З рівності (2.46) видно, що АКФ є перетворенням Фур'є енергетичного спектра комплексної обвідної сигналу. Відповідно до оберненого перетворення Фур'є енергетичний спектр:

$$|G(\omega)|^2 = 2E \int R(\tau) e^{-i\omega t} d\tau \quad (2.47)$$

Розглянемо приклад, що ілюструє властивості автокореляційної функції. На рис. 2.14 *а, б* зображені простий сигнал у вигляді прямокутного імпульсу та його автокореляційна функція. Максимум $R(\tau)$ припадає на момент закінчення сигналу ($t = T$). Це пояснюється тим, що $R(\tau)$ є нормованою напругою на виході фільтра, погодженого з вхідним сигналом, максимум якого збігається з моментом закінчення сигналу, тобто. при $\tau = T$ максимум $R(\tau) = 1$ відповідно до (2.41).

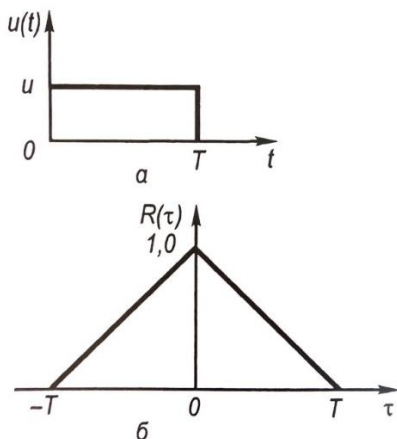


Рис. 2.14. Прямокутний імпульс (*а*) та його автокореляційна функція (*б*)

Для автокореляційної функції, зображеної на рис. 2.14 б, енергетичний спектр (квадрат модуля амплітудного спектра) відповідно до (2.47) описується функцією

$$|G(\omega)|^2 = (\sin \omega T / 2)^2 / (\omega T / 2)^2 \quad (2.48)$$

На рис. 2.15, а зображений фазоманіпульований шумоподібний сигнал (ФМ ШПС) тривалістю T , а на рис. 2.15, б - його АКФ. Елементарний імпульс має тривалість $\tau_0 = T/N$, де N - число імпульсів. Для ФМ ШПС, зображеного на рис. 2.15, а $N=15$. Автокореляційна функція ФМ має центральний пік в інтервалі $(-\tau_0, \tau_0)$ з одиничною амплітудою та бічні піки, розподілені на інтервалах $(-T, -\tau_0)$ і (τ_0, T) .

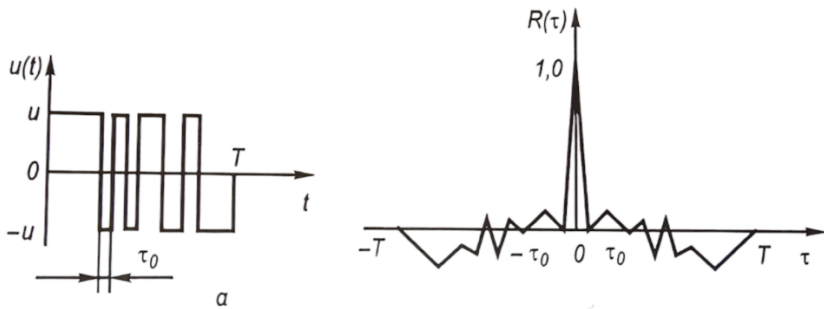


Рис. 2.15. Фазоманіпульований шумоподібний сигнал (а) та його автокореляційна функція (б)

Амплітуди бічних піків набувають різних значень, але в сигналах з «добrimi» кореляційними властивостями вони малі, тобто істотно менші за амплітуди центрального піка, що дорівнює одиниці. Існують різні оцінки бічних піків як АКФ, так і ВКФ, ВФН, ФН, проте всі вони описуються однаковими за формою відношенням. Для ФМ ШПС оцінка бічних піків має вигляд

$$R = \alpha / \sqrt{N} \quad (2.49)$$

де α - деяка величина, яка залежить від вигляду оцінки, класу сигналу і, в загальному випадку, від N . Для довільних ШПС із базою B оцінка бічних піків

$$R = \beta / \sqrt{B} \quad (2.50)$$

де β , як і α в (2.49), - деяка стала величина. Співвідношення (2.49), (20.50) визначають однакову залежність оцінок величини бічних піків від бази ШПС, оскільки N у ФМ ШПС пропорційною базі B . Чим більша база, тим менші бічні піки. У граничному переході, коли $B \rightarrow \infty$, АКФ має вигляд трикутного імпульсу зображеного на рис. 2.16. Бічні піки на рис. 2.16 не зображені, оскільки при $B \rightarrow \infty$ вони наближаються до нуля відповідно до (2.49), (2.50). Тривалість центрального піку АКФ також прямує до нуля, оскільки $\tau_0 = T/N$, а з ростом бази B число імпульсів $N \rightarrow \infty$. АКФ, зображена на рис. 2.16 називається *ідеальною*, тому що вона не має бічних піків. Саме таку АКФ мають тривалі реалізації шуму, що пояснює назву «шумоподібні» сигнали.

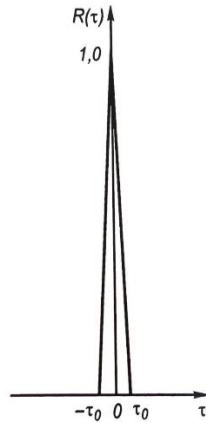


Рис. 2.16. Ідеальна АКФ

Частотна кореляційна функція (ЧКФ) - Перетин ФН при $\tau = 0$. Взявши $\tau = 0$, з (2.21) отримуємо

$$R(\Omega) = \frac{1}{2E} \int_{-\infty}^{\infty} |U(t)|^2 e^{i\Omega t} dt = \frac{1}{4\pi E} \int_{-\infty}^{\infty} G(\omega - \Omega) G^*(\omega) d\omega \quad (2.51)$$

З першої рівності (2.51) випливає, що ЧКФ є перетворенням Фур'є квадрата обвідної сигналу. Вона залежить від фазової структури сигналу, а визначається лише квадратам модуля обвідної. Наприклад, для простого сигналу (рис. 2.14, а) та для фазоманіпульованого шумоподібного сигналу (рис. 2.15, а) квадрат обвідної дорівнює одиниці (рис. 2.17).

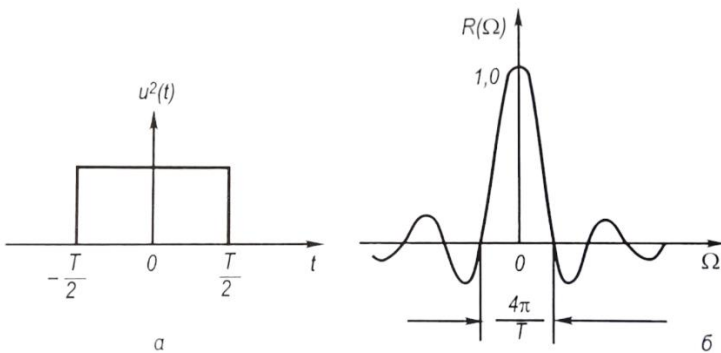


Рис. 2.17. Квадрат обвідної фазоманіпульованого сигналу (а) та його ЧКФ(б)

Тому ПКФ сигналів, зображених на рис. 2.14, а і 2.15, а, однакова і записується у вигляді

$$R(\Omega) = (\sin \Omega T/2) / (\Omega T/2) \quad (2.52)$$

Нульові значення (рис. 2.17, б) ідуть з інтервалом $2\pi/T$.

Максимум і симетрія кореляційних функцій. У функції (2.40), (2.42), (2.46) називаються, як було зазначено раніше, *кореляційними функціями* (КФ). Відомо, що максимум КФ має місце лише при $j = k$ і $\tau = 0$, $\Omega = 0$, тобто тільки в центрі ФН (або АКФ та ЧКФ). Максимум:

$$R_j(0,0) = R_k(0,0) = 1 \quad (2.53)$$

що аналогічно (2.41), а

$$|R_{jk}(\tau, \Omega)|_{j \neq k} < 1, \quad |R(\tau, \Omega)|_{\Omega \neq 0, \tau \neq 0} < 1$$

Властивість симетрії КФ полягає у тому, що

$$R_{jk}(-\tau, -\Omega) = R_{jk}^*(\tau, \Omega)e^{i\Omega\tau} \quad (2.54)$$

Із (2.54) випливає, що

$$|R_{jk}(-\tau, -\Omega)| = |R_j(-\tau, -\Omega)| = |R_j(\tau, \Omega)|, \quad (2.55)$$

$$R_j(-\tau) = R_j^*(\tau), \quad R_j(\Omega) = R_j^*(\Omega) \quad (2.56)$$

Обсяг та середньоквадратичні значення ВФН і ФН. Об'єм між поверхнею, що описується квадратом модуля ВФН, і площиною невизначеності (або просто об'єм ВФН), дорівнює одиниці, тобто.

$$\frac{1}{2\pi} \iint_{-\infty}^{\infty} |R_{ik}(\tau, \Omega)|^2 d\tau d\Omega = 1 \quad (2.57)$$

і не залежить від форми та номерів сигналів. Вважаючи $j = k$ і відкинувши індекси, маємо, що об'єм ФН також не залежить від форми сигналу і дорівнює одиниці, тобто.

$$\frac{1}{2\pi} \iint_{-\infty}^{\infty} |R(\tau, \Omega)|^2 d\tau d\Omega = 1 \quad (2.58)$$

Формули (2.57), (2.58) дозволяють знайти ефективні значення ВФН і ФН. Позначимо ці значення через $R_{jk \text{ еф}}$ та $R_{\text{еф}}$. Припустивши, що ВФН і ФН приблизно розподілені на прямокутник зі сторонами $3T$ і $4\pi f$, згідно з (2.57), (2.58) можемо записати, що $R_{jk \text{ еф}}^2 \cdot 4FT = R_{\text{еф}}^2 \cdot 4FT = 1$. Звідси знаходимо

$$R_{\text{еф}} = R_{jk \text{ еф}} = R_{j \text{ еф}} = \frac{1}{2\sqrt{FT}} = \frac{1}{2\sqrt{B}} \quad (2.59)$$

Із (2.59) видно, що чим більша база сигналу, тим менші ефективні значення ВФН і ФН. Формули (2.57) - (2.59) дуже важливі. Оцінка ефективного значення (2.59) збігається за формою з (2.50), але має визначений коефіцієнт, рівний 1/2. Як зрозуміло з наступного матеріалу, оцінка (2.59) дає нижню границю, тобто найменше ефективне значення, оскільки отримана за умови рівномірного розподілу ВФН і ФН на частотно-часовій площині. Насправді, для реальних сигналів розподіл цих функцій є нерівномірний, і тому ефективні значення ВФН і ФН перевищуватимуть знайдені за формулою (2.59).

Інтегральні рівності. Для визначення оцінок КФ широко використовують інтегральні рівності, що пов'язують між собою КФ різних сигналів. Однією із загальних інтегральних рівностей є:

$$\int_{-\infty}^{\infty} R_{ml}(\tau_1, \Omega_1) R_{jk}^*(\tau_1, \Omega_1) e^{-iz\tau} d\tau = \int_{-\infty}^{\infty} R_{kl}(\tau, z) R_{jm}^*(\tau, z) e^{-i\Omega_1\tau} d\tau \quad (2.60)$$

(Надалі індекс l буде опущений).

З формули (2.60) можна вивести окремі інтегральні рівності. Розглянемо їх.

А. Візьмемо $j = k = l = m, z = 0$. Маємо рівність Бакулева

$$\int_{-\infty}^{\infty} |R(\tau, \Omega)|^2 d\tau = \int_{-\infty}^{\infty} |R(\tau)|^2 e^{-i\Omega\tau} d\tau \quad (2.61)$$

Середня потужність сигналу ФН у перетині $\Omega = const$ є перетворенням Фур'є від квадрата АКФ.

Б. Візьмемо, $j = m, k = l, z = 0$. Маємо рівність Сталдера - Канна

$$\int_{-\infty}^{\infty} |R_{jk}(\tau, \Omega)|^2 d\tau = \int_{-\infty}^{\infty} R_j(\tau) R_k^*(\tau) e^{-i\Omega\tau} d\tau$$

В. Візьмемо, $j = m, k = l, z = \Omega = 0$. Маємо

$$\int_{-\infty}^{\infty} |R_{jk}(\tau)|^2 d\tau = \int_{-\infty}^{\infty} |R_j(\tau) R_k^*(\tau)| d\tau \quad (2.62)$$

З (2.62) випливає, що середнє значення квадрата модуля ВКФ сигналів з номерами j і k дорівнює середньому значенню добутку їх АКФ. Позначимо квадрати ефективних значень ВКФ через

$$R_{jk\text{ еф}}^2 = \frac{1}{2T} \int_{-\infty}^{\infty} |R_{jk}(\tau)|^2 d\tau, \quad R_{q\text{ еф}}^2 = \frac{1}{2T} \int_{-\infty}^{\infty} |R_q(\tau)|^2 d\tau \quad (2.63)$$

де T - тривалість сигналу, а $q = j$ або $q = k$. Використовуючи нерівність Буняковського-Шварца, з (2.63) отримуємо

$$R_{jk\text{ еф}}^2 \leq R_{j\text{ еф}} R_{k\text{ еф}} \quad (2.64)$$

З (2.64) випливає, що для зменшення ефективного значення ВКФ необхідно зменшувати ефективні значення АКФ.

Використання наведених інтегральних рівностей для оцінки КФ буде проілюстровано далі.

2.7. Основні типи ШПС

Відомо велика кількість різних ШПС, властивості яких відбито у багатьох книгах і журнальних статтях. ШПС можна розбити на частотно-модульовані (ЧМ) сигнали; багаточастотні (МЧ) сигнали; фазоманіпульовані (ФМ) сигнали (сигнали з кодовою фазовою модуляцією - КФМ сигнали); дискретні частотні (ДЧ) сигнали (сигнали з кодовою частотною модуляцією - КЧМ сигнали, частотно-маніпульовані (ЧМ) сигнали); дискретні складові частотні (ДСЧ) (складові сигнали з кодовою частотною модуляцією - СКЧМ сигнали).

У дужках вказані інші назви. Іноді ФМ сигнали називають просто ШПС, ДЧ сигнали - сигнали з частотою, що «стрибає».

Частотно-модульовані сигнали (ЧМ-сигнали) є неперервними сигналами, частота яких змінюється за заданим законом. На рис. 2.18, а зображений ЧМ-сигнал, частота якого змінюється за V-подібним законом від $f_0 - F/2$ до $f_0 + F/2$, де f_0 - центральна (несуча) частота сигналу, F - ширина спектра, яка дорівнює дев'ятій частоти $F = \Delta F$. Тривалість сигналу T .

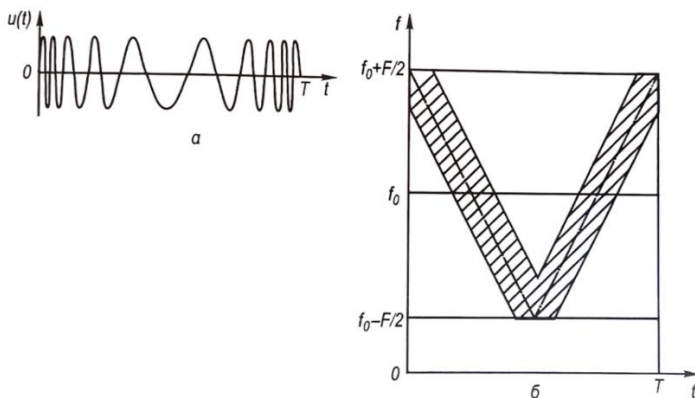


Рис. 2.18. Частотно-модульований сигнал (а) та частотно-часова площина (б)

На рис. 2.18, б зображена частотно-часова (f, t) площина, на якій заштриховано приблизний розподіл енергії ЧМ-сигналу за частотою і за часом. База ЧМ сигналу.

$$B = FT = \Delta f_d T \quad (2.65)$$

де Δf_d - девіація частоти

Частотно-модульовані сигнали знайшли широке застосування в радіолокаційних системах, оскільки для конкретного ЧМ-сигналу можна створити погоджений фільтр на приладах з поверхневими акустичними хвилями. Для систем зв'язку необхідно мати безліч сигналів. Необхідність швидкої зміни сигналів і переключення апаратури формування й обробки призводять до того, що реалізація зміни частоти стає дискретною. При цьому від ЧМ-сигналів переходять до БЧ-сигналів.

Багаточастотні сигнали (БЧ-сигнали) (рис. 2.19, а) є сумою N гармонік $u_1(t) \dots u_N(t)$, амплітуди і фази яких визначаються відповідно до законів формування сигналів. На частотно-часовій площині (рис. 2.19, б) штрихуванням виділений розподіл енергії одного елемента (гармоніки) БЧ-сигналу на частоті f_k . Всі елементи (гармоніки) цілком перекривають виділений квадрат зі сторонами F і T . База сигналу B дорівнює площі квадрата. Ширина спектру елемента $F_0 \approx 1/T$. Тому база БЧ-сигналу

$$B = FT = F/F_0 = N \quad (2.66)$$

тобто збігається з числом гармонік. БЧ-сигнали є неперервними: для їх формування й обробки важко пристосувати методи цифрової техніки. Крім цього, вони мають і такі недоліки, як поганий пік-фактор (нерізке вираження амплітуди сигналу, див. рис. 3.19, а) та потреба у великому числі частотних каналів N для одержання великої бази B .

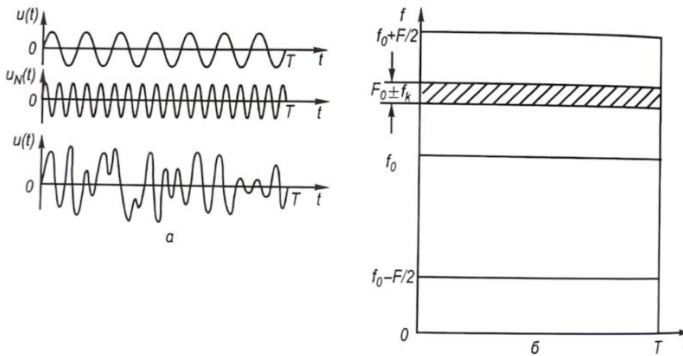


Рис. 2.19. Багаточастотний сигнал (а) і частотно-часова площина (б)

Фазоманіпульовані сигнали (ФМ-сигнали) є послідовністю радіоімпульсів, фази яких змінюються за заданим законом. Звичайно фаза набуває двох значень: 0 або π . При цьому радіочастотному ФМ-сигналові відповідає відео ФМ-сигнал (рис. 2.20 а), що складається з позитивних і негативних імпульсів. Якщо число імпульсів N , то тривалість одного імпульсу дорівнює $\tau_0 = T/N$, а ширина його спектра дорівнює приблизно ширині спектра сигналу: $F_0 = 1/\tau_0 = 1/N$. На частотно-часовій площині (2.20 б) штрихуванням виділений розподіл енергії одного елемента (імпульсу) ФМ-сигналу.

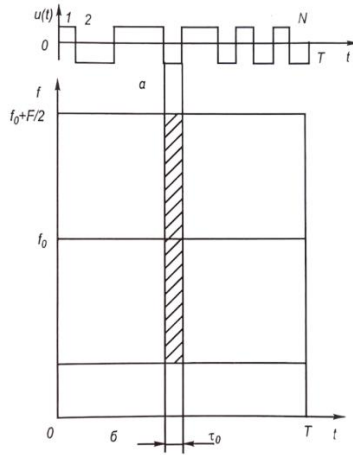


Рис. 2.20. Фазомодульований сигнал з кодовою фазовою модуляцією (а) і частотно-часова площа (б)

Всі елементи перекривають виділений квадрат зі сторонами F та T . База ФМ-сигналу

$$B = FT = T/\tau_0 = N \quad (2.67)$$

тобто дорівнює числу імпульсів у сигналі.

Можливість застосування ФМ-сигналів у якості ШПС із базами $B = 10^4 \dots 10^6$ обмежена в основному апаратурою обробки. При використанні погоджених фільтрів у вигляді приладів на поверхнево-акустичних хвилях можливий оптимальний прийом ФМ-сигналів з максимальними базами $B_{\max} = 1000 \dots 2000$. ФМ-сигнали, оброблювані такими фільтрами, мають широкі спектри (порядку $10 \dots 20$ МГц) і відносно короткі тривалості ($50 \dots 100$ мкс). Обробка ФМ-сигналів за допомогою відеочастотних ліній затримки при перенесенні спектра сигналів в область відеочастот дозволяє одержувати бази $B = 100$ при $F \approx 1$ МГц, $T \approx 100$ мкс.

Досить перспективними є погоджені фільтри на приладах із зарядовим зв'язком (ПЗЗ). Згідно з опублікованими даними за допомогою погоджених фільтрів

ПЗЗ можна обробляти ФМ-сигнали з базами $10^2 \dots 10^3$ при тривалості сигналів $10^{-4} \dots 10^{-1}$ с. Цифровий корелятор на ПЗЗ здатний обробляти сигнали до бази $4 \cdot 10^4$.

Слід зазначити, що ФМ-сигнали з великими базами доцільно обробляти за допомогою кореляторів (на великих інтегральних схемах або ПЗЗ). При цьому $B = 4 \cdot 10^4$ є граничною базою. Але при використанні кореляторів необхідно в першу чергу вирішити питання про прискорене входження в синхронізм.

Оскільки ФМ-сигнали дозволяють широко використовувати цифрові методи і техніку формування й обробки та реалізувати відносно великі бази їх, то вони є одним з перспективних видів ШПС.

Дискретні частотні сигнали (ДЧ-сигнали) є послідовністю радіоімпульсів (рис. 2.21 а), несучі частоти яких змінюються за заданим законом. Нехай число імпульсів у ДЧ-сигналі дорівнює M , тривалість імпульсу дорівнює $T_0 = T/M$, його ширина спектра $F_0 = 1/T_0 = M/T$. Над кожним імпульсом (рис. 2.21 а) зазначена його несуча частота. На частотно-часовій площині (рис. 2.21 б) штрихуванням виділені квадрати, в яких розподілена енергія імпульсів ДЧ-сигналу. Як видно із рис. 2.21 б, енергія ДЧ-сигналу розподілена нерівномірно на частотно-часовій площині.

База ДЧ-сигналів

$$B = FT = MF_0 \cdot MT_0 = M^2 F_0 T_0 = M^2 \quad (2.68)$$

оскільки база імпульсу $F_0 T_0 = 1$. Із (2.68) випливає основна перевага ДЧ-сигналів: для одержання необхідної бази B число каналів $M = \sqrt{B}$, тобто значно менше, ніж для БЧ-сигналів. Саме це і обумовило застосування таких сигналів у системах зв'язку. Разом з тим для великих баз ($B=10^4 \dots 10^6$) використовувати тільки ДЧ-сигнали недоцільно, тому що число частотних каналів $M = 10^2 \dots 10^3$, що видається надмірно велике.

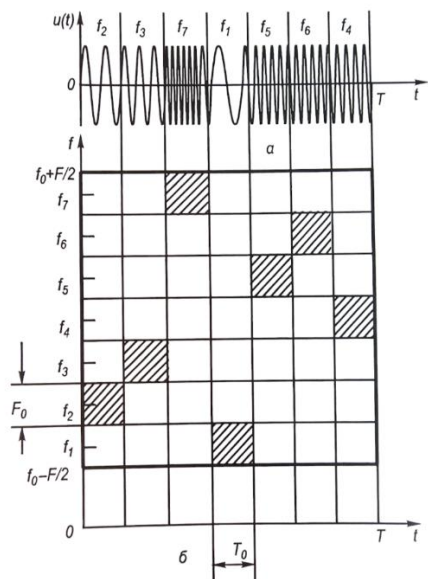


Рис. 2.21. Дискретний сигнал з кодовою частотною модуляцією (а)
і частотно-часова площина (б)

Дискретні складові частотні сигнали (ДСЧ-сигнали) є ДЧ сигналами, кожен імпульс яких замінений шумоподібним сигналом. На рис. 2.22, а зображений відеочастотний ФМ-сигнал, окремі частини якого передаються на різних несучих частотах. Номери частот зазначені над ФМ-сигналом. На рис. 2.22, б зображена частотно-часова площина, на якій штрихуванням виділено розподіл енергії ДСЧ-сигналу. Рис. 2.22, б за структурою не відрізняється від рис. 2.21, б, але на першому з них площа $F_0 T_0 = N_0$ - дорівнює числу імпульсів ФМ-сигналу в одному частотному елементі ДСЧ-сигналу. База ДСЧ-сигналу

$$B = FT = M^2 F_0 T_0 = N_0 M^2 \quad (2.69)$$

Число імпульсів повного ФМ-сигналу

$$N = N_0 M \quad (2.70)$$

Зображений на рис. 2.22 ДСЧ-сигнал містить як елементи Ф-сигнали. Тому такий сигнал скорочено називатимемо ДСЧ-ФМ-сигнал. За елементи ДСЧ-сигналу можна взяти ДЧ-сигнали. Розподіл енергії такого сигналу на частотно-часовій площині зображено на рис. 2.23. Якщо база елемента ДЧ-сигналу

$$B_0 = F_0 T_0 = M_0^2 \quad (2.71)$$

то база всього сигналу

$$B = M_0^2 M^2 \quad (2.72)$$

Такий сигнал можна скорочено позначати ДСЧ-ЧМ. Число частотних каналів у ДСЧ-ЧМ-сигналі дорівнює $M_0 M$. Якщо ДЧ-сигнал (рис. 2.22) і ДСЧ-ЧМ-сигнал мають рівні бази, то вони мають і однакову кількість частотних каналів. Тому особливих переваг ДСЧ-ЧМ-сигнал перед ДЧ-сигналом немає. Але принципи побудови ДСЧ-ЧМ-сигналу можуть виявитися корисними при побудові великих систем ДЧ-сигналів.

Таким чином, найбільш перспективними ШПС для систем зв'язку є ФМ-, ДЧ-, ДСЧ-ФМ-сигнали.

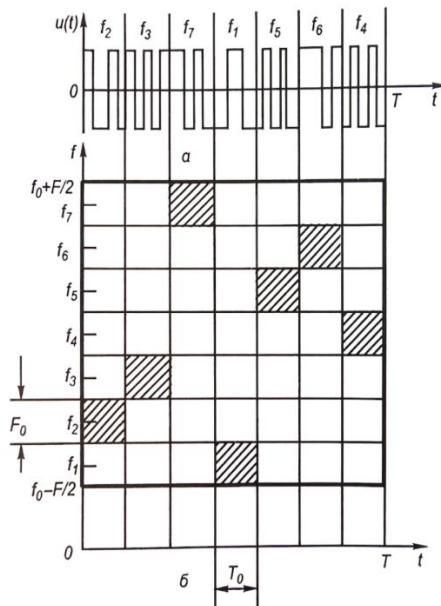


Рис. 2.22. ДСЧ-ФМ-сигнал (а) і частотно-часова площина (б)

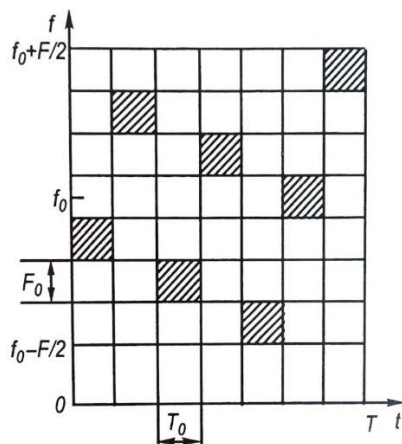


Рис. 2.23. Частотно-часова площина дискретного складеного сигналу з частотною маніпуляцією (ДСЧ-ЧМ-сигналу)

Контрольні питання

1. Дайте визначення шумоподібного сигналу.
2. Як визначається база сигналу?
3. Дайте визначення завадозахищеності систем зв'язку.
4. Як визначається завадостійкість широкосмужової системи зв'язку?
5. Як визначити скритність системи зв'язку?
6. У чому полягає особливість асинхронних адресних систем зв'язку з використанням ШПС?
7. Як визначається ефективність стільникових рухомих систем зв'язку?
8. Запишіть умови поділу променів при боротьбі з багатопротеневістю
9. Які особливості електромагнітної сумісності при використанні шумоподібних сигналів?
10. Які структурні схеми широкосмужових систем зв'язку Ви знаєте?

РОЗДІЛ 3. МЕТОДИ ОБРОБКИ СИГНАЛІВ

3.1. Загальні відомості про прийом сигналів

На приймальній стороні про передаються сигнали зазвичай є деякі попередні (априорні) відомості. Можуть бути відомими, наприклад, частота несучої, вид модуляції тощо. Сигнал, про який все відомо, не несе інформації, а абсолютно невідомий сигнал не можна було б прийняти.

Відомі параметри сигналу використовуються у приймачі для кращого відокремлення сигналів від перешкод. Чим більше ми знаємо про сигнал, тим досконалішими можуть бути методи прийому.

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

Залежно від виду та призначення передачі при прийомі сигналів виникають такі основні завдання:

- Виявлення сигналів;
- Розрізнення сигналів;
- Відновлення сигналів.

При виявленні сигналів завдання зводиться до отримання відповіді питанням, є на вході приймача сигнал чи ні, точніше, чи є на вході сигнал плюс шум чи лише шум. З таким завданням ми зазвичай зустрічаємось у радіолокації, вона також має місце і в деяких системах зв'язку. Якщо ми можемо виявити сигнал, то з'являється можливість передачі за допомогою двійкового коду. Наявність сигналу (надсилка) буде відповідати символу 1, відсутність сигналу (пауза) — символу 0. Цей принцип використовується в системах з пасивною паузою.

При передачі двох сигналів S_1 і S_2 виникає завдання виявлення, а розрізнення сигналів. Тут необхідно дати відповідь на запитання: чи є на вході приймача сигнал S_1

або сигнал S_2 . Відповідь це питання визначається не властивостями кожного сигналу окремо, які різницею. Сигнали можуть відрізнятися один від одного своїми параметрами. Очевидно, потрібно прагнути до того, щоб відмінність була найбільшою та стійкою до впливу перешкод. Випадок виявлення може розглядатися як вироджений випадок розрізнення двох сигналів, коли один з них тотожно дорівнює нулю.

Передача двійковим кодом, у якому кожному символу (1 та 0) відповідає певний сигнал (S_1 і S_2), не рівний нулю, називається передачею з активною паузою. Випадок розрізнення багатьох сигналів мало відрізняється від випадку розрізнення двох сигналів.

Завдання відновлення повідомлення істотно відрізняється від задач виявлення та розрізнення сигналів. Вона полягає в тому, щоб отримати вихідний відеосигнал $v(t)$, що найменш відрізняється від переданого повідомлення $u(t)$. При цьому повідомлення $u(t)$ наперед невідоме; відомо лише, що воно належить до деякої множини або є реалізацією деякого випадкового процесу.

На рис. 3.1 наведено функціональну схему обробки дискретних сигналів. Прийнятий сигнал, спотворений на заваді, в приймачі піддається певній обробці, детектується і надходить для розпізнавання на вирішальний пристрій. Вочевидь, що можливість правильного розпізнавання сигналів істотно залежить відношення сигналу до перешкоди на вході вирішального устрою. У зв'язку з цим основним завданням обробки сигналів у приймачі є збільшення відношення сигнал/шум. Обробка аналів, зазвичай, зводиться до тих чи інших методам фільтрації.

У каскадах обробки проводиться також посилення сигналів до величин, за яких можуть нормально працювати детектор і вирішальний пристрій.

У звичайному приймачі безперервних сигналів додетекторна обробка сигналів здійснюється за допомогою підсилювачів резонансних, що забезпечують необхідну частотну селекцію. Функції післядетекторної обробки виконуються відеопідсилювачем (або підсилювачем низької частоти). Вирішальний пристрій у таких приймачах відсутній. Замість нього на виході є пристрій, що відтворює або записує (що реєструє) прийняте повідомлення.

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

Стробування сигналу є найпростішим видом обробки. Воно давно відоме і широко застосовується у практиці. Обробку сигналів стробування часто називають прийомом за методом укорочених контактів. При стробированні даного елемента сигналу виробляється відлік його поточного значення (напруги чи / струму) у певний час. Останній вибирається в тій частині елемента, яка найменше схильна до спотворень. Так як статистичні характеристики перешкод мало залежить від вибору моменту часу, то момент стробування необхідно вибирати тоді, коли корисний сигнал має максимальне значення. Внаслідок перехідних процесів у фільтрах межі елементів сигналу значною мірою спотворено і тому максимальне значення сигналу, як правило, відповідає середній частині елемента. Стробирование проводиться з допомогою спеціальних сигналів, які від системи синхронізації.

Фільтрування сигналів може виконуватися як до, так і після детектора. При використанні синхронного детектора фільтрація до і після детектора може дати один і той же результат. Практично ж зазвичай проводять фільтрацію сигналу двічі: як до, так і після детектора.

Операція інтегрування може розглядатися як процес накопичення або як визначення середнього значення сигналу. Будь-який фільтр тією чи іншою мірою проводить 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 прийому можна класифікувати за видами детекторів, що застосовуються, за способами додетекторної і післядетекторної обробки. Розрізняють такі методи прийому:

- когерентний;
- некогерентний;
- Кореляційний;
- Автокореляційний.

3.2. Прийом сигналів із накопиченням

Одним із найбільш відомих методів прийому сигналів за наявності перешкод є метод накопичення. При використанні методів накопичення сигнал, що передається, багаторазово повторюється (дублюється). Кожен із сигналів, що передаються, по-різному спотворюється перешкодами, так як переданий сигнал і перешкода взаємно незалежні. На приймальному кінці каналу зв'язку кілька зразків сигналу порівнюються між собою і в результаті такого порівняння сигналів вдається більш точно відновити правильну форму переданого сигналу і за рахунок цього підвищити ймовірність його правильного прийому. Практично порівняння різних зразків прийнятого сигналу здійснюється шляхом їх підсумовування. З цієї причини вказаний метод прийому сигналів отримав назву методу накопичення.

При використанні методу накопичення сигнал отримує нову властивість - властивість періодичного повторення.

При підсумовуванні сигналів, що приймаються, енергія корисного сигналу збільшується при l -кратному повторенні сигналу l^2 разів. Енергія перешкоди, що являє собою випадковий сигнал, збільшується при підсумовуванні l разів. Таким

чином, при використанні методу накопичення сигналів відношення $\frac{P_c}{P_{\Pi}}$ на виході накопичувача зростає в l разів. Таке поліпшення відношення сигнал/перешкода досягається за рахунок збільшення часу передачі сигналу l разів, тобто за рахунок зниження ефективності системи зв'язку.

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

У разі телеграфного зв'язку, кожна кодова комбінація, що складається з елементів 0 і 1, передається кілька разів. Якщо ймовірність збою символів 1 і 0 однакова, то на прийомі рішення виноситься за «більшістю», тобто відтворюється

символ 1 на даній позиції, коли їх число на цій позиції більше від числа символів 0, і, навпаки, відтворюється 0, коли число "нулів" більше числа "одиниць":

передана комбінація 01001

перша прийнята » 00001

друга » » 11010

третя » » 01101

відтворена комбінація 01001

Зауважимо, що можна було б отримати п зразків сигналу не шляхом їх повторення в часі, а шляхом передачі незалежними каналами, розділеним по частоті, або яким-небудь іншим способом.

Існують інші різновиди методу накопичення. До них, зокрема, відноситься метод синхронного накопичення, що знайшов застосування радіолокації. У цьому методі протягом посилки сигналу береться не один відлік, а кілька. На прийомі ці відліки підсумовуються у накопичувачі.

Розглянемо як приклад метод синхронного накопичення, застосовуваного виявлення слабких періодичних сигналів. При використанні цього методу 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 зі

зрушенням щодо початку відліку $\frac{K}{m} T$ подаються на накопичувач та підсумовуються (рис. 3.2, а).

Визначимо відношення сигнал/перешкода K -ом накопичувачі. Позначимо миттєві значення перешкоди, що надходить на накопичувач, через

$$n_i = n(t + \kappa \frac{T}{m} + iT), \quad (3.1)$$

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 перешкоди дорівнюватиме

$$\overline{n_i} = \frac{1}{n} \sum_{i=1}^t n_i. \quad (3.2)$$

Так як n_i - Випадкова величина, то при малому числі $l \overline{n_i}$ сильно флюктує близько свого середнього значення. Зі збільшенням l ймовірність великих флюктуації зменшується.

Корисний сигнал $x(t)$, що надходить на накопичувач l циклів накопичення, збільшується в l раз, тобто стає рівним $lx(t)$. Помилка прийому можлива до тих пір, поки флюктуації можуть перевершити корисний сигнал по абсолютній величині.

Визначимо ймовірність того, що відхилення $\overline{n_i}$ від середнього значення не перевищить величини a , що дорівнює середньому значенню сигналу за l повторень. Ця ймовірність визначається нерівністю Чебишева

$$P\left\{\left|\overline{n_i} - \overline{n}\right| < a\right\} > 1 - \frac{D(n)}{lx^2}, \quad (3.3)$$

де $D(n)$ - дисперсія $n(t)$, рівна $P_{II}^2 = \sigma_i^2 x^2 = P_c$.

Розв'язуючи нерівність (3.3) щодо l , отримаємо

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

Таким чином, верхня межа частоти повторення l при заданому P , P_c та P_{II} може бути з нерівності (3.4).

Ймовірність $P\left\{\left|\overline{n_i} - \overline{n}\right| < x\right\}$ може бути знайдена за наближеною формулою

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

Ймовірність помилки прийому сигналу методом синхронного накопичення дорівнює

$$P_{out} = 1 - P\left\{\left|\overline{n_i} - \overline{n}\right| < x\right\} \approx 1 - \Phi\left(\frac{x}{\sigma} \sqrt{\frac{l}{2}}\right).$$

З цієї рівності випливає, що стійкість до перешкод системи з синхронним накопиченням тим більше, чим більше відношення сигнал/перешкода на вході накопичувача і чим більше число циклів накопичення l .

Розглянуті вами методи прийому слабких сигналів не вичерпують можливих методів прийому слабких сигналів.

Нехай окремі відліки прийнятого сигналу:

$$x_1 = s + \omega_1, x_2 = s + \omega_2, \dots, x_n = s + \omega_n; \quad (3.6)$$

тоді сума відліків

$$x = \sum_{k=1}^n x_k = \sum_{k=1}^n (s + \omega_k) = ns + \sum_{k=1}^n \omega_k = b + \xi. \quad (3.7)$$

Величина $b=ns$ є корисним сигналом на виході приймача. Випадкова величина

$$\xi = \sum_{k=1}^n \omega_k$$

є перешкодою. Відношення сигналу до перешкоди на виході приймача

$$P\left\{\left|\overline{n_i} - \overline{n}\right| < x\right\}$$

одно:

$$q_{\text{вых}} = \frac{b^2}{D\xi} = \frac{n^2 s^2}{D(\sum \omega_k)} = \frac{n^2 s^2}{\sum D\omega_k} = \frac{n^2 s^2}{nD\omega_k} = \frac{n^2 s^2}{\sigma_{\Pi}^2} = nq_{\text{вых}}. \quad (3.8)$$

Тут ми вважали, що ω_k некорельовані та мають однаковий розподіл,

$$q_{\text{ex}} = \frac{s^2}{\sigma_{\Pi}^2} = \frac{P_c}{P_{\Pi}}$$

відношення сигналу до перешкод на вході приймача.

Таким чином, за описаних умов накопичення відліків сигналу дозволяє збільшити відношення сигналу до перешкоди на виході приймача рівно в n разів. Суть справи зводиться до того що, що потужність сигналу під час підсумовування зростає пропорційно n^2 (складається напруги), а потужність перешкоди - пропорційно n (сумуються потужності). Тому відношення сигналу до перешкоди збільшується в n раз, якщо перешкоди незалежні. За наявності кореляції між значеннями перешкоди цей вигравш буде меншим.

Метод накопичення можна здійснити, беручи не суму відліків x_k , а інтеграл функції, що безперервно змінюється $x(t) = s + \omega(t)$ час T , що дорівнює тривалості сигналу,

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

Тут b - також постійна, що виражає сигнал на виході накопичувача (інтегратора), а ξ - випадкова величина, що визначає перешкоду на виході інтегратора.

Визначимо дисперсію випадкової величини ξ

$$\begin{aligned} D\xi &= \overline{\left[\int_0^T \omega(t)dt \right]^2} = \overline{\int_0^T \int_0^T \omega(t)\omega(t_1)dt dt_1} = \\ &= \int_0^T \int_0^T \omega(t)\omega(t_1)dt dt_1 = \int_0^T dt \int_0^T B_{\omega}(t-t_1)dt_1, \end{aligned} \quad (3.9)$$

де $B_{\omega}(t-t_1)$ - функція кореляції перешкоди $\omega(t)$. Якщо спектр перешкоди рівномірний у досить широкій смузі частот F , тобто інтервал кореляції перешкоди $\Delta\tau \ll T$, то кінцеві межі інтегрування можуть бути замінені на нескінченні:

$$\int_0^T B_{\omega}(t-t_1)dt_1 \approx \int_{0-\infty}^{\infty} B_{\omega}(\tau)d\tau = 2 \int_0^{\infty} B_{\omega}(\tau)d\tau, \quad (3.10)$$

де $t-t_1 = \tau$. За визначенням інтервал кореляції

$$\Delta\tau = \frac{1}{B(0)} \int_{-\infty}^{\infty} B(\tau)d\tau = \frac{G(0)}{B(0)}. \quad (3.11)$$

У цьому випадку $B(0) = G(0)2F$ і $\Delta\tau = \frac{1}{2F}$. Отже,

$$D\xi = B(0)\Delta\tau T = B(0) \frac{T}{2F}.$$

Тоді ставлення сигналу до перешкод на виході інтегратора буде

$$q_{\text{вых}} = \frac{b^2}{D\xi} = \frac{s^2 T^2}{B(0)\Delta\tau T} = \frac{T}{\Delta\tau} = 2TFq_{\text{вх}}. \quad (3.12)$$

Отже, виграш, який отримується при інтегруванні, тим більше, чим більше ставлення $T/\Delta\tau$. Описаний спосіб прийому називається інтегральним.

Зауважимо, що $\frac{T}{\Delta\tau} = n$ є кількість незалежних відліків перешкоди на інтервалі T . Це означає, що ф-ли (3.8) і (3.12) збігаються, т. є. заміна підсумовування

незалежних значень безперервним інтегруванням додаткового виграшу не дає. Однак практична реалізація методу інтегрування здійснюється простіше, ніж підсумовування дискретних значень. Так, при прийомі телеграфних сигналів як інтегратор широко використовується ланцюжок RC, де синхронно розряджається після закінчення кожної елементарної посилки (рис. 3.2, б). В кінці кожної посилки заряд на ємності приблизно пропорційний інтегралу вхідного сигналу. Додетекторне інтегрування можна здійснити за допомогою великої добротності резонатора.

У найпростішому випадку таким резонатором може бути коливальний контур.

3.3. Кореляційний прийом

Структурну схему кореляційного приймача наведено на рис. (3.3,а). Вона складається з перемножувача MPL, генератора опорного колювання Р та інтегратора І.

При кореляційному прийомі деякий час T вимірюється значення функції взаємної кореляції $y(T)$ прийнятого сигналу $x(t) = s(t) + \omega(t)$ та опорного колювання $s_0(t)$. Якщо опорне колювання тотожно переданого сигналу, тобто, $s_0(t) \equiv s(t)$, то цю функцію можна записати в наступному вигляді:

$$\dot{y}(T) - \frac{1}{T} \int_0^T \dot{x}(t) \dot{s}^*(t) dt = \frac{2}{T} \int_0^T \dot{x}(t) \dot{s}^*(t) dt + i \frac{2}{T} \int_0^T x(t) \hat{s}(t) dt, \quad (3.13)$$

де $\dot{x}(t)$; $\dot{s}(t)$ - аналітичні сигнали, що відповідають $x(t)$; $s(t)$; $\dot{s}^*(t)$ - функції комплексно-сполучена із сигналом $s(t)$.

Залежно від методу реєстрації сигналу на виході, кореляційний прийом може бути когерентним і некогерентним. При когерентному прийомі на момент часу T відраховується дійсне значення функції $y(T)$, тобто.

$$\operatorname{Re}[y(T)] = \frac{2}{T} \int_0^T x(t)s(t)dt. \quad (3.14)$$

При некогерентному прийомі відраховується модуль цієї функції

$$\frac{1}{T} = \left| \int_0^T \dot{x}(t) \dot{s}^*(t) dt \right| = \frac{2}{T} \sqrt{\left[\int_0^T x(t)s(t)dt \right]^2 + \left[\int_0^T x(t) \wedge s(t) \right]^2}. \quad (3.15)$$

Справжнє значення сигналу на виході кореляційного приймача можна подати у вигляді суми двох складових

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

$$b = \frac{1}{T} \int_0^T s^2(t)dt = P_c \quad \xi = \frac{1}{T} \int_0^T s(t)\omega(t)dt$$

де - Корисний сигнал, - завада

на виході приймача. Визначимо дисперсію перешкоди

$$\begin{aligned} D\xi &= \overline{\left[\frac{1}{T} \int_0^T s(t)\omega(t)dt \right]^2} = \frac{1}{T^2} \int_0^T \int_0^T s(t)s(t_1)\omega(t)\omega(t_1)dt dt_1 = \\ &= \frac{1}{T^2} \int_0^T s(t)dt \int_0^T s(t_1)\overline{\omega(t)\omega(t_1)}dt_1 = \frac{1}{T^2} \int_0^T s(t)dt \int_0^T B_\omega(t-t_1)dt_1 dt. \end{aligned} \quad (3.10)$$

Припустимо, що інтервал кореляції перешкод $\Delta\tau$ настільки малий, що сигнал протягом цього інтервалу практично не змінюється (ширина спектра сигналу мала в порівнянні з шириною спектра перешкоди). Тоді з урахуванням співвідношення (3.10) та (3.11) отримаємо

$$D\xi = \frac{1}{T^2} \int_0^T s^2(t) dt \int_0^T B_{\omega}(t-t_1) dt_1 \approx \frac{\Delta\tau}{T} P_c P_{\Pi}, \quad (3.17)$$

де $P_{\Pi} = B_{\omega}(0)$ - Потужність на вході приймача.

Відношення сигналу до перешкоди на виході кореляційного приймача при когерентному прийомі дорівнюватиме:

$$q_{кз} = \frac{b^2}{D\xi} = \frac{T}{\Delta\tau} q_{вх} \approx 2TFq_{вх}. \quad (3.18)$$

При некогерентному прийомі це ставлення буде вдвічі менше:

$$q_{нкз} = \frac{b^2}{2D\xi} = \frac{1}{2} \frac{T}{\Delta\tau} \approx TFq_{вх}. \quad (3.19)$$

Порівняння ф-л (3.12) та (3.18) показує, що кореляційний спосіб прийому можна розглядати як узагальнення методу накопичення сигналів довільної форми.

3.4. Автокореляційний прийом

У схемі автокореляційного приймача відсутній спеціальний генератор опорних коливань (рис. 3.3, б). Зате в ній є лінія затримки ЛЗ, що здійснює затримку сигналу, що приймається на час τ . Затримане вагання $x(t-\tau)$ використовується як опорний.

У разі автокореляційного прийому виконується інтегральна операція

$$y = \frac{1}{T} \int_0^T x(t)x(t-\tau)dt = \frac{1}{T} \int_0^T [s(t) + \omega(t)][s(t-\tau) + \omega(t-\tau)]dt. \quad (3.20)$$

Розкриваючи дужки та вважаючи $\tau \ll T$, отримаємо

$$\begin{aligned} y &= \frac{1}{T} \int_0^T s(t)s(t-\tau)dt + \frac{1}{T} \int_0^T s(t)\omega(t-\tau) + \frac{1}{T} \int_0^T \omega(t)s(t-\tau)dt + \\ &+ \frac{1}{T} \int_0^T c(t)\omega(t-\tau)dt = B_{ss} + B_{s\omega} + B_{\omega ss} + B_{\omega\omega} = b + \xi. \end{aligned} \quad (3.21)$$

Тут:

$$b = B_{ss}(\tau) = P_c \rho_{II},$$

$$\xi = B_{s\omega}(\tau) + B_{\omega s}(\tau) + B_{\omega\omega}(\tau).$$

Далі можна показати, що відношення сигналу до перешкоди на виході приймача автокореляційного при малому рівні перешкод ($q_{\text{вх}} \gg 1$) буде приблизно таким самим, як і для кореляційного приймача. При великому рівні перешкод ($q_{\text{вх}} \ll 1$) автокореляційний приймач за перешкодою наближається до приймача з квадратним детектором. Нижча завадостійкість автокореляційного приймача порівняно з кореляційним за будь-яких значень $q_{\text{вх}}$ обумовлена наявністю перешкоди у тракці опорної напруги. Автокореляційний прийом можливий і в тому випадку, коли відсутні відомості не лише про фазу сигналу, а й про частоту. Це можливо завдяки тому, що опорне коливання породжується самим сигналом, а не створюється спеціальним генератором в місці прийому.

Наприклад розгляду автокореляційний прийом періодичного сигналу і натовість адитивної флюктуаційної перешкоди. Нехай $s(t)$ - періодичний сигнал, а $\omega(t)$ - перешкода, яку вважатимемо ергодичним процесом. Сумарне коливання сигналу та перешкоди на вході приймача $x(t) = s(t) + \omega(t)$. Визначимо функцію кореляції цього коливання

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

Оскільки сигнал і перешкоду вважатимуться незалежними, їх взаємно кореляційні функції $B_{s\omega}$ і $B_{\omega s}$ звертаються у нуль.

Тоді

$$B_x(\tau) = B_s(\tau) + B_\omega(\tau)$$

Функція автокореляції перешкод B_ω як ергодичного процесу зменшується при зростанні аргументу τ , а функція автокореляції періодичного сигналу є періодичною функцією з тим самим періодом. Функція кореляції прийнятого сигналу є сумою цих двох функцій (рис.6.4). Таким чином, за досить великого τ можна отримати періодичну криву, що відповідає переданому сигналу $s(t)$, тобто. виділити сигнал із шуму. Величина τ , за якої вже можна знехтувати $B_\omega(\tau)$ залежить від інтервалу кореляції перешкоди та відношення сигналу до перешкоди. Вочевидь, що більше рівень перешкод, то більше вписувалося необхідний час затримки.

3.5. Когерентний та некогерентний прийоми

Функціональна схема когерентного наведена на рис. (3.5 а). Вона складається з перемножувача МНЛ, генератора опорної напруги Г та фільтра ФНЧ. Опорне коливання $s_0(t)$ при когерентному прийомі є точною копією переданого сигналу $s(t)$.

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

При некогерентному прийомі апріорні відомості про початкову фазу прийнятого сигналу не враховуються, тому в схемі приймача можна застосувати синхронний, а амплітудний детектор (рис. 3.5, б).

Нехай на приймач впливає регулярний сигнал

$$s(t) = A_0 \cos \omega_0 t, \quad 0 \leq t \leq T \quad (3.22)$$

та флуктаційна перешкода $\omega(t)$. Перешкоду із симетричним щодо частоти сигналу спектром можна записати у вигляді квазігармонічного коливання:

$$\omega(t) = U_1 \cos \omega_0 t + U_2 \sin \omega_0 t = U \cos[\omega_0 t + \varphi(t)], \quad (3.23)$$

де U_1, U_2 - випадкові незалежні величини, що мають нормальний розподіл ймовірностей із середнім значенням, рівним нулю, та дисперсією

$\sigma_{\Pi}^2 = \overline{U_1^2(t)} = \overline{U_2^2(t)} = N_0 F$, де N_0 - спектральна щільність перешкоди, F - Ефективна ширина спектра перешкоди. Сумарне колювання сигналу та перешкоди.

$$x(t) = s(t) + \omega(t) = (A_0 + U_1) \cos \omega_0 t + U_2 \sin \omega_0 t. \quad (3.24)$$

Відповідно до цього виразу на рис. 6.6. побудована векторна діаграма сигналу та перешкоди. Відповідно до цієї діаграми перешкода подається щодо сигналу у вигляді двох складових, що знаходяться в квадратурі: U_1 і U_2 .

При когерентному прийомі детектор виділяє сигнал та синфазну складову перешкоди U_1 . Квадратурна складова перешкоди U_2 . У цьому повністю виключається. Помилки при когерентному прийомі обумовлені флуктаціями амплітуди перешкоди U_1 , що має нормальний розподіл ймовірностей

$$P(\omega) = \frac{1}{\sqrt{2\pi}\sigma_{\Pi}} e^{-\omega^2/(2\sigma_{\Pi}^2)} \quad (3.25)$$

При некогерентному прийомі детектора реагує на обидві складові перешкоди U_1 і U_2 і виділяє огинаючу сумарного колювання сигналу та перешкоди. Помилки в цьому випадку обумовлені флуктаціями огинаючої U сумарного колювання $x(t)$, яка розподілена за узагальненим законом Релея:

$$P(\omega) = \frac{U}{\sigma_{\Pi}^2} I_0\left(\frac{AU}{\sigma_{\Pi}^2}\right) e^{-\frac{U^2+A^2}{2\sigma_{\Pi}^2}}, \quad (3.26)$$

де $I_0(x)$ - модифікована функція Бесселя нульового порядку, σ_{Π}^2 - Дисперсія перешкоди.

Розподіл (3.26) називається узагальненим розподілом Релея чи розподіл Райсу. При $A = 0$ розподіл (3.26) перетворюється на звичайний розподіл Релея.

Вважаючи $z = U/\sigma_{\Pi}$ і $a = A/\sigma_{\Pi}$, розподіл Райсу можна записати в наступному вигляді:

$$p(z) = z e^{-\frac{a^2+z^2}{2}} I_0(az). \quad (3.27)$$

На рис. 6.6. зображено графіки цього розподілу для різних значень $a = A/\sigma_{\Pi} = \sqrt{2P_c/P_{\Pi}}$, де $P_c = A_2/2$ - Потужність сигналу, $P_{\Pi} = \sigma_{\Pi}^2$ - Потужність потужності.

Вираз для щільності ймовірності фази сумарного коливання має такий вигляд:

$$P(\phi) = \frac{1}{2\pi} e^{-A^2/2\sigma_{\Pi}^2} + \frac{1}{2} \frac{A \cos \phi}{\sqrt{2\pi\sigma_{\Pi}^2}} \left[1 - \Phi\left(\frac{A \cos \phi}{\sqrt{2\pi\sigma_{\Pi}^2}}\right) \right] e^{-\frac{A^2 \sin^2 \phi}{2\sigma_{\Pi}^2}}, \quad (3.28)$$

де $\Phi(x)$ - Функція Крампа.

При некогерентному прийомі використовуються різні типи детекторів, які прийнято класифікувати за їхніми характеристиками: лінійний детектор, квадратичний і т.п.

Розглянемо вплив синусоїдального сигналу і нормального вузькосмугового шуму на амплітудний детектор і визначимо відношення сигналу до шуму на його виході. Зазначене завдання вирішується за допомогою добре розробленого кореляційного методу. Цей метод є досить загальним та зводиться до наступного. Для заданої характеристики детектора $y = f(x)$ визначається середнє значення y (струму або напруги):

$$\overline{y(t)} = \int_{-\infty}^{\infty} f(s + \omega) p(\omega) d\omega. \quad (3.29)$$

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

$$y_0 = \overline{\overline{y(t)}} = \overline{\int_{-\infty}^{\infty} f(s + \omega) p(\omega) d\omega}. \quad (3.30)$$

Флуктуаційна складова на виході детектора визначається різницею $y - \overline{y} = \xi$, а регулярна змінна складова $b = \overline{y} - y_0$. Під сигналом зазвичай розуміється низькочастотна складова величини b або прирошування величини y_0 під час подачі на вхід сигналу. Таким чином, коливання на виході детектора можна подати у вигляді суми

$$y = y_0 + b + \xi. \quad (3.31)$$

Перше доданок є постійну складову, друге – регулярну складову (сигнал) і третє – перешкоду на виході детектора. Перевищення сигналу над перешкодою можна визначити як відношення b_2 до дисперсії перешкоди $D\xi$.

$$q_{\text{вих}} = \frac{b_2}{D\xi}. \quad (3.32)$$

Нехай характеристика детектора є квадратичною $y = x^2$. На детектор впливає сигнал (3.22) та перешкода (3.23). Тоді коливання на виході детектора буде

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

Високочастотні складові $y_{\text{вч}}$ відфільтровується і тому можуть не братися до уваги. На виході фільтра нижніх частот має коливання:

$$y_{\text{нч}}(t) = \frac{A_0^2}{2} + \frac{U^2(t)}{2} + U(t) A_0 \cos \phi(t) = b + \xi. \quad (3.33)$$

Перший доданок $b = \frac{A_0^2}{2}$ визначає сигнал, друге та третє $\xi = \frac{U^2}{2} + U A_0 \cos \phi$ – перешкоду на виході. Визначимо дисперсію шуму на виході детектора

$$D\xi = (\xi - \bar{\xi})^2 = \sigma_{\Pi}^4 + A_0^2 \sigma_{\Pi}^2. \quad (3.34)$$

Тут враховано, що $\overline{U^2(t)} = 2\sigma_{\Pi}^2$ і $\overline{U(t) \cos \phi(t)} = 0$. Відношення сигналу до перешкод на виході квадратичного детектора

$$q_{\text{вих}} = \frac{b_2}{D\xi} = \frac{A_0^4}{4(\sigma_{\Pi}^4 + A_0^2 \sigma_{\Pi}^2)} = \frac{q_{\text{вх}}^2}{1 - 2q_{\text{вх}}} \quad (3.35)$$

Де $q_{\text{вх}} = \frac{A_0^2}{2\sigma_{\Pi}^2}$. Відношення сигналу до перешкоди на вході детектора.

При значеннях $q_{\text{вх}} \gg 1$

$$q_{\text{вих}} \approx \frac{1}{2} q_{\text{вх}},$$

а при малих значення $q_{\text{вх}} \ll 1$,

$$q_{\text{вих}} \approx q^2_{\text{вх}}$$

Підвищення сигналу над завадою цієї області різко зменшується $q_{\text{вх}}$. Тут має місце пригнічення слабкого сигналу на заваді. Так, при $q_{\text{вх}} = 0,1$, ставлення $q_{\text{вих}} \approx 0,01$ - Придушення вдесятеро.

Аналогічно вирішується завдання для лінійного та інших видів детекторів. Не зупиняйся на їхньому аналізі, зазначимо, що при слабких сигналах будь-який амплітудний детектор (крім синхронного) є квадратичним.

При когерентному (синхронному) детектуванні опорного коливання $s_0(t)$ вибираємо подібним очікуваним сигналом:

$$s_0(t) = B_0 \cos \omega_0 t. \quad (3.38)$$

На виході детектора отримуємо добуток двох коливань

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

Відповідно до (3.22), (3.23) та (3.38) маємо

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

Звідси

$$y_{\text{нч}}(t) = \frac{A_0 B_0}{2} + \frac{1}{2} B_0 U \cos \phi = b + \xi, \quad (3.39)$$

де

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

Дисперсія шуму на виході

$$D\xi = \frac{1}{4} \overline{B_0^2 U^2 \cos^2 \phi} = \frac{1}{8} B_0^2 \overline{U^2} = \frac{1}{4} B_0^2 \sigma_{\Pi}^2. \quad (3.40)$$

Перевищення сигналу над перешкодою на виході синхронного детектора

$$q_{\text{вх}} = \frac{b^2}{D\xi} = \frac{A_0^2}{\sigma_{\Pi}^2} = 2q_{\text{вх}}. \quad (3.41)$$

Отримане співвідношення показує, що з когерентному детектуванні ставлення сигналу до перешкоді виході лінійно залежить від відношення сигналу до перешкоді вхіді. Явлення придушення сигналу на заваді відсутнє. Лінійні властивості когерентного приймача визначає його високі показники щодо завадостійкості. Переваги когерентного детектора в порівнянні з амплітудним особливо помітно при малих значеннях $q_{\text{вх}}$.

Так, якщо $q_{\text{вх}} = 0,1$, то при квадратичному детектуванні $q_{\text{вх}} = 0,2$. Виграш $q_{\text{кг}}/q_{\text{кв}} = 20$.

Хоча наведений вище розгляд відноситься до гармонійного сигналу, отримані висновки можуть бути повністю поширені і модульовані (маніпульовані) сигнали.

3.6. Прийом на узгоджений фільтр

Існує великий клас завдань, у яких потрібно виявити сигнал, якщо його відома. До них належить завдання прийому телеграфних сигналів, сигналів при імпульсно-кодової модуляції, радіолокаційних сигналів. У цих випадках важливим параметром, що характеризує якість виявлення, є відношення сигналу до перешкод. Лінійний фільтр, що максимізує це відношення, називається оптимальним узгодженим фільтром.

Нехай на вході фільтра діє сума сигнал $s(t)$ та перешкоди $\omega(t)$, тобто. вагання $x(t) = s(t) + \omega(t)$. Корисний сигнал $s(t)$ розглядається не як випадковий процес, а як функція відомої форми зі спектральною густиною $G_{\Pi}(\omega) = \frac{1}{2}N_0$.

Коефіцієнт передачі лінійного фільтра запишемо у вигляді

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

Сигнал на виході фільтра, очевидно, дорівнює сумі корисного сигналу $y_c(t)$ та перешкоди $y_{\Pi}(t)$:

$$y(t) = y_c(t) + y_{\Pi}(t).$$

Корисний сигнал на виході можна записати як

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

Пікова потужність сигналу у певний момент t_0 дорівнюватиме:

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

а потужність перешкоди

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

Тоді перевищення сигналу над перешкодою на момент часу t_0 буде визначатися наступним виразом:

$$q = \frac{P_{\text{ст}}}{P_{\Pi}} = \frac{\left| \int_{-\infty}^{\infty} S(i\omega) K(i\omega) e^{i\omega t_0} d\omega \right|^2}{\pi N_0 \int_{-\infty}^{\infty} K^2(\omega) d\omega}. \quad (3.43)$$

Необхідно знайти, яким має бути коефіцієнт передачі фільтра, щоб відношення сигналу до перешкод q з його виході було максимальним. На підставі нерівності Буняковського-Шварца, яка має такий вигляд:

$$\left| \int f_1(x) f_2(x) dx \right|^2 \leq \int |f_1(x)|^2 dx \int |f_2(x)|^2 dx. \quad (3.44)$$

отримуємо

$$q \leq \frac{\int_{-\infty}^{\infty} |S(i\omega)|^2 d\omega \int_{-\infty}^{\infty} |K(i\omega)|^2 d\omega}{\pi N_0 \int_{-\infty}^{\infty} |K(i\omega)|^2 d\omega} = \frac{1}{\pi N_0} \int_{-\infty}^{\infty} |S(i\omega)|^2 d\omega. \quad (3.45)$$

Таким чином, за будь-якої характеристики фільтра $K(i\omega)$ відношення сигналу до перешкоди не може перевищувати максимального значення

$$q_{\max} = \frac{1}{\pi N_0} \int_{-\infty}^{\infty} |S(i\omega)|^2 d\omega = \frac{2E}{N_0} \quad (3.46)$$

де E - Повна енергія сигналу. Вказана максимальна величина q досягається в тому випадку, коли коефіцієнт передачі фільтра має такий вираз:

$$K(i\omega) = cS(-i\omega)e^{-i\omega t_0} = cS(\omega)e^{-i[\omega t_0 + \phi(\omega)]}, \quad (3.47)$$

де $S(-i\omega) = S(\omega)e^{-i\phi(\omega)}$ - функція, комплексно-пов'язана із спектром сигналу $S(i\omega)$, c - довільна стала. У цьому неважко переконатися шляхом безпосередньої підстановки виразу (6.47) у рівність (6.43).

Вираз (6.47) можна записати у вигляді двох рівностей:

$$\begin{aligned} K(i\omega) &= cS(\omega) \\ \phi(\omega) &= -[\phi(\omega) + \omega t_0], \end{aligned} \quad (3.48)$$

з яких випливає, що амплітудно-частотна характеристика узгодженого фільтра з точністю до постійного множника збігається з амплітудним спектром сигналу, а фазо-частотна характеристика визначається фазовим спектром сигналу $\phi(\omega)$ та лінійною функцією частоти ωt_0 . Таким чином, частотна характеристика оптимального фільтра повністю визначається спектром сигналу, "узгоджена" з ним. Звідси і назва - узгоджений фільтр.

Фаза сигналу на виході узгодженого фільтра дорівнює:

$$\theta(t) = \omega t + \phi(\omega) + \psi(\omega) = \omega t + \phi(\omega) - \phi(\omega) - \omega t_0 = \omega(t - t_0). \quad (3.49)$$

При $t = t_0, \theta = 0$, тобто, в момент t_0 всі гармонійні складові сигналу мають однакову фазу і складаються арифметично, утворюючи в цей момент пік сигналу на виході фільтра. Спектральні складові перешкоди на виході фільтра мають випадкову фазу. Цим пояснюється доведене вище положення про те, що узгоджений фільтр максимізує ставлення сигналу до перешкоди на виході.

Імпульсний відгук узгодженого фільтра легко визначається виходячи з перетворення Фур'є. Згідно (6.47) маємо

$$\begin{aligned}
 g(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} K(i\omega) e^{i\omega t} d\omega = \frac{c}{2\pi} \int_{-\infty}^{\infty} S(-i\omega) e^{-i\omega(t_0-t)} d\omega = \\
 &= \frac{c}{2\pi} \int_{-\infty}^{\infty} S(i\omega) e^{i\omega(t_0-t)} d\omega = cs(t - t_0).
 \end{aligned}
 \tag{3.50}$$

Отже, імпульсним відгуком узгодженого фільтра є дзеркальне відображення сигналу щодо t_0 у масштабі «с» (рис. 6.7). З рис. 6.7 видно, що t_0 не може бути менше моменту закінчення сигналу T . Це означає, що для практично реалізованого фільтра повинна виконуватися умова $t_0 \geq T$. Для зменшення часу аналізу доцільно прийняти $t_0 = T$.

Узгоджений фільтр як корелятор. Нехай $u_{\text{вх}}(t)$ - деякий вхідний сигнал, що у загальному випадку не збігається з сигналом $s(t)$. По відношенню до якого аналізований лінійний фільтр є узгодженим. Відгук фільтра на цей вхідний вплив

$$\begin{aligned}
 u_{\text{вх}}(t) &= \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) y_{\text{согл}}(t - \tau) d\tau = \\
 &= k \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) s[t - (t - \tau)] d\tau = \\
 &= k \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) s[\tau - (t - t_0)] d\tau.
 \end{aligned}
 \tag{3.51}$$

Легко помітити, що останній інтервал є взаємкореляційною функцією сигналів $u_{\text{вх}}(t)$ і $s(t)$, тобто.

$$u_{\text{вх}}(t) = kB_{\text{us}}(t - t_0) \tag{3.51}$$

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

$$u_{\text{вх}}(t) = k \int_{-\infty}^{\infty} u_{\text{вх}}(\tau) s(\tau) d\tau. \tag{3.52}$$

Припустимо тепер, що $u_{\text{вх}} = S(t)$, тобто. на вході фільтра є сигнал, стосовно якого цей фільтр узгоджений. З формули (6.51) випливає, що в цьому випадку

$$S_{\text{ввх}}(t) = kB_s(t - \tau_0), \tag{3.53}$$

тобто, вихідний сигнал пропорційний автокореляційної функції вхідного сигналу. Як приклад на рис. 6.8 зображено побудову сигналу на виході конкретного узгодженого фільтра.

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

$$s(\omega) = \int_{-\infty}^{\infty} s(t) e^{-j\omega t} dt = U_0 \int_0^{\tau_n} e^{-j\omega t} dt = \frac{U_0}{j\omega} (1 - e^{-j\omega \tau_n}). \quad (3.53)$$

Звідси знаходимо частотний коефіцієнт передачі узгодженого фільтра, поклавши для конкретності $t_0 = \tau_n$, тобто. що відгук фільтра максимальний в момент закінчення імпульсу:

$$K_{\text{согл}}(j\omega) = k \frac{1 - e^{-j\omega \tau_n}}{-j\omega} e^{-j\omega \tau_n} = \frac{k}{j\omega} (1 - e^{-j\omega \tau_n}). \quad (3.54)$$

Отриманий результат дозволяє синтезувати узгоджений фільтр. Дійсно, відповідно до виразу (3.54) такий фільтр повинен являти собою каскадне з'єднання трьох лінійних ланок: а) масштабного підсилювача з коефіцієнтом посилення k ; б) ідеального інтегратора; в) пристрої з коефіцієнтом передачі $K'(j\omega) = 1 - \exp(-j\omega \tau_n)$. Останній пристрій реалізується за допомогою ланки затримки сигналу на час τ_n , інвертора, що змінює знак сигналу, та суматора. Структурну схему фільтра зображено на (рис. 3.9.)

Узгоджений 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_n кожен; інтервал між імпульсами дорівнює T . Якщо $S_0(\omega)$ - спектральна щільність окремого імпульсу, то спектральна щільність пачки імпульсів

$$S_{\Pi}(\omega) = S_0(\omega) [1 + e^{-j\omega T} + e^{-j2\omega T} + \dots + e^{-j(N-1)\omega T}]. \quad (3.55)$$

Синтезуючи структуру узгодженого фільтра для пачки імпульсів, потрібно, щоб максимальний відгук виникав у момент закінчення останнього імпульсу пачки, звідки $t_0 = (N - 1)T + \tau_n$. Знаходимо частотний коефіцієнт передачі узгодженого фільтра:

$$K_{\text{согл}}(j\omega) = kS_0(\omega)e^{-j\omega\tau}[1 + e^{j\omega T} + e^{j2\omega T} + \dots + e^{j(N-1)\omega T}]e^{j(N-1)\omega T} = \\ = K_{0\text{ согл}}(j\omega)[1 + e^{-j\omega T} + e^{-j2\omega T} + \dots + e^{-j(N-1)\omega T}], \quad (3.56)$$

де $K_{0\text{ согл}}(j\omega)$ - коефіцієнт передачі узгодженого фільтра для одиночного відеоімпульсу.

Формула (3.56) безпосередньо визначає структурну схему узгодженого фільтра, зображену на (рис. 3.10).

На вході розміщено узгоджений фільтр для одиночного відеоімпульсу. Основою пристрою є багатовідвідна лінія затримки, що забезпечує запізнення сигналів на відрізок часу $T, 2T, \dots, (N-1)T$. Сигнали з усіх відводів надходять до суматора. Легко бачити, що максимальний відгук на виході суматора буде спостерігатися тоді, коли корисні сигнали від усіх імпульсів пачки одночасно виявляться на його входах. Ефективність роботи пристрою всіх його входах. Ефективність роботи пристрою тим вища, чим довша пачка.

Практично виконувані виявлячі сигналів радіолокаційних містять також спеціальний нелінійний пороговий елемент, вхід якого з'єднаний з виходом суматора узгодженого фільтра. Рівень порога дещо перевищує середньоквадратичне значення шуму в отауті.

Якщо сплеск вихідного сигналу фільтра досягає порогового рівня, пристрій індикації надходить управляючий сигнал, що свідчить про наявність імпульсу, відбитого від мети.

Узгоджений фільтр прямокутного радіоімпульсу. Нехай виділений сигнал є радіоімпульсом виду

$$s(t) = \{0, t < 0, | \{U_0 \sin \omega_0 t, 0 \leq t \leq \tau_n, | \quad (3.57)$$

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

Як було показано, імпульсна характеристика узгодженого фільтра $q_{\text{согл}}(t) = k s(t_0 - t)$. Покладемо $t_0 = \tau_n$ і будемо вважати для простоти тривалість імпульсу кратною періоду високочастотного заповнення, тож $\sin \omega_0 \tau_n = 0$.

Тоді

$$q_{\text{согл}}(t) = \{0, t < 0, | \quad \{k \sin \omega_0 t, 0 \leq t \leq \tau_n, | \quad (3.58)$$

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

Таку імпульсну характеристику можна приблизно реалізувати за допомогою системи, структурна схема якої наведена на (рис. 3.11.)

На вході фільтра розміщується коливальна ланка (наприклад, високочастотний коливальний контур) з імпульсною характеристикою

$$q_{\text{согл}}(t) = \{0, t < 0, |$$

де b - постійна величина.

Для того щоб імпульсна характеристика узгодженого фільтра дорівнювала нулю при $t > \tau_n$, передбачені суматор, на один із входів якого сигнал з виходу коливальної ланки подається безпосередньо, а на інший – через ланку затримки на τ_n секунд, і фазообертач, що змінює фазу сигналу на 180° . При такому включенні елементів починаючи з часу $t = \tau_n$ до входів суматора додані два гармонійних коливання з однаковими амплітудами та протилежними фазами, що перетворює на нуль сигнал на виході суматора.

Узгоджений фільтр для сигналу Баркер. Гідність сигналів Баркера – високе значення головної пелюстки автокореляційної функції та гранично низький рівень бічних пелюсток.

На рис.6.12 зображено структурну схему узгодженого фільтра, призначеного для виявлення M – позиційного сигналу Баркера з фазовим кодуванням. Такий сигнал має вигляд послідовності відрізків гармонійних коливань із фазовими зсувами $\phi_1, \phi_2, \dots, \phi_M$, рівними 0 або 180° .

При синтезі виходять з того, що імпульсна характеристика узгодженого фільтра повинна являти собою «дзеркальну» копію сигналу, що виділяється з зверненим у часі порядком проходження окремих позицій.

На вході пристрою є допоміжний фільтр Φ_1 . Узгоджений стосовно однієї позиції складного фазоманіпульованого сигналу, тобто. до прямокутного радіоімпульсу. На виході цього фільтра під впливом вхідного дельта-імпульсу виникає радіоімпульс з прямокутної форми, що огинає. Цей імпульс подається на лінію затримки з відводами, що є зазвичай хвильовою (розподіленою) системою. Затримка у часі між відведеннями дорівнює тривалості T кожної позиції сигналу.

Для правильного функціонування пристрою необхідно, щоб послідовність фазових зсувів $\phi_M, \phi_{M-1}, \dots, \phi_1$ (див. рис. 3.12) відповідала значенням фаз в окремих позиціях сигналу Баркера за рахунку від кінця сигналу на початок.

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

Контрольні питання

1. Які основні завдання виникають під час прийому сигналів?
2. У чому полягає суть прийому сигналів із накопиченням?
3. Які різновиди способу накопичення Ви знаєте?
4. У чому полягає суть кореляційного прийому?
5. Чим автокореляційний метод прийому відрізняється від кореляційного?
6. Чим когерентний метод прийому відрізняється від некогерентного?
7. Як змінює відношення сигнал/шум на виході детектора за різних значень відношення сигнал/шум на виході?
8. Який фільтр називається узгодженим?
9. Яке максимальне відношення сигнал/шум матиме вихід узгодженого фільтра?
10. Намалюйте структурну схему узгодженого фільтра для одиночного відеоімпульсу?
11. Як виглядає узгоджений фільтр для прямокутного радіоімпульсу?
12. Намалюйте структурну схему узгодженого фільтра для Баркера?

РОЗДІЛ 4. РОЗРОБКА АЛГОРИТМІВ ОПТИМАЛЬНОГО ПРИЙОМУ СИГНАЛІВ ІЗ ФАЗОРІЗНИЦЕВОЮ МОДУЛЯЦІЄЮ ВИСОКОЇ КРАТНОСТІ

Постановка задачі

Звичайно, коли досліджуються сигнали з фазорізнцевою модуляцією і методи прийому цих сигналів, то це відноситься до часткового випадку ФРМ, названому фазорізнцевою модуляцією першого чи порядку кратному ФРМ-1. Підкреслимо, що мова тут йде не про позиційність системи (не про кратність модуляції), обумовленої числом дискретних значень чи сигналу його інформаційного параметра в каналі зв'язку, а про порядок різностей фази, використовуваних як інформаційний параметр.

У розділі вводиться поняття порядку різниці фази сигналу. Показано, що перехід до ФРМ-2 дозволяє досягти повної нечутливості не тільки до довільної початкової фази, але і до довільних зрушень частоти.

У розділі розглядаються принципи побудови багатопозиційних фазомодульованих сигналів і сигналів з комбінованими способами модуляції.

Досліджуються загальні алгоритми формування й обробки багатопозиційних сигналів з фазорізнцевою модуляцією.

Показано, що властивості ФРМ, які називають властивостями інваріантності до відповідних заважаючих впливів, є насправді наслідком фундаментальних властивостей кінцевих різностей функцій, що відображають реальні сигнали.

Розроблено загальні алгоритми формування й обробки сигналів з фазорізнцевою модуляцією k -го порядку.

Представлено відповідні структурні схеми.

Досліджено систему інваріантну до адитивної перешкоди у виді гармонійного коливання з випадковими амплітудою, частотою і фазою.

4.1. Фазорізницева модуляція першого та другого порядків та властивості інваріантності

Поняття порядку різниць фази сигналу вводиться наступним чином. Нехай ми маємо послідовність посилок, що передаються за допомогою гармонійного ФМ сигналу з початковими фазами:

$$\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_{n-1}, \varphi, \varphi_{n+1}, \dots \quad (4.1)$$

Складемо різниці фаз між кожною парою сусідніх посилок:

$$\begin{aligned} \Delta_1^1 \varphi &= \varphi_1 - \varphi_0, \\ \Delta_2^1 \varphi &= \varphi_2 - \varphi_1, \\ &\dots\dots\dots \\ \Delta_n^1 \varphi &= \varphi_n - \varphi_{n-1}, \\ \Delta_{n+1}^1 \varphi &= \varphi_{n+1} - \varphi_n, \\ &\dots\dots\dots \end{aligned} \quad (4.2)$$

Різниці фаз (4.2) називають ще різницями фаз першого порядку (чи просто першими різницями фаз), тому що вони отримані з вихідної послідовності фаз (4.1) шляхом однократної операції віднімання. Ця обставина підкреслюється верхнім індексом 1 при операторі обчислення різниці Δ . Таким чином, позначення $\Delta_n^1 \varphi$ показує, що мова йде про розрахунок різниці фаз першого порядку між n - та $(n - 1)$ посилками сигналу. Послідовність різниць фаз першого порядку, так як і послідовність початкових фаз, розвертається в часі в міру передачі посилок сигналу:

$$\Delta_0^1 \varphi, \Delta_1^1 \varphi, \Delta_2^1 \varphi, \dots, \Delta_{n-1}^1 \varphi, \Delta_n^1 \varphi, \Delta_{n+1}^1 \varphi, \dots \quad (4.3)$$

З цієї послідовності чисел можна скласти нові різниці за тим же правилом, за яким з (4.1) складались (4.2):

$$\begin{aligned}
\Delta_1^2 \varphi &= \Delta_1^1 \varphi - \Delta_0^1 \varphi , \\
\Delta_2^2 \varphi &= \Delta_2^1 \varphi - \Delta_1^1 \varphi , \\
&\dots\dots\dots \\
\Delta_n^2 \varphi &= \Delta_n^1 \varphi - \Delta_{n-1}^1 \varphi , \\
\Delta_{n+1}^2 \varphi &= \Delta_{n+1}^1 \varphi - \Delta_n^1 \varphi , \\
&\dots\dots\dots
\end{aligned}
\tag{4.4}$$

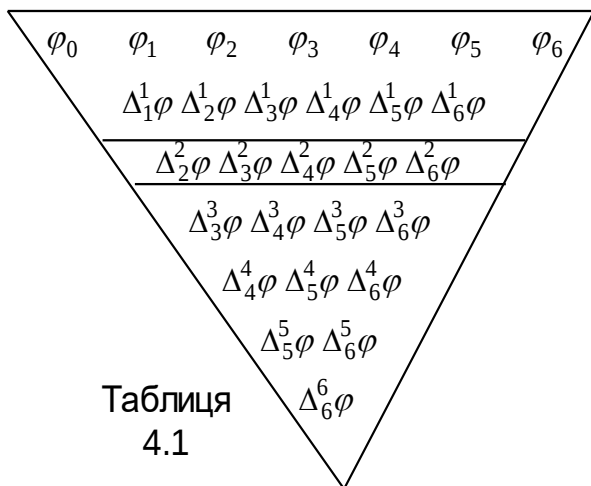
Різниці фаз (4.4) називають різницями другого порядку (чи просто другими різницями фаз), тому що вони отримані з вихідної послідовності фаз (4.1) шляхом двократного застосування операції віднімання. Ця обставина підкреслюється верхнім індексом 2 при операторі віднімання Δ . Різниці фази другого порядку утворюють часову послідовність, аналогічну послідовностям (4.1) і (4.3):

$$\Delta_0^2 \varphi, \Delta_1^2 \varphi, \Delta_2^2 \varphi, \dots, \Delta_{n-1}^2 \varphi, \Delta_n^2 \varphi, \Delta_{n+1}^2 \varphi, \dots
\tag{4.5}$$

Можна й надалі продовжувати обчислення різниць фаз третього та більш високих порядків. Процес формування різниць фази високого порядку ілюструється в табл.4.1. Кожен елемент цієї таблиці дорівнює різниці двох сусідніх елементів рядку, розміщеного вище [95].

Зосередимо однак увагу на використанні різниць фаз перших двох порядків.

Зараз можна визначити введену раніше ФРМ наступним чином:



Таблиця
4.1

фазорізницевою модуляцією першого порядку називається спосіб формування фазомодульованого сигналу, при якому інформація вкладається в значення різниць першого порядку початкової фази посилок. Інформаційним параметром сигналу при ФРМ-1 є різниця фаз, що визначається двома посилками сигналу:

$$\Delta_n^1\varphi = \varphi_n - \varphi_{n-1}, \quad (4.6)$$

звідки випливає, що початкова фаза чергової n -посилки переданого в канал зв'язку сигналу:

$$\varphi_n = \varphi_{n-1} + \Delta_n^1\varphi \quad (4.7)$$

Вирази (4.7) і (4.6) визначають загальний алгоритм формування й обробки фазомодульованого сигналу при ФРМ-1, який ілюструється схемою, приведеною на рис.4.1,а. На передаючій стороні системи цифрової передачі дискретному інформаційному символу J_n ставиться у відповідність одне з припустимих значень

різниці фаз першого порядку $\Delta_n^1\varphi$, потім за допомогою елемента затримки на посилку та суматора відповідно до (4.7) формується початкова фаза наступної n -ї посилки сигналу, який передається. На прийомній стороні після вимірювання початкових фаз

двох сусідніх посилок сигналу за допомогою елементу затримки та пристрою віднімання обчислюється інформаційна різниця фаз $\Delta_n^1 \varphi$, яка ототожнюється з переданим дискретним символом.

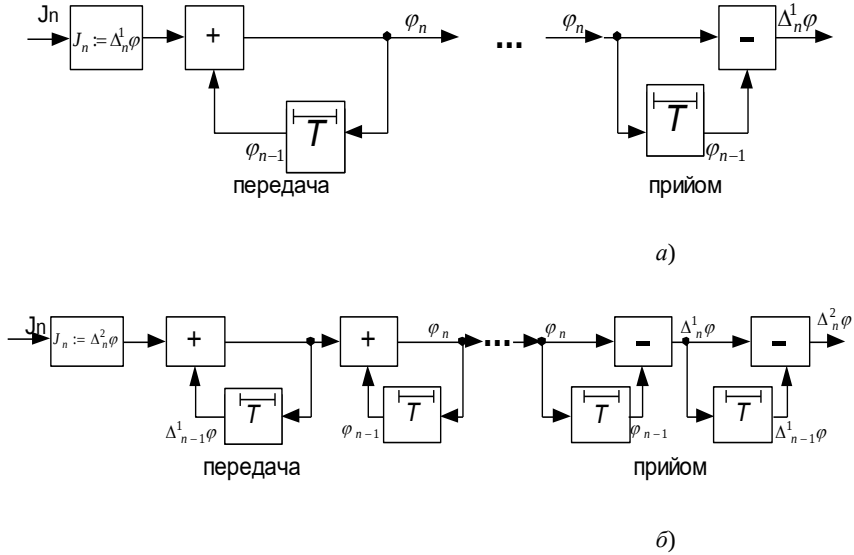


Рис.4.1,а,б. Принципи формування і обробки фазомодульованих сигналів:

а) ФРМ-1; б) ФРМ-2

Визначимо фазорізницеву модуляцію другого порядку (ФРМ-2) як спосіб формування ФМ сигналу, при якому інформація вкладається в значення різниць другого порядку початкової фази посилок сигналу. Інформаційним параметром сигналу при ФРМ-2 є різниця між різницями фаз, що визначається трьома посилками:

$$\Delta_n^2 \varphi = \Delta_n^1 \varphi - \Delta_{n-1}^1 \varphi = (\varphi_n - \varphi_{n-1}) - (\varphi_{n-1} - \varphi_{n-2}) = \varphi_n - 2\varphi_{n-1} + \varphi_{n-2} \quad (4.8)$$

З (4.8) випливає, що початкова фаза наступної n -ї послілки переданого в канал зв'язку сигналу:

$$\varphi_n = \Delta_n^2 \varphi + 2\varphi_{n-1} - \varphi_{n-2}. \quad (4.9)$$

Її також можна представити аналогічно (4.7) у вигляді двох рекурентних співвідношень:

$$\left. \begin{aligned} \varphi_n &= \varphi_{n-1} + \Delta_n^1 \varphi; \\ \Delta_n^1 \varphi &= \Delta_{n-1}^1 \varphi + \Delta_n^2 \varphi \end{aligned} \right\}. \quad (4.10)$$

Формувач початкових фаз посилок сигналу при ФРМ-2, який працює за алгоритмом (4.10), складається з двох послідовно підключених формувачів фази при ФРМ-1, а пристрій обробки – з двох послідовно підключених обчислювачів різниць фази першого порядку (рис.4.1,б).

Звернемося до основної властивості ФРМ-1 і ФРМ-2 - властивості інваріантності. У техніці цим терміном позначають нечутливість системи до тих чи інших зовнішніх дій або дестабілізуючих факторів. При цьому деяка характеристика чи параметр системи є інваріантом перетворень, які викликаються дією завад чи дестабілізуючими факторами [94].

За допомогою ФРМ-1, як уже відзначалося, вдається усунути неоднозначність рішення на виході демодулятора, викликану невизначеністю початкової фази прийнятого сигналу. Це пояснюється тим, що різниця фази першого порядку є інваріантом перетворення, яке полягає в додаванні до інформаційних фаз посилок сигналу довільної загальної початкової фази. Дійсно, якщо до інформаційних фаз $(n-1)$ -ї та n -ї посилок сигналу φ_{n-1} і φ_n додалася довільна і невідома в місці прийому початкова фаза φ_n , то різниця фаз $(n-1)$ -ї та n -ї посилок ніяк не зміниться від такого перетворення:

$$\Delta_n \varphi = (\varphi_n + \varphi_0) - (\varphi_{n-1} + \varphi_0) = \varphi_n - \varphi_{n-1} = \text{in var } \varphi_0. \quad (4.11)$$

Таким чином, фазорізницева модуляція першого порядку інваріантна до початкової фази сигналу. Важлива позитивна сторона ФРМ-1 полягає ще й в тому, що зазначена властивість інваріантності фактично зберігається при всіх методах прийому відповідних сигналів. При оптимальному некогерентному й автокореляційному прийомі не потрібно ніякої інформації про початкову фазу, а при когерентному прийомі початкова фаза повинна бути відома з точністю до фіксованих зсувів, які залежать від кратності модуляції, наприклад з точністю до 180° при однократній ФРМ-1, що цілком можливо здійснити. Найпростіші схеми модемів з ФРМ-1, які реалізують властивість інваріантності до початкової фази сигналу, приведені в попередньому параграфі. Можна сказати, що розглянуті некогерентні демодулятори сигналів з ФРМ-1 абсолютно інваріантні, тобто цілком нечутливі до початкової фази сигналу.

Перехід до ФРМ-2 дозволяє досягнути повної нечутливості не тільки до довільної початкової фази, але й до довільних зсувів частоти. Це пояснюється тим, що різниця фази другого порядку є інваріантом перетворення, яке полягає в довільному зсуві частоти несучого коливання. Дійсно, припустимо, що номінальна частота несучого коливання $\omega = m2\pi/T$, де m — ціле число, змінилася на довільну величину $\Delta\omega$. Тоді якщо фаза $(n - 1)$ -ї посилки дорівнює $(\varphi_{n-1} + \varphi_0)$, то фаза n -ї посилки стане рівною $(\varphi_{n-1} + \varphi_0 + \Delta\omega T)$, а фаза $(n + 1)$ -ї посилки $(\varphi_{n-1} + \varphi_0 + 2\Delta\omega T)$. Неважко тепер побачити, що різниці фази першого порядку

$$\Delta_{n+1}^1 \varphi = \varphi_{n+1} - \varphi_n + \Delta\omega T; \quad \Delta_n^1 \varphi = \varphi_n - \varphi_{n-1} + \Delta\omega T$$

не залежать від початкової фази φ_0 , однак залежать від зсуву частоти $\Delta\omega$, у той час як різниця фази другого порядку не залежить ні від φ_0 , ні від $\Delta\omega$:

$$\Delta_{n+1}^2 \varphi = \Delta_{n+1}^1 \varphi - \Delta_n^1 \varphi = \varphi_{n+1} - 2\varphi_n + \varphi_{n-1} = \text{in var}(\varphi_0, \Delta\omega). \quad (4.12)$$

Таким чином, фазорізницева модуляція другого порядку інваріантна до частоти несучого коливання. Ця унікальна властивість ФРМ-2 значно розширює можливості систем цифрового зв'язку з фазомодульованими сигналами. При абсолютній ФМ прийом сигналів можливий тільки при точно відомій початковій фазі, що надзвичайно обмежує сферу практичного використання абсолютної фазової модуляції. При ФРМ-1 з'являється можливість прийому фазомодульованих сигналів з невизначеною початковою фазою. При ФРМ-2 до цього додається можливість прийому фазомодульованих сигналів з невизначеною частотою несучого коливання.

Сигнали з ФРМ-2 так як і сигнали з ФРМ-1, можна приймати за допомогою алгоритмів когерентного, оптимального некогерентного й автокореляційного прийому. При перших двох методах властивість інваріантності до частоти, природно, не реалізується, оскільки когерентний і оптимальний некогерентний методи прийому досягають своїх потенційних можливостей лише при точно відомій частоті несучого коливання, а при відхиленнях частоти сигналу від частоти опорних коливань демодуляторів швидко втрачають працездатність. Важливою перевагою ФРМ-2 є те, що при відомій частоті сигналу вона володіє такою ж завадостійкістю, як ФРМ-1 при когерентному прийомі, і більш високою завадостійкістю, ніж ФРМ-1 при оптимальному некогерентному прийомі.

Властивість абсолютної інваріантності до частоти несучого коливання досягається при автокореляційному прийомі сигналів із ФРМ-2. У найпростішому випадку однократної ФРМ-2 із другими різницями фази $\Delta^2\varphi_1=0$, $\Delta^2\varphi_2=\pi$ алгоритм роботи відповідного демодулятора можна одержати, представивши косинус другої різниці фази прийнятого сигналу через скалярні добутки сусідніх посилок аналогічно тому, як це було зроблено в попередньому параграфі для ФРМ-1. Очевидно, що в даному випадку переданий інформаційний символ J_n визначається знаком косинуса другої різниці фаз $\Delta_n^2\varphi_\xi$ на n -й посилю:

$$j_n = \text{sgn} \cos \Delta_n^2 \varphi_\xi = \text{sgn}(\cos \Delta_n^1 \varphi_\xi \cos \Delta_{n-1}^1 \varphi_\xi + \sin \Delta_n^1 \varphi_\xi \sin \Delta_{n-1}^1 \varphi_\xi). \quad (4.13)$$

Вхідні в (4.13) косинуси і синуси перших різниць фази надаються формулами векторної алгебри аналогічно (4.14) у вигляді:

$$\left. \begin{aligned} \cos \Delta_n^1 \varphi_\xi &= \frac{(X_n X_{n-1})}{|X_n| |X_{n-1}|}, \quad \sin \Delta_n^1 \varphi_\xi = \frac{(X_n X_{n-1}^*)}{|X_n| |X_{n-1}|}; \\ (X_n X_{n-1}) &= \int_0^T x_n(t) x_{n-1}(t); \\ (X_n X_{n-1}^*) &= \int_0^T x_n(t) x_{n-1}^*(t), \end{aligned} \right\} \quad (4.14)$$

де $x_n(t)$ та $x_{n-1}(t)$ - n -а і $(n - 1)$ -а послілки сигналу на вході демодулятора. Підставивши (4.14) у (4.13), одержимо шуканий алгоритм автокореляційної обробки сигналів з однократною ФРМ-2, котрий, як неважко перевірити, абсолютно інваріантний до частоти несучого колювання. На рис.4.2 відповідний демодулятор представлений у складі найпростішої системи цифрової передачі з ФРМ-2. У цій системі є також модулятор сигналів ФРМ-2, який складається, як і модулятор сигналів ФРМ-1, з генератора несучого колювання, перемножувача та перекодовуючого пристрою. Останній містить елемент затримки двійкового символу на дві послілки і перемножувач двійкових символів ± 1 . Алгоритм роботи даного перекодовуючого пристрою, неважко отримати із співвідношення (4.9). Оскільки при однократній ФРМ-2 всі вхідні в це співвідношення фази рівні або 0, або π , операції їхнього додавання можна замінити додаванням по модулю двох чисел 0 і 1 або перемноженням чисел 1 і -1. В останньому випадку алгоритм перекодування переданих двійкових символів J_n приймає вигляд рекурентного співвідношення:

$$J_n' = J_n J_{n-2}'. \quad (4.15)$$

Алгоритм (4.15) реалізований у схемі, приведений на рис.4.2.

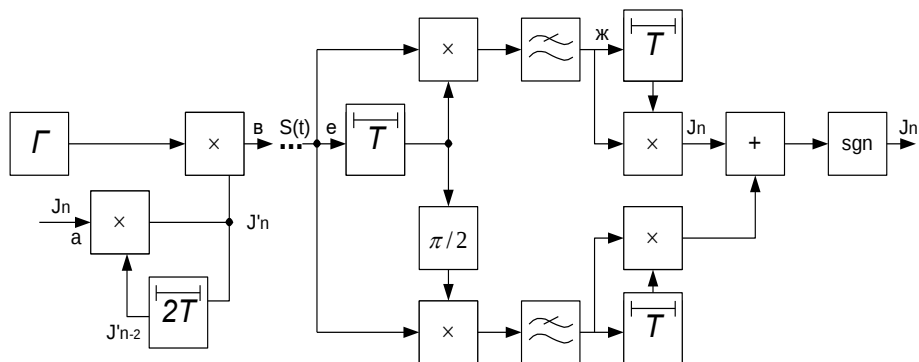


Рис.4.2. Модем системи з однократною ФРМ-2, забезпечуючий інваріантність до частоти несучого коливання

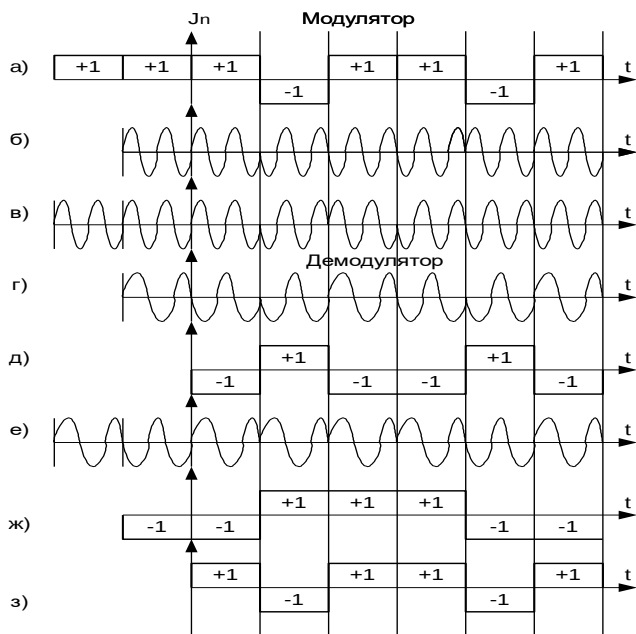


Рис.4.3. Часові діаграми сигналів при ФРМ-1 та ФРМ-2

Припустимо, що в каналі частота несучого коливання змінилася таким чином, що в посилювачі укладається не ціле число періодів несучого коливання, як було на виході модулятора, а непарне число напівперіодів. Тоді сигнали з ФРМ-1 і ФРМ-2 на вході відповідних демодуляторів будуть мати вигляд, показаний на рис.4.3, *г* і *е* відповідно.

На рис.4.3, *д* приведений вихідний сигнал автокореляційного демодулятора сигналів із ФРМ-1. Виходячи з порівняння діаграм рис.4.3, *д* і *а*, при даній зміні частоти несучого коливання автокореляційний демодулятор сигналів із ФРМ-1 видає двійкові символи, протилежні переданим, тобто поводить себе так як і когерентний демодулятор сигналів з абсолютною ФМ при невідповідності фаз прийнятого сигналу й опорного коливання [96].

Звернемося тепер до автокореляційного демодулятора сигналів із ФРМ-2. Сигнал на виході верхнього корелятора (точка *ж* на рис.4.2) представлений на діаграмі рис.4.3, *ж*. Сигнал на виході нижнього корелятора в даному випадку дорівнює нулеві, так як при обраному співвідношенні між частотою сигналу і тривалістю посилки сигнали на входах цього корелятора ортогональні. Враховуючи сказане, одержуємо, що вихідний сигнал демодулятора виходить у результаті перемножування знаків сусідніх посилок сигналу, зображеного на діаграмі рис.4.3, *ж*. Цей вихідний сигнал показаний на діаграмі рис.4.3, *з*. Як видно, він співпадає з переданим сигналом.

Така коротка процедура обробки сигналів із ФРМ-2 дозволяє демодувати їх безпомилково при невідомих зсувах несучого коливання.

4.2. Багатопозиційні фазомодульовані сигнали

Існують різні способи формування і обробки фазомодульованих сигналів, які відповідають ФМ, ФРМ-1 та ФРМ-2, ілюструвалися головним чином на прикладі двопозиційних (однократних) сигналів, в яких фаза чи різниця фаз (першого або другого порядків) приймає тільки два значення: 0 або π . Насправді ж багато які з розглянутих властивостей фазомодульованих сигналів, наприклад властивості інваріантності ФРМ-1 до початкової фази і ФРМ-2 до несучої частоти, відносяться в рівній мірі до всіх багатопозиційних сигналів.

Розглянемо принципи побудови багатопозиційних фазомодульованих сигналів і сигналів з комбінованими способами модуляції. Як правило, якщо це спеціально не обговорено, не робляться відмінності між ФМ, ФРМ-1 чи ФРМ-2. Якщо, наприклад, розглядається деяка сукупність сигнальних векторів, то вони можуть відноситися до кожного з цих способів модуляції, а варіанти фаз сигналів можуть бути в однаковій мірі варіантами різниці фаз. Це застереження необхідне для того, щоб не підкреслювати щоразу, що сигнал, який розглядається, може бути використаний і при ФМ, і при ФРМ лише з тією відмінністю, що в першому випадку відкладені на векторній діаграмі - передані абсолютні фази, а у другому - різниці фаз сусідніх посилок [115].

Принцип побудови m -позиційних фазомодульованих сигналів дуже простий: m сигнальних точок розташовуються по окружності із радіусом, залежним від потужності (або енергії посилки) сигналу на рівних відстанях з кутовим інтервалом $2\pi/m$ радіан. Такі сукупності сигнальних точок-векторів для випадків $m = 2, 3, 4, 8$ показані на рис.4.4. Радіуси всіх окружностей, на яких розташовані сигнальні точки, співпадаючи з кінцями відрізків, прирівняні до кореня квадратного з енергії посилок сигналу: $R = \sqrt{E}$.

Якщо на посилці передається гармонійне коливання з параметрами a, ω і φ , то ця величина дорівнює:

$$R = \sqrt{E} = \sqrt{\int_0^T a^2 \sin^2(\omega t + \varphi) dt} = \frac{a\sqrt{T}}{\sqrt{2}}. \quad (4.16)$$

Значення (4.16) збігається з евклідовою відстанню між центром окружності і будь-якою точкою на ній.

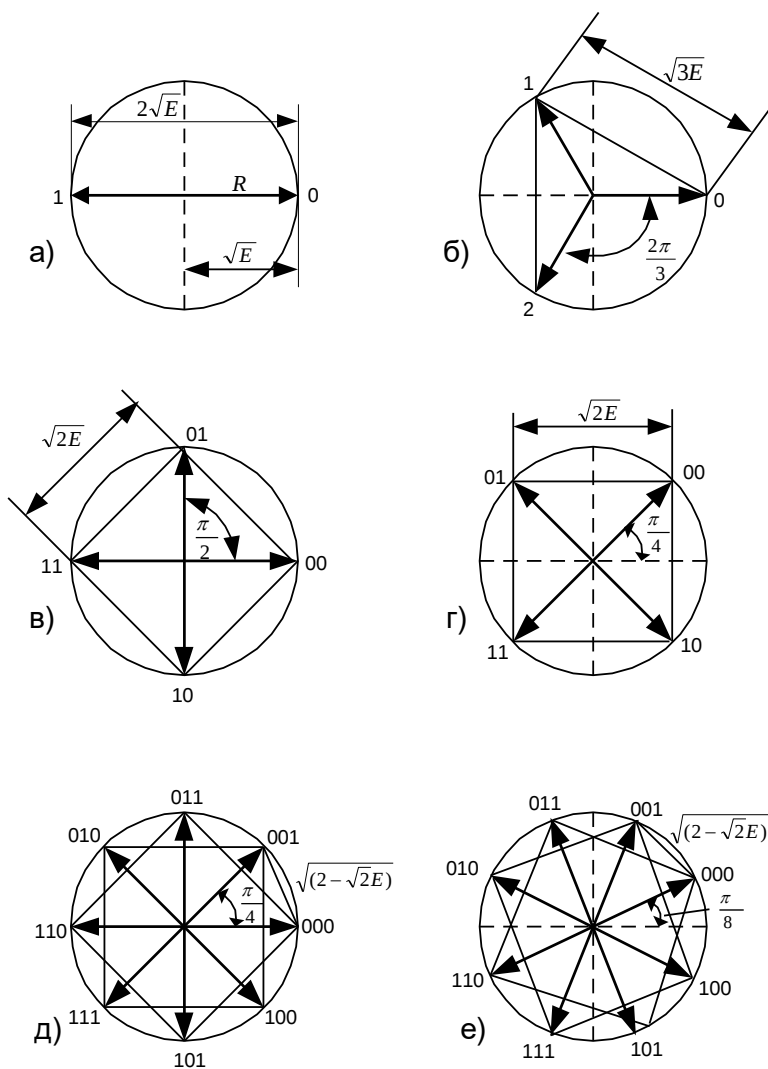


Рис.4.4. Багато позиційні фазомодульовані сигнали: а) війкові (однократні); б) трійкові; в,г) четвіркові (двократні); д,е) восьмирічні (трикратні)

На рис.4.4,а показаний двопозиційний (однократний) сигнал з фазами 0 і π .

Відстань між сигнальними точками в даному випадку $d_2 = 2\sqrt{E}$ - це максимально

можлива відстань між точками кола з радіусом $\sqrt{E}$. Вона цілком визначає потенційну завадостійкість даної двопозиційної системи. Відстань $d_2 = 2\sqrt{E}$ є відліковою величиною, з якою зіставляється відстань між варіантами сигналу в різних багато позиційних системах. Двопозиційний сигнал кодується елементарно: до одного з його варіантів ставиться у відповідність символ 0, а іншому - символ 1, як показано на рис.4.4,а.

На рис.4.4,б показаний трипозиційний ФМ сигнал з векторами, що знаходяться один від одного на 120° . Це єдина з багатопозиційних систем із ФМ, у якій відстані між усіма варіантами сигналу однакові (еквідистантні сигнали).

Відмітимо що відстань між двома гармонійними сигналами S_1 і S_2 тривалістю T , які відрізняються за фазою на φ :

$$d(S_1, S_2) = \sqrt{\int_0^T [S_1(t) - S_2(t)]^2 dt} = \sqrt{\int_0^T [a \sin(\omega t + \varphi) - a \sin \omega t]^2 dt} =$$

$$\sqrt{a^2 T - a^2 T \cos \varphi} = \sqrt{2E} \sqrt{1 - \cos \varphi},$$

(4.17)

де $E = a^2 T / 2$. У табл. 1.2 приведені розраховані за формулою (1.50) відстані d_m між ній ближніми варіантами сигналу в m -позиційних системах з ФМ і програш (за мінімальною сигнальною відстанню) m -позиційної системи двопозиційної. У загальному випадку мінімальна відстань між сигналами не є характеристикою, що однозначно визначає завадостійкість. Ймовірність похибки в ідеальному приймачі залежить від системи сигналів у цілому, а не тільки від відстані між найближчими варіантами сигналу. Однак для заданої конфігурації сигнальних точок ймовірність

похибок є монотонно спадаючою функцією мінімальної відстані. Цей параметр зручний для якісного співставлення різних систем сигналів.

Кратність модуляції N	Число фаз m	Мінімальна різниця фаз	Мінімальна евклідова відстань між сигналами d_m	d_2 / d_m , дБ
1	2	Π	$2\sqrt{E}$	0
$\log_2 3$	3	$2\pi/3$	$\sqrt{3E} \approx 1,73\sqrt{E}$	1.25
	4	$\pi/2$	$\sqrt{2E} \approx 1,41\sqrt{E}$	3.01
2	8	$\pi/4$	$\sqrt{(2 - \sqrt{2})E} \approx 0,765\sqrt{E}$	8.34
3	16	$\pi/8$	$\sqrt{(2 - \sqrt{2 + \sqrt{2}})E} \approx 0,39\sqrt{E}$	14.2
4	32	$\pi/16$	$\sqrt{(2 - \sqrt{2 + \sqrt{2 + \sqrt{2}}})E} \approx 0,196\sqrt{E}^{20.2}$	

У випадку трипозиційної системи, який розглядається, (див. рис.4.4,б) відстань між сигналами $d_3 = \sqrt{3E}$ і відрізняється від d_2 на 1.25 дБ. Тому при рівній від d_2 на 1.25 дБ. Тому при рівній тривалості посилки трипозиційна система незначно поступається двопозиційній за завадостійкістю, перевершуючи її по швидкості передачі в $\log_2 3 = 1,58$ рази. Як буде показано нижче, розглянута трипозиційна система ФМ сигналів при заданій ймовірності похибки є оптимальною по енергетичних витратах, які припадають на передачу одного біта інформації. Незважаючи на це, вона

використовується порівняно рідко через необхідність спеціального перетворення двійкового коду у трійковий.

Найбільш широке розповсюдження на практиці одержали чотирипозиційні (двократні) сигнали, показані на рис.4.4,в і г. Двократна система представлена в двох найбільш популярних варіантах: з фазами (або різницями фаз) $0, \pi/2, \pi$ і $3\pi/2$ (рис.4.4,в) або $\pi/4, 3\pi/4, 5\pi/4$ і $7\pi/4$ (рис.4.4,г). В обох випадках мінімальна відстань між варіантами сигналу $d_4 = \sqrt{2E}$, що відповідає відстані між ортогональними сигналами.

Показані чотирипозиційні сигнали є найкращими за завадостійкістю з усіх двовимірних чотирипозиційних сигналів. Крім того, ці сигнали мають чудову властивість: при однаковій швидкості передачі вони забезпечують таку ж завадостійкість, як і двопозиційні сигнали з ФМ (див. рис.4.4,а), тому що зменшення відстані між сигналами в $\sqrt{2}$ рази цілком компенсується подовженням посилки сигналу вдвічі. Кожна посилка з двократною ФМ чи ФРМ несе два біти інформації. На рис.4.4,в і г показана типова система кодування двократних сигналів дворозрядними двійковим комбінаціями. Ця система називається кодом Грея, який має ту властивість, що кодові комбінації, які відповідають найближчим фазам (або різницям фаз), відрізняються тільки одним двійковим символом. При двократній ФМ код Грея є єдиним маніпуляційним кодом, який має таку властивість. У результаті його використання помилковий прийом, як правило, приводить до спотворення тільки одного з двох переданих двійкових символів.

Відзначимо, що системи сигналів, показані на рис.4.4,в і г, мають однакову потенційну завадостійкість і відрізняються тільки реалізаційними аспектами. У системі рис.4.4,г мінімальне фазовий зсув рівний $\pi/4$; в ній завжди, тобто на кожній посилці, є той чи інший ненульовий стрибок фази на відміну від системи рис.4.4,в, у якій при передачі комбінації 00 фазовий зсув дорівнює нулеві. Це визначає деякі особливості реалізації модулятора і демодулятора, про що ще буде йти мова надалі.

На рис.4.4,д і е представлені два варіанти восьмипозиційних (трикратних) фазомодульованих сигналів. У першому варіанті фази чи різниці фаз приймають значення $(i-1)\pi/4$, де $i = 1, 2, \dots, 8$, а в другому варіанті — значення $(2i-1)\pi/8$. Перший варіант начебто поєднує сигнали двох чотирипозиційних систем (рис.4.4,в і г),

зсунутих один по відношенню до іншого на $\pi/4$. Варіант не містить нульової фази і має постійний зсув щодо першого на $\pi/8$. Кожна посилка восьмипозиційного сигналу несе три двійкові одиниці інформації. В трикратній системі, так як і в двократній, існує єдиний оптимальний маніпуляційний код, у якому трирозрядні двійкові комбінації, відповідні найближчим один до одного сигнальним векторам, відрізняються тільки в одному двійковому символі — це код Грея. На рис.4.4 комбінації маніпуляційного коду Грея проставлені у відповідних векторах.

Відстань між найближчими векторами в трикратних системах сигналів, показаних на рис.4.4, $d_8 = \sqrt{(2 - \sqrt{2})E}$, що на 8.3 дБ менше, ніж в однократній системі. Ця відстань уже не є мінімальною, як у розглянутих раніше сигналах меншої кратності.

Рівномірне розміщення всіх сигнальних точок на окружності, тобто використання рівнопотужних сигналів, які відрізняються лише фазою, є оптимальним тільки для двох-, трьох- і чотирьох-позиційних випадків (див. рис.4.4,а-г). При $m > 4$ оптимальними є нерівнопотужні сигнали, що відрізняються як фазою, так і амплітудою і розміщені рівномірно всередині окружності, радіус якої визначається максимально припустимою енергією сигналу. З точки зору теорії модуляції такі сигнали відносяться до комбінованих способів модуляції, при яких одночасно змінюються декілька параметрів сигналу - у випадку, який розглядається, амплітуда і фаза чи різниця фаз. Сигнали з амплітудно-фазовою чи амплітудно-фазорізницевою модуляцією (АФМ) розглянуті в [16, 39, 107, 130, 175, 177 та ін.].

Найпростіший принцип побудови сигналів з АФМ виявляється в тому, що сигнальні точки розміщуються на двох концентричних окружностях. Цей шлях, однак, не завжди приводить до оптимального результату. В якості прикладу на рис.4.5 показаний побудований таким чином восьмипозиційний сигнал з АФМ - сигнальні точки позначені хрестиками. Чотири сигнали розміщені на окружності радіусом $R = \sqrt{E}$, а чотири — на окружності радіусом $r < R$ із зсувом по фазі на $\pi/4$. Дана сукупність сигналів оптимізується за критерієм максимуму мінімальної відстані між сигналами шляхом вибору відношення радіусів R і r . Дійсно, чим більше r при фіксованому R , тим більша відстань між точками внутрішньої окружності, але тим

менша відстань між точками внутрішньої і зовнішньої окружностей. Тому шуканий максимум досягається, коли ці відстані стають рівними, тобто коли трикутник abc на рис.4.5 – рівносторонній:

$$ab = bc = ca \quad (4.18)$$

Найдемо співвідношення R/r , при якому виконується умова (4.18).

Очевидні наступні рівності:

$$\frac{ad}{2} = r \sin \frac{\pi}{4}, \quad \frac{bc}{\sin \pi/4} = \frac{R}{\sin(\pi - \pi/4 - \pi/6)}. \quad (4.19)$$

Підставивши (4.19) у (4.18), одержимо:

$$R/r = 2 \cos \pi/12 \approx 1/932 \quad (4.20)$$

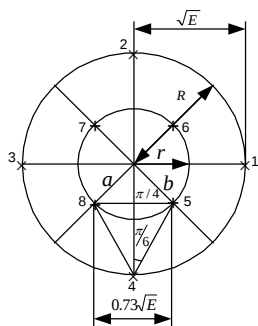


Рис.4.5. Восьми позиційний сигнал
позиційний

з АФМ із розташуванням сигнальних
точок на двох концентричних
окружностях

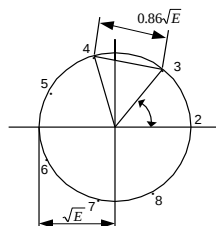


Рис.4.6. Найкращий восьми

сигнал у двовимірному просторі

При цьому оптимальному відношенні R/r мінімальна відстань між сигналами

$$d_8 = ab = \frac{\sqrt{E} \sin \pi / 4}{\cos \pi / 12} \approx 0.73\sqrt{E}. \quad (1.54)$$

Ця відстань трохи менша, ніж у системи восьмипозиційних ФМ сигналів, розташованих на одній окружності радіусом $R = \sqrt{E}$ (див. рис.4.4, д і е і табл.4.2). Таким чином, у випадку трикратної системи розміщення сигнальних векторів на двох концентричних окружностях замість однієї не дає виграшу. Оптимальною за критерієм максимуму мінімальної відстані виявляється найпростіша восьмипозиційна система з АФМ, у якій сім сигналів розміщені рівномірно на окружності радіусом $R = \sqrt{E}$, а

восьмий сигнал дорівнює нулеві [130]. Така сукупність сигнальних точок показана на рис.4.6. Мінімальна відстань у цій системі:

$$d_8 = ab = \frac{\sqrt{E} \sin 2\pi / 7}{\cos \pi / 7} \approx 0.86\sqrt{E}. \quad (4.22)$$

Недолік сигналу, показаного на рис.4.6, той же, що й у всіх сигналів з АФМ, - нерівнопотужність варіантів. Нерівнопотужність сигналів викликає певні труднощі як при їх передачі по каналі зв'язку (особливо при наявності нелінійних перетворень), так і при реалізації оптимальної обробки. Не дивлячись на відомі недоліки АФМ, цей спосіб модуляції широко використовується, так як при зростанні m рівнопотужні фазомодульовані сигнали швидко втрачають завадостійкість.

Розглянемо 16-позиційні сигнали з ФМ і АФМ.

На рис.4.7 у якості першого прикладу показані дві системи подібних сигналів.

Точками на окружності радіусом $R = \sqrt{E}$ позначені 16 варіантів сигналу з ФМ. Сигнальні точки в цьому випадку розміщені рівномірно вздовж окружності з кроком $\pi/8$. Зазвичай вони кодуються чотирирозрядними двійковими комбінаціями чотирикратного коду Грея. Однак у даному випадку код Грея є не єдиним оптимальним маніпуляційним кодом (див. гл. 2).

Відстань між найближчими сигнальними точками в 16-позиційній системі з ФМ чи ФРМ $d_{16}(\Phi M) = 0.39\sqrt{E}$ (див. табл.4.2), що відповідає програшу більше 14 дБ у порівнянні з двопозиційною системою. Зрозуміло, що в колі даного радіусу можна побудувати кращу систему сигналів з АФМ.

Система з 16 сигналів з АФМ позначена на рис.4.7 хрестиками.

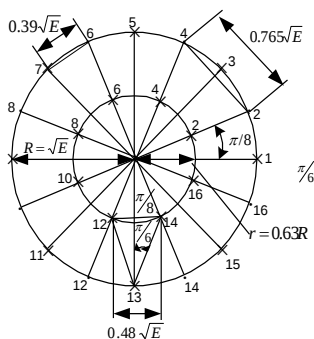


Рис.4.7. Шістнадцятипозиційний сигнал із АФМ із сигнальними точками на двох концентричних окружностях

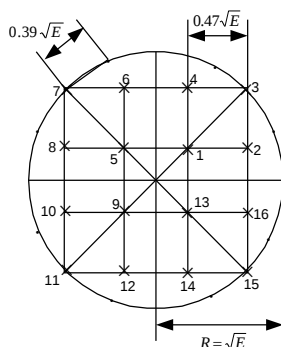


Рис.4.8. Шістнадцятипозиційні сигнали із ФМ та АФМ із сигнальними точками у вузлах квадратної решітки при обмеженій максимальній потужності

Непарні сигнали цієї системи рівномірно розміщені на окружності більшого радіусу з інтервалом $\pi/4$. Непарні сигнальні точки системи з АФМ збігаються з непарними сигнальними точками системи з ФМ. Парні сигнали системи з АФМ рівномірно розміщені на окружності меншого радіусу $r < R$ з тим же інтервалом $\pi/4$ і з загальним зсувом відносно непарних варіантів на $\pi/8$. Мінімальна відстань між точками системи максимізується вибором такого співвідношення R/r при якому відстань між сусідніми точками на окружності радіусом r дорівнює відстані між найближчими точками на двох окружностях: $ab = bc = ca$. Аналогічно (4.18) - (4.20):

$$ab = 2r \sin \pi / 8; \quad (4.23)$$

$$bc = \frac{R \sin \pi / 8}{\sin(\pi - \pi / 8 - \pi / 6)}. \quad (4.24)$$

Порівнюючи (4.23) і (4.24), одержуємо:

$$R / r = 2 \cos 5\pi / 24 \approx 1.587. \quad (4.25)$$

При даному оптимальному співвідношенні R / r мінімальна відстань між сигналами:

$$d_{16}(A\Phi M) = ab = \frac{\sqrt{E} \sin \pi / 8}{\cos 5\pi / 24} \approx 0.482\sqrt{E}, \quad (4.26)$$

яка, як видно, більша, ніж $d_{16}(\Phi M)$.

В якості другого прикладу на рис.4.8 показана 16-позиційна система сигналів з АФМ, у якій сигнальні точки, позначені хрестиками і пронумеровані, розташовані у вузлах квадратної решітки. Розмір решітки визначається тим, що діагональ великого квадрата рівна діаметру описаної довкола нього окружності $2R = 2\sqrt{E}$. На окружності для зіставлення розміщені 16 сигнальних точок системи з ФМ. У розглянутому прикладі обидві системи, що зіставляються, мають однакове обмеження на максимальну потужність сигналу, яка у системі з АФМ досягається при передачі сигналів з номерами 3, 7, 11 і 15. За таких умов мінімальна відстань між сигнальними точками:

$$d_{16}(A\Phi M) = \sqrt{2E} / 3 \approx 0.47\sqrt{E}. \quad (4.27)$$

Це більше, ніж у системі з 16-позиційною ФМ, і майже стільки ж, скільки в системі сигналів з АФМ, представленої на рис. 4.7.

Найбільший ефект дає перехід від ФМ до АФМ у системах з обмеженою середньою потужністю. У цьому випадку сигнальні точки 3, 7, 11 і 15 системи сигналів,

показаної на рис.4.8, будуть лежати на окружності радіусом, більшим, ніж $R = \sqrt{E}$

, і, отже, мінімальна відстань буде більшою, ніж $0.47\sqrt{E}$, у той час як для сигналів із ФМ усі параметри, у тому числі мінімальна відстань, зберігається попереднім, так як для цих сигналів середня і максимальна потужності рівні один одному.

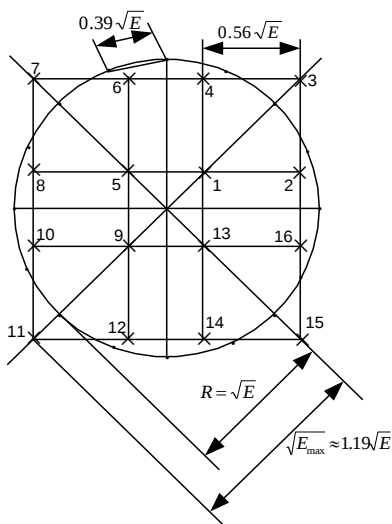


Рис.4.9. Шістнадцятипозиційні сигнали з ФМ і АРМ із сигнальними точками у вузлах квадратної решітки при обмеженій середній потужності

На рис.4.9 у якості прикладу представлені ті ж, що і на рис.4.8, 16-позиційні сигнали з ФМ і АФМ, але для випадку обмеженої середньої потужності. На окружності радіусом $R = \sqrt{E}$, де E — енергія посилки сигналу з ФМ, показані 16 сигнальних точок чотирикратного фазомодульованого сигналу. Шістнадцять

сигнальних точок системи з АФМ, відзначені хрестиками, розташовані у вузлах квадратної решітки, яка у даному випадку виходить за межі окружності радіусом $\sqrt{E}$. Для знаходження порівняльних енергетичних параметрів сигналів необхідно виходити з рівності середньої потужності або середньої енергії сигналів із ФМ і АФМ. У сигналів із ФМ середня і максимальна енергії посилок сигналу

$$E_{cp}(\Phi M) = E_{\max}(\Phi M) = E. \quad (4.28)$$

У сигналів з АФМ середня енергія виражається через енергію варіантів:

$$E_{cp}(A\Phi M) = \frac{1}{16} \sum_{i=1}^{16} E_i. \quad (4.29)$$

Максимальна енергія має місце в сигналах з номерами 3, 7, 11 і 15:

$$E_3 = E_7 = E_{11} = E_{15} = E_{\max}. \quad (4.30a)$$

Енергія інших сигналів виражається через $E_{\max}$:

$$E_1 = E_5 = E_9 = E = E_{\max} / 3; \quad (4.30b)$$

$$E_2 = E_4 = \dots = E_{16} = \sqrt{5}E_{\max} / 3. \quad (4.30\text{в})$$

Таким чином, підставивши (4.30) у (4.29), одержимо:

$$E_{cp}(A\Phi M) = (2 + \sqrt{5})E_{\max} / 6. \quad (4.31)$$

Прирівнюючи (4.28) і (4.31), маємо:

$$E_{\max} = \sqrt{6/(2 + \sqrt{5})}\sqrt{E} \approx 1.19\sqrt{E}. \quad (4.32)$$

Як видно, при обмеженій середній потужності максимальна довжина вектора сигналу розглянутої системи з АФМ приблизно на 20% більша від довжини вектора сигналу з ФМ. При цьому мінімальна відстань між векторами збільшується з $0.47\sqrt{E}$ при обмеженій максимальній потужності (див. рис.4.8) до

$$d_{16}(A\Phi M) \approx 0.56\sqrt{E} \quad (4.33)$$

при обмеженій середній потужності (див. рис.4.9).

Розрахунки, аналогічні приведеним, показують, що для випадку сигналів з АФМ, розташованих на двох концентричних окружностях (див. рис.4.7), ті ж, що й у (4.32) і (4.33) параметри при обмеженій середній потужності:

$$\sqrt{E_{\max}} \approx 1.2\sqrt{E}; \quad d_{16}(AFM) \approx 0.58\sqrt{E}. \quad (4.34)$$

Таким чином, розглянуті 16-позиційні сигнали з АФМ - один варіант на рис.4.7, а другий на рис.4.8 і 4.9 - приблизно еквівалентні по потенційної завадостійкості при обмежених як максимальній, так і середньої потужностях. Ці варіанти не є строго оптимальними, однак близькі до оптимального за завадостійкістю [130]. Вони зручні для реалізації і внаслідок цього часто використовуються на практиці.

Розглянуті багатопозиційні фазомодульовані сигнали і багатопозиційні сигнали з амплітудно-фазовою (фазорізницевою) модуляцією найбільш часто використовуються у системах цифрового зв'язку. В радіоканалах майже виняткове застосування знаходять сигнали з одно-, дво- і трикратною ФМ і ФРМ, а також засновані на них сигнально-кодові конструкції. У каналах проводного і радіорелейного зв'язку крім них використовуються чотирикратні сигнали з АФМ і сигнали більш високої кратності.

Докладні відомості про багатопозиційні сигнали з амплітудно-фазовою модуляцією, включаючи питання їхнього кодування й аналіз завадостійкості, можна знайти в [16, 130].

4.3. Фазорізницева модуляція високого порядку

Фазорізницева модуляція є найбільш важливим у практичному відношенні частковим випадком різничної модуляції. Володіючи достоїнствами класичної фазової модуляції в області досягнення високої швидкості і вірності передачі інформації, ФРМ як різнична модуляція дозволяє досягнути інваріантності системи передачі інформації до змін фази і частоти сигналу. Останньою властивістю володіє і частотно-різнична модуляція (ЧРМ), однак за інших рівноправних умов ФРМ забезпечує більш високі питому швидкість передачі інформації та завадостійкість.

При ФРМ k -го порядку інформація вкладається в послідовність k -х різниць фаз несучого коливання, які можуть приймати кінцеве число значень: $\Delta^k \varphi_1, \Delta^k \varphi_2, \dots, \Delta^k \varphi_m$. Відповідна система називається m -позиційною системою з ФРМ k -го порядку чи при $m = 2^N$ системою з N -кратною ФРМ k -го порядку. У N -кратних системах варіанти переданих k -х різниць фаз приймають значення $\Delta^k \varphi_i = \Delta^k \varphi_0 + (i-1)\pi / 2^{N-i} \quad (i = 1, 2, \dots, 2^N)$, де $\Delta^k \varphi_0$ - початкове значення k -ї різниці фаз, яке обирається виходячи зі зручності реалізації операцій модуляції та демодуляції.

Вибір порядку ФРМ або ЧРМ у тій чи іншій системі визначається характером неадитивних дій, що приводять до нестабільності частоти і фази сигналу, і апріорними відомостями про ці дії. Розглянемо це питання докладніше.

Нехай сигнал-переносник має коливання виду:

$$S(t) = a \sin \psi(t.) \quad (4.35)$$

Виділимо із аргументу даної тригонометричної функції лінійну складову $\omega_0 t$, яка не приводить до змін початкової фази в моменти, кратні тривалості послілки T , що, справедливо при $\omega_0 = 2\pi l / T$, де l - ціле число. Тоді різниця фаз $\psi(t) - \omega_0 t = \varphi(t)$ буде характеризувати зміни початкової фази сигналу в зазначені моменти, а сигнал (4.35) можна представити у вигляді:

$$S(t) = a \sin[\omega_0 t + \varphi(t)]. \quad (4.36)$$

При цифровій передачі повідомлень по каналах зв'язку фазомодульованими сигналами функція $\varphi(t)$ відображає як інформацію, що передається, так і збурюючі дії. Останні називаються завадами, а також нелінійними і параметричними перетвореннями сигналу в каналі зв'язку. Ці перетворення, як правило, можна представити на кінцевому інтервалі поліноміальною моделлю з випадковими коефіцієнтами $\varphi_0, \varphi_1, \varphi_2, \dots$, що характеризують постійні, лінійні, квадратичні і т.д. зміни початкової фази сигналу,

$$\varphi(t) = \sum_{i=0}^{k-1} \varphi_i t^i + \varphi_u(t), \quad (4.37)$$

де φ_u - інформаційна фаза сигналу. Так як відхилення частоти сигналу ω від номінального значення ω_0 дорівнює похідній від (4.37), то:

$$\omega(t) = \frac{d\varphi}{dt} = \sum_{i=0}^{k-1} i \varphi_i t^{i-1} + \omega_u(t), \quad (4.38)$$

де ω_u - інформаційне відхилення частоти сигналу. З (4.37) та (4.38):

$$\frac{d^k[\varphi(t) - \varphi_u(t)]}{dt^k} \equiv 0, \quad \frac{d^{k-1}[\omega(t) - \omega_u(t)]}{dt^{k-1}} \equiv 0, \quad (4.39)$$

звідки у відповідності з доведеною вище властивістю інваріантності дискретно різничного перетворення випливає, що k -е різниці фаз сигналу і $(k-1)$ -е різниці частоти не залежать від збурюючих дій, які описуються функцією (4.37).

Таким чином, якщо паразитна зміна фази сигналу на інтервалі тривалістю $(k+1)T$ можна представити у вигляді полінома $(k-1)$ -го порядку, то абсолютну інваріантність системи до цих змін може забезпечити ФРМ k -го порядку чи $(k-1)$ -го порядку.

Розглянемо найбільш характерні часткові випадки:

1. Невизначена, але незмінена початкова фаза сигналу:

$$\varphi(t) = \varphi_u(t) + \varphi_0; \quad \omega(t) = \omega_u(t). \quad (4.40)$$

Ця ситуація виникає внаслідок некогерентності задаючих генераторів у лінії зв'язку або нестрокої когерентності сигналу та опорного коливання в демодуляторі (когерентність з точністю до дискретного зсуву). Так як у цьому випадку

$$\frac{d[\varphi(t) - \varphi_u(t)]}{dt} \equiv 0; \quad \omega(t) - \omega_u(t) \equiv 0,$$

то

$$\Delta^1 \varphi = \Delta^1 \varphi_u = \text{in var } \varphi_0; \quad \Delta^0 \omega = \omega_u = \text{in var } \varphi_0. \quad (4.41)$$

Це означає, що абсолютну інваріантність до постійних відхилень фази сигналу забезпечує ФРМ першого порядку і ЧРМ нульового (тобто ЧМ) порядку.

2. Невизначена і лінійно змінна початкова фаза (невизначена, але постійна частота сигналу):

$$\varphi(t) = \varphi_u(t) + \varphi_0 + \varphi_1 t; \quad \omega(t) = \omega_u(t) + \varphi_1. \quad (4.42)$$

Ця ситуація виникає в наслідок нестабільності та несталонності задаючих генераторів у лінії зв'язку та ефектом Допплера при зв'язку з рухомим чи нестационарним об'єктом із рівномірною у часі зміною відстані між передавачем і приймачем. Так як у даному випадку:

$$d^2[\varphi(t) - \varphi_u(t)]/dt^2 \equiv 0; \quad d[\omega(t) - \omega_u(t)]/dt \equiv 0$$

то

$$\Delta^2 \varphi = \Delta^2 \varphi_u = \text{in var}(\varphi_0, \varphi_1); \quad \Delta^1 \omega = \Delta^1 \omega_u = \text{in var}(\varphi_0, \varphi_1). \quad (4.43)$$

Це означає, що абсолютну інваріантність до лінійних відхилень фази сигналу може забезпечити ФРМ другого порядку і ЧРМ першого порядку.

3. Невизначена і квадратично змінна фаза сигналу (невизначена і лінійно змінна частота сигналу):

$$\varphi(t) = \varphi_u(t) + \varphi_0 + \varphi_1 t + \varphi_2 t^2; \quad \omega(t) = \omega_u(t) + \omega_1 + 2\omega_2 t. \quad (4.44)$$

Ця ситуація виникає внаслідок нестабільності задаючих генераторів у лінії зв'язку та через ефект Доплера при рівноприскореній зміні відстані між передавачем і приймачем. Так як у даному випадку

$$d^3[\varphi(t) - \varphi_u(t)]/dt^3 \equiv 0; \quad d^2[\omega(t) - \omega_u(t)]/dt^2 \equiv 0$$

то

$$\Delta^3 \varphi = \Delta^3 \varphi_u = \text{invar}(\varphi_0, \varphi_1, \varphi_2); \quad \Delta^2 \omega = \Delta^2 \omega_u = \text{invar}(\varphi_0, \varphi_1, \varphi_2) \quad (4.45)$$

Це означає, що абсолютну інваріантність до квадратичних змін фази (лінійних відхиленнях частоти) сигналу можуть забезпечити ФРМ третього порядку і ЧРМ другого порядку.

Деякі додаткові зауваження слід зробити з приводу умов і засобів досягнення інваріантності в розглянутих ситуаціях.

Щодо умов відзначимо, що апроксимаційна модель змін фази і частоти, яка задається співвідношеннями (4.37) і (4.38), повинна бути справедливою для інтервалів спостереження, тривалість яких дорівнює тривалості $k+1$ посилки сигналу. Якщо, наприклад, використовується ФРМ-1, то для досягнення абсолютної інваріантності до початкової фази остання повинна бути незмінною на інтервалі двох посилок. Якщо ж використовується ФРМ-2, то для досягнення абсолютної інваріантності до частоти несучого коливання ця частота повинна бути незмінною на інтервалі трьох посилок.

Якщо зазначені умови не виконуються і випадкові зміни фази не виражаються точно через поліном

($k - 1$)-го ступеня на інтервалі $k + 1$ послілки, то застосування ФРМ- k може забезпечувати лише відносну інваріантність до відповідних паразитних змін фази і частоти сигналу, тобто послабити, а не усунути вплив збурюючих дій. Ступінь такого ослаблення потрібно розраховувати для кожної конкретної ситуації окремо.

Засобами досягнення інваріантності до змін частоти та фази сигналу є адекватні алгоритми формування (модуляції) і обробки (демодуляції) відповідних сигналів.

Реалізація модулаторів у системах із ФРМ- k не містить ніяких принципових труднощів. Із загальних алгоритмів сигналів з ФРМ неважко одержати відповідні алгоритми для ФРМ- k . Загальні схеми формування сигналів з різничною модуляцією показані на рис.4.1,а і 4.2, також містять вичерпну інформацію, необхідну для складання функціональних схем модулаторів сигналів із ФРМ- k . У якості прикладу на рис.4.4 показані схеми модулаторів сигналів із ФРМ-2 і ФРМ-3. Вони містять кодери, фазові модулатори (ФМ) і пристрої перетворення різниць фаз другого або третього порядку в значення поточної фази сигналу, які, зазвичай, називають перекодовуючими пристроями (обведені штриховими лініями). Питання побудови кодерів, що входять у ці схеми, розглядаються в наступному параграфі.

Більш складним є питання про реалізацію демодуляторів сигналів з ФРМ- k . Загальні схеми обробки сигналів з різничною модуляцією (див. рис.4.1,б і 4.3) містять вимірювач інформаційного параметра, алгоритм якого важко отримати в загальному вигляді. При ФРМ - це вимірювач-демодулятор сигналу, що будується по-різному в залежності від апріорних відомостей про прийнятий сигнал і з врахуванням порядку обчислюваних надалі різниць фази. Так, при ФРМ-1, як відзначалося вище, властивість абсолютної інваріантності, стосовно невизначеної початкової фази сигналу, досягається за допомогою будь-якого некогерентного демодулятора сигналів із ФРМ-1, у тому числі за допомогою оптимального некогерентного й автокореляційного демодуляторів. Алгоритми роботи і схеми таких демодуляторів докладно розглядаються нижче. При ФРМ-2 властивість абсолютної інваріантності до частоти несучого коливання

досягається тільки при використанні автокореляційних демодуляторів, що також докладно вивчаються нижче. При ФРМ-3 автокореляційний демодулятор, як буде наведено, забезпечує тільки відносну інваріантність до лінійних змін частоти, а для досягнення абсолютної інваріантності потрібно синтезувати спеціальний недовимірний вимірювач поточної фази сигналу. У загальному випадку вимірювач фази в системах із ФРМ високого порядку являє собою пристрій оптимального оцінювання фази сигналу при апіорних відомостях, які відповідають тій ступені невизначеності, яку дана система повинна подолати [137].

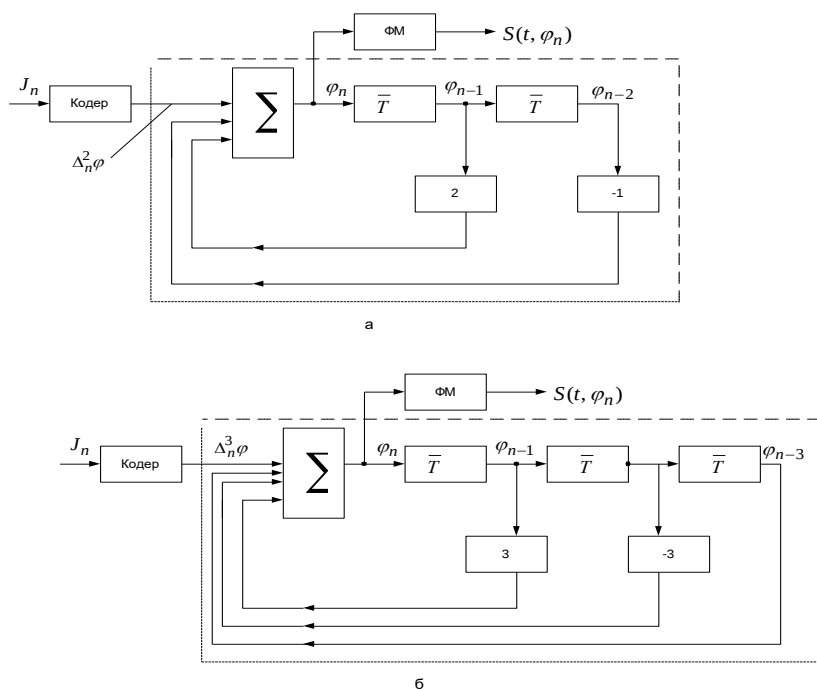


Рис.4.4. Модулятори сигналів з ФРМ другого (а) та третього (б) порядків

4.4. Система з перемінними параметрами, інваріантна до адитивної перешкоди

Розглянемо систему, інваріантну до адитивної перешкоди у виді гармонійного коливання з випадковими амплітудою, частотою і фазою, однак без обмеження на стійкість параметрів прийомного пристрою.

Досліджуємо, чи можна в системах з перемінними параметрами одержати кращі результати з погляду досягнення інваріантності, чим у системах з постійними параметрами.

Ясно, що системи з перемінними параметрами, у принципі, мають великі можливості. Наприклад, потенційна завадостійкість систем передачі дискретної інформації досягається при використанні когерентного приймача. Когерентний, же прийом може бути практично реалізований тільки в адаптивному пристрої, оскільки когерентне із сигналом опорне коливання формується з прийнятої суміші сигналу з завадою за допомогою спеціальних пристроїв автоматичного підстроювання частоти і фази.

З погляду досягнення інваріантності до зосередженого по спектру завад системи з перемінними параметрами також мають переваги в порівнянні із системами з постійними параметрами.

Як було наведено вище, у системі з постійними параметрами інваріантність до такої завади (абсолютна чи відносна) можна досягти тільки шляхом ускладнення сигналу, збільшення його бази ΔfT . Як правило, базу не можна збільшувати безмежно, тому що при заданій швидкості передачі символів, рівної $1/T$, це можна зробити тільки розширивши смугу частот каналу, що завжди важко. Наприклад, у короткохвильовому радіоканалі смуга частот, що відводиться до однієї станції, не може бути більше декількох десятків кілогерців. Якщо навіть $\Delta f = 100$ кГц, то при швидкості маніпуляцій 300 посилок/з ($T=3,33$ МС) база системи $\Delta fT \approx 330$. При такій базі максимально припустиме перевищення потужності завади над потужністю сигналу дорівнює тільки десяти ($P_n/P_c = 10$), якщо вважати за можливе зменшення вдвічі ($q = Q/Q_{\text{зкв}} = 2$) еквівалентної енергії сигналу.

Крім того, при розширенні смуги пропускання каналу збільшується імовірність влучення в неї декількох вузькосмугових завад, що викликає додаткові труднощі.

Таким чином, можливості досягнення інваріантності в рамках систем з постійними параметрами обмежені (з чого, однак, не випливає, що цими можливостями потрібно зневажати!).

Розглянемо тепер широкосмугову систему зі складеним сигналом і адаптивним прийомним пристроєм (рис.4.5). Нехай як елементи складеного сигналу використовуються гармонійні коливання з частотами $\omega_1, \omega_2, \dots, \omega_m$, а сам сигнал являє собою суму цих коливань.

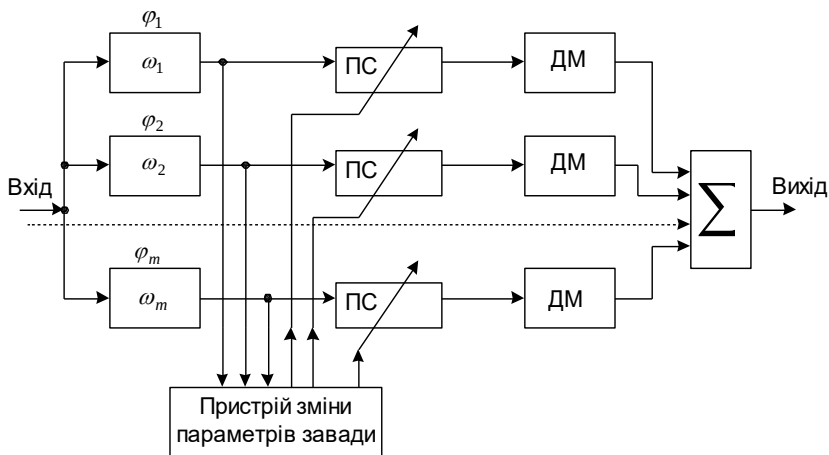


Рис.4.5. Структурна схема адаптивного приймача рівнобіжного складному сигналу

Такий сигнал називається рівнобіжним складним, тому що кожен інформаційний символ переноситься в той самий час всіма елементами сигналу. Вид модуляції елементів складного сигналу в даному прикладі не істотний: це може бути,

наприклад, частотна чи фазорізнічна модуляція. Важливо тільки, щоб спектри модульованих елементів складного сигналу не перекривалися.

У прийомному пристрої (рис.4.5) сигнал проходить через систему голосових фільтрів $\Phi_1, \Phi_2, \dots, \Phi_m$ з частотами настроювання $\omega_1, \omega_2, \dots, \omega_m$, у результаті чого окремі частотні складові (елементи складного сигналу) виявляються цілком розділеними на виходах смугових фільтрів. Кожен елемент складного сигналу проходить потім через підсилювач ПС з регульованим коефіцієнтом передачі і надходить на демодулятор. Потім модульовані елементи складеного сигналу спільно обробляються з метою ухвалення рішення про переданий інформаційний символ. Тому що вони несуть ту саму інформацію, то переданий символ можна визначати «методом голосування» по більшості результатів демодуляції елементів сигналу - у цьому випадку суматор $\sum$ виконує роль мажоритарної логічної схеми. Можна також скласти аналогові напруги на виходах демодуляторів (до ухвалення рішення про переданий символ у кожнім демодуляторі), тоді суматор являє собою пристрій додавання аналогових сигналів.

Не будемо поглиблюватися в деталі роботи вихідної частини розглянутого адаптивного приймача; істотним зараз є те, що в ньому є m галузей обробки сигналів, що несуть ту саму інформацію, і що результуючий ефект складається з приватних ефектів на виходах розділових фільтрів.

Можливість досягнення інваріантності в розглянутій системі заснована на тім, що зосереджена по спектру завад проходить тільки через один з розділових фільтрів і, отже, пошкоджує тільки одну з m галузей приймача. Якщо виключити з подальшої обробки цю пошкоджену галузь, то система буде абсолютно інваріантна до завади ξ , т. ч., для неї можна $p = \text{in var } \xi = 0$.

Для визначення пошкодженої завадою галузі приймача служить пристрій виміру параметрів завад. Алгоритми його роботи можуть бути дуже різними. Оскільки в складному сигналі є велика надмірність, то, порівнюючи результат сумарної обробки сигналу з результатами обробки в кожній з галузей, можна виявляти найменш «якісну» галузь. Пристрій виміру параметрів завади може також визначати рівень завад в галузях

шляхом порівняння рівня сигналу на виході кожного фільтра із середнім рівнем сигналу на виході всіх розділових фільтрів. При будь-якому алгоритмі роботи пристрою виміру завади воно повинно виробити команду на установку коефіцієнтів передачі відповідного підсилювача; при дуже сильній заваді в i -й галузі коефіцієнт передачі i -го підсилювача практично дорівнює нулю.

При наявності, крім зосередженої завади ξ , а і флуктуаційної в розглянутій системі з перемінними параметрами так само, як у системі з постійними параметрами, можлива тільки відносна інваріантність до завади ξ . Дійсно, при запиранні однієї з галузей приймача разом з завадою усувається і частина корисного сигналу, тому імовірність помилки при наявності зосередженої завади трохи більше, ніж без її. Правда, це збільшення імовірності похибки можна зробити дуже незначним шляхом збільшення надмірності сигналу і відповідно числа галузей прийому. Однак, як вказувалося в попередньому прикладі, на шляху збільшення надмірності (бази) складного сигналу є визначені завади.

Перевага інваріантної до зосередженої завади системи з перемінними параметрами в порівнянні із системою з постійними параметрами, інваріантної до тієї ж завади, полягає в тім, що вона може забезпечити інваріантність (абсолютну чи відносну) до завад зі значно великим діапазоном амплітуд. У системі з постійними параметрами інваріантності забезпечується стосовно завади з амплітудою, що не перевершує визначену величину, що залежить від бази (надмірності) складного сигналу. У розглянутій системі з перемінними параметрами, як неважко помітити, максимальна припустима амплітуда завади не залежить від бази сигналу і визначається винятково здатністю смугових фільтрів придушувати сигнали, що лежать по частоті поза смугою пропускання. Тому що неважко побудувати фільтри з загасанням у сотні разів поза смугою пропускання, та така система може забезпечити, придушення практично великої завади. Правда, для забезпечення малого програшу в завадостійкості при дії флуктуаційного шуму в системі з перемінними параметрами також потрібно досить велика база сигналу (щоб придушення сильної зосередженої завади не приводило до значного зменшення енергії корисного сигналу).

Таким чином, розглянута система з перемінними параметрами є, інваріантною до зосередженою по спектру заваді; у порівнянні із системами аналогічного

призначення з постійними параметрами вона інваріантна до більш широкого класу зосереджених завад. Інваріантість широкосмугової системи з рівнобіжним складним сигналом досягнуть шляхом застосування складного сигналу з великою базою (так само, як у розглянутій у параграфі 1.2 системі з послідовним зіставним сигналом) і ускладнення приймача, що повинний бути адаптивним.

4.5. Система з постійними параметрами, варіантна до неадитивної завади

У загальному випадку неадитивна завада викликає випадкові зміни параметрів сигналу. Тут розглядається сигнал з випадковою частотою.

Причини, що викликають зміни частоти сигналу, дуже різноманітні: нестабільність генераторів, що задають, у каналі зв'язку, швидке переміщення джерела електромагнітних коливань чи середовища, що відбиває ці коливання (ефект Доплера), і т.д. Бувають, крім того, випадки, коли частота сигналу строго фіксована і стабільна, але точно не відома в місці прийому. Абстрагуючи від конкретної причини, що викликає невизначеність частоти сигналу, будемо вважати, що на вхід демодулятора з постійними параметрами надходить сигнал

$$x(t) = a \sin[(\omega_0 + \xi)t + \phi], \quad (4.46)$$

де ξ - випадкова величина, рівна відхиленню частоти сигналу від середнього значення ω_0 .

Помітимо, що оскільки частота сигналу невідома, то в класі систем з постійними параметрами не можна здійснити прийом сигналу (4.46) когерентним чи оптимальним некогерентним методом і варто використовувати автокореляційний (енергетичний) прийом [3, 12].

Алгоритм автокореляційного прийому сигналу з однократною фазорізницевою модуляцією першого порядку (ФРМ -1) має вид [13,14].

$$I = \text{sign} \int_0^T x_n(t) x_{n-1}(t) dt , \quad (4.47)$$

де $I = \pm 1$ - переданий інформаційний символ, $x_n(t)$ і $x_{n-1}(t)$ - дві послідовні посилки сигналу, рівні відповідно до формули (4.46):

$$x_{n-1}(t) = a \sin[(\omega_0 + \xi)t + \varphi_{n-1}], (n-1)T \leq t \leq nT;$$

$$x_n(t) = a \sin[(\omega_0 + \xi)t + \varphi], nT \leq t \leq (n+1)T.$$

Структурна схема автокореляційного приймача сигналів із ФРМ-1 представлена на рис.4.6. Тривалість затримки τ дорівнює тривалості посилки T , але в реальних пристроях ці величини завжди трохи різняться через похибки реалізації.

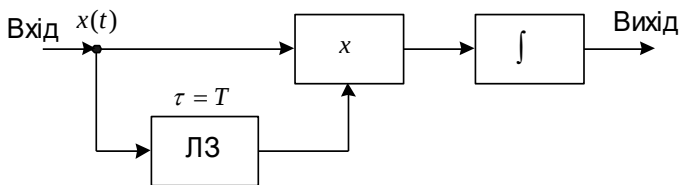


Рис. 4.6. Структурна схема автокореляційного приймача сигналів із ФРМ-1

Визначимо величину напруги на виході приймача (рис.4.6), для чого обчислимо інтеграл у правій частині виразу (4.47). Врахуємо, що $(n-1)$ -а послідовність сигналу після сполучення її за часом з n -ї (за допомогою лінії затримки) матиме вигляд

$x_{n-1}(t) = a \sin[(\omega_0 + \xi)t + (t + \tau) + \varphi_{n-1}]$. У результаті обчислення одержуємо

$$J = \int_0^T a \sin[(\omega_0 + \xi)t + \varphi_{n-1}] \sin[(\omega_0 + \xi)(t + \tau) + \varphi_{n-1}] dt =$$

$$= \frac{a^2 T}{2} \cos(\varphi_n - \varphi_{n-1} + \xi \tau) + \frac{a^2}{4(\omega_0 + \xi)} [\sin(\varphi_n - \varphi_{n-1} + \xi \tau) - \sin(\varphi_n - \varphi_{n-1} + \xi \tau + 2\xi T)]$$

(4.48)

При обчисленні виразу (4.48) прийнято до уваги, що $\omega_0 \tau \approx \omega_0 T \approx 2\pi k$.

Для спрощення подальших викладень знехтуємо другим доданком у формулі (4.48) у порівнянні з першим. Це можна зробити, коли

$$\omega_0 + \xi \geq 2\pi/T, \quad (4.49)$$

т.ч. у випадку вузькосмугового сигналу. Тоді

$$J \approx \frac{a^2}{2} \cos(\varphi_n - \varphi_{n-1} + \xi \tau). \quad (4.50)$$

Як видно з отриманого виразу, при ФРМ-1 результат «обробки сигналу на виході автокореляційного приймача залежить від зміни частоти, сигналу ξ . Якщо ж $\xi \tau > \pi/2$, то знак величини J зміниться і відповідно до алгоритму (4.47) відбудеться помилка в прийомі повідомлення. Таким чином, система з ФРМ-1

неінваріантна до перешкод, що викликають зміну частоти сигналу: $p(\text{ФРМ-1}) \neq \text{in var } \xi$ [40].

Розглянемо тепер систему з фазорізницевою модуляцією другого порядку (ФРМ-2), при якій інформація закладена в другу різницю фаз сигналу, рівну:

$$\Delta^2 \varphi = (\varphi_{n+1} - \varphi_n) - (\varphi_n - \varphi_{n-1}) = \varphi_{n+1} - 2\varphi_n + \varphi_{n-1}.$$

Схема рис.4.7 містить у собі два автокореляційних приймачі сигналів із ФРМ-1. На один з них сигнал надходить через додатковий фазообертач, що змінює фазу на $\pi/2$.

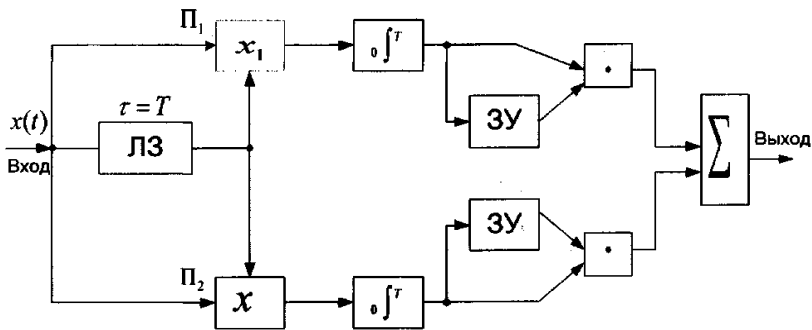


Рис.4.7. Структурна схема автокореляційного приймача сигналів із ФРМ-2

Відлік напруг на виходах інтеграторів пропорційні косинусу і синусу першої різниці фаз, тобто $\cos(\varphi_n - \varphi_{n-1})$ і $\sin(\varphi_n - \varphi_{n-1})$. Інша частина схеми, що складає з елементів пам'яті (на тривалість послілки T) постійних напруженості (ЗУ), множників постійних напруг ($\bullet$) і суматора, обчислює косинус другої різниці фаз по формулі

$$\cos \Delta^2 \varphi_n = \cos(\varphi_n - \varphi_{n-1}) \cos(\varphi_{n+1} - \varphi_n) + \sin(\varphi_n - \varphi_{n-1}) \sin(\varphi_{n+1} - \varphi_n).$$

У цілому приймач рис.4.7 реалізує наступний математичний алгоритм обробки трьох послідовних посилок сигналу $X_{n-1}(t), X_n(t), X_{n+1}(t)$:

$$I = \text{sign}(X_n X_{n-1} + Y_n Y_{n-1}) , \quad (4.51)$$

де

$$\left. \begin{aligned} X_n &= \int_0^T x_{n+1}(t) x_n(t) dt \\ X_{n-1} &= \int_0^T x_{n+1}(t) x_{n-1}(t) dt \\ Y_n &= \int_0^T x_{n+1}(t) x_n^*(t) dt \\ Y_{n-1} &= \int_0^T x_{n+1}(t) x_{n-1}^*(t) dt \end{aligned} \right\} . \quad (4.52)$$

Зірочкою тут відзначені перетворення по Гілберту відповідних посилок (цю операцію виконує фазообертач на схемі рис.4.6). Запис (4.52) так само, як і (4.47), означає, що відповідні посилки сигналу сполучені в часі на деякому умовному інтервалі $(0, T)$.

Роблячи по формулах (4.52) обчислюючи, аналогічні обчисленням (4.48)-(4.50), одержимо:

$$\left. \begin{aligned} X_n &\approx \frac{a^2 T}{2} \cos(\varphi_{n+1} - \varphi_n + \xi \tau) \\ X_{n-1} &\approx \frac{a^2 T}{2} \cos(\varphi_{n+1} - \varphi_{n-1} + \xi \tau) \\ Y_n &\approx \frac{a^2 T}{2} \cos(\varphi_{n+1} - \varphi_n + \xi \tau) \\ Y_{n-1} &= \int_0^T x_{n+1}(t) x_{n-1}(t) dt \end{aligned} \right\} . \quad (4.53)$$

Нагадаємо, що тут ξ - випадкове відхилення частоти сигналу тривалість затримки сигналу в схемі рис.4.5, приблизно тривалості посилки T .

Підставивши (4.53) у (4.51), одержимо, що величина напруги на виході автокореляційного приймача сигналів із ФРМ-2

$$J[x(t)] = X_n X_{n-1} + Y_n Y_{n-1} = \frac{a^2 T}{2} \cos(\varphi_{n-1} - 2\varphi_n + \varphi_{n-1}) = \text{in var } \xi .$$

Отже, вихідна напруга приймача сигналів із ФРМ-2 пропорційно косинусу другої різниці фаз і не залежить від частоти сигналу. Отже, $p(\text{ФРМ-2}) = \text{in var } \xi$.

Система передачі дискретної інформації з ФРМ-2, абсолютно інваріантна до частоти сигналу; розглядається нижче при вивченні питань синтезу інваріантних систем. По цьому тут обмежимося, як і в попередніх прикладах, тільки вказівкою на те, якою ціною досягнута в даному випадку інваріантність.

У попередніх прикладах інваріантність стосовно адитивної перешкоди досягалася шляхом збільшення бази сигналу й ускладнення приймального пристрою. У даному випадку сигнал із ФРМ-2 по своїх характеристиках нічим не відрізняється від сигналів з іншими видами модуляції, а ускладнення приймача в порівнянні, наприклад, із приймачем ФРМ -1 дуже незначне. Однак і в даному випадку інваріантність до перешкоди отримана не безкоштовно: платою за інваріантність є зниження завадостійкості стосовно адитивних завад. Іншими словами, у системі ФРМ-2 ймовірність похибки інваріантна до частоти, сигналу, але вона більше, ніж ймовірність похибки в системі з ФРМ-1 при незмінній частоті сигналу. Це положення ілюструється рис.4.8, на якому представлена якісна картина співвідношення між завадостійкості інваріантної (ФРМ-2) і неінваріантної (ФРМ-1) систем.

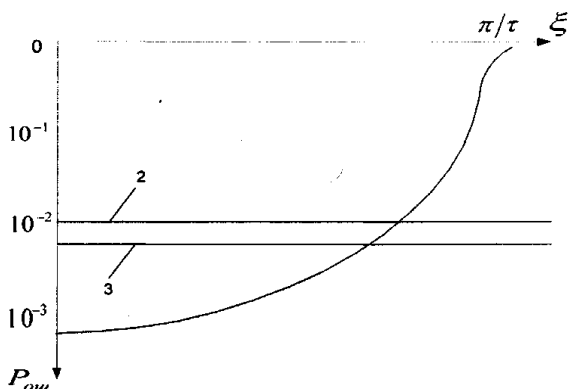


Рис.4.8. Характеристики завадостійкості інваріантної і неінваріантної систем:

1 - імовірність помилки в системі з ФРМ-1 ($p \neq \text{in var } \xi$); 2 - припустима імовірність помилки; 3 - імовірність помилки в системі з ФРМ-2 ($p = \text{in var } \xi$).

Як видно, при відсутності розстройки частоти ($\xi = 0$) ймовірність похибки в неінваріантній системі менше, ніж імовірність помилки в інваріантній. Однак якщо вимога до завадостійкості системи передачі інформації має вигляд $p < p_{\text{доп}}$ (див. пунктирну лінію на рис.4.8), то система з ФРМ-2 задовольняє цій вимозі, а система з ФРМ-1 - ні. Таке типове співвідношення між характеристиками інваріантних і неінваріантних систем зв'язку. Воно ілюструє висловлені в параграфі 1.1 розуміння про призначення і роль інваріантних систем зв'язку.

На закінчення визначимо, що в розглянутому окремому випадку - канал з невизначеною частотою сигналу, - очевидно, не тільки система з ФРМ-2 є інваріантною. Дійсно, неважко придумати багато систем, інваріантних до частоти сигналу. Наприклад, завадостійкість системи з амплітудною модуляцією (АМ) також не залежить від частоти сигналу, якщо його обробка здійснюється автокореляційним демодулятором (енергетичний приймач).

Таким чином, у ряді випадків існує трохи інваріантних (причому абсолютно, інваріантних!) до визначеної перешкоди систем і, природно, виникає, питання про те, яку з них варто вибрати. Якщо імовірність помилки в декількох систем є інваріантом деякої перешкоди, то варто вважати найкращою, оптимальною інваріантною системою ту, у якій ця імовірність менше. Таким чином, поряд з поняттями абсолютно і щодо інваріантних систем доцільно ввести поняття *оптимальної інваріантної системи*. У наступній главі ці поняття визначені більш строго.

4.6. Адаптивна система, інваріантна до адитивної перешкоди

Попередні приклади, будучи досить переконливими, володіють разом з тим одним загальним недоліком: у них розглядаються порівняно «прості» перешкоди. Це може навести на неправильну думку про те, що тільки для дуже простих перешкод можна побудувати інваріантну систему.

Тепер розглянемо адитивну перешкоду самого загального виду - флуктуаційний шум. Будемо думати, що потужність шуму змінюється необмежено, тобто від нуля до нескінченності. Математичною моделлю такої перешкоди є нестационарний гауссовський випадковий процес з необмеженою дисперсією.

У подібному нестационарному каналі зв'язку інваріантість до зухвалого цю нестационарність перешкоді можна забезпечити за допомогою адаптивної системи зв'язку. Адаптивна система зв'язку також має перемінні параметри, але на відміну від системи, розглянутої в параграфі 1.3, у ній змінюється не тільки алгоритм роботи приймача, але і, узгоджено з ним, алгоритм роботи передавача.

Розглянемо адаптивну систему передачі дискретної інформації маючої канал зворотного зв'язку. По прямому каналі зв'язку цієї системи передаються кодові комбінації довжини n , причому k : символів з кожної комбінації інформаційні, а інші $(n - k)$ - перевіірочні. Декодувальне пристрій приймача працює в режимі виявлення помилок: якщо в даній комбінації виявлений помилка, по каналі зворотного зв'язку посилається запит на повторення цієї комбінації. Помилково прийнята комбінація передається вдруге; якщо помилка при повторному прийомі не виявляється, те

передається наступна комбінація, якщо ж помилка знову виявляється, посиляється повторний запит і т.д.

Вище відтворений (дуже спрощено) алгоритм функціонування широко розповсюджених систем передачі дискретних повідомлень з так названою «вирішальною зворотним зв'язком» чи просто «систем з перепопитом» [6]. У таких системах однієї з характеристик завадостійкості є імовірність невиявленої помилки $p_{\text{но}}$, тому що споживачу невірна інформація видається тільки в тому випадку, коли помилка не виявлена.

Припустимо, що в розглянутій системі задана максимальна припустима імовірність невиявленої помилки

$$p_{\text{но}} < p_{\text{но доп}} \quad (4.54)$$

Будемо вважати систему інваріантною стосовно описаний ний вище адитивної перешкоді (флуктуаційному шуму), якщо нерівність (4.54) виконується при всіх можливих значеннях пари метрів перешкоди.

Якщо імовірність невиявленої помилки є монотонною функцією потужності перешкоди, то значення $p_{\text{но}}$ буде максимально при максимальній потужності перешкоди. Знайдемо це максимальне значення, впливаючи міркуванням роботи [6].

При потужності перешкоди, що прагне до нескінченності, імовірність помилкової реєстрації одного двійкового символу кодової комбінації прагне до 1/2. У цьому випадку на вході декодера з рівною імовірністю виникає будь-яка двійкова комбінація. Помилка не буде виявлена в тому випадку, коли з випадкових, двійкових символів складеться комбінація, що належить коду. Отже, максимум $p_{\text{но}}$ дорівнює відношенню числа кодових комбінацій даного (n, k) коду до загального числа двійкових комбінацій довжини n , тобто $\max p_{\text{но}} = 2^k / 2^n = 2^{k-n}$.

Максимум невиявленої помилки не залежить від характеристик перешкоди, а визначається винятково параметрами використовуваного коригувального коду, тобто якщо позначити перешкоду, як і колись, через ξ , те $\max p_{\text{но}} = \ln \text{var } \xi$. Якщо до

того ж $2^{k-n} \leq p_{\text{но доп}}$, то дана адаптивна система зв'язку інваріантна в зазначеному вище змісті до флуктуаційного шуму.

Підкреслимо ще раз відмінність поняття, що тут фігурує, інваріантності від аналогічного поняття в попередніх прикладах, де імовірність помилки не змінювалася при зміні параметрів перешкоди; у даному ж прикладі імовірність помилки змінюється, але не перевершує деякого максимального припустимого значення ні при яких параметрах перешкоди. Іншими словами, колись інваріантом перешкоди була імовірність помилки, а тепер інваріантом перешкоди є припустимий максимум цієї імовірності. Однак у даному прикладі можна говорити про інваріантність системи, оскільки якість зв'язку (з погляду завадостійкості) не може стати гірше заданого ні при яких обставинах.

На закінчення відзначимо, як це робилося і раніш, якою ціною отримана інваріантність у розглянутій системі. На жаль, ціна ця чимала: зниження швидкості передачі інформації.

Мало того, що в системі застосовується код з надмірністю; при збільшенні потужності перешкоди збільшується частота «перепопитів» і темп передачі інформації сповільнюється. У межі при дуже могутній перешкоді фактична швидкість передачі інформації падає до нуля: власне кажучи, у цьому режимі система «направляє зусилля» не на передачу інформації, а на «недопущення» помилкових комбінацій до споживача. Таким чином, хоча властивість інваріантності системи зі зворотним зв'язком гарантує задану вірогідність інформації, він не гарантує наперед задану вірогідність передачі інформації. Це природно, тому що гару дуже могутній перешкоді пропускну здатність каналу зв'язку дорівнює нулю, і єдине, чого можна досягти, - це не приймати помилкову інформацію.

Контрольні питання

1. Дайте означення фазорізницевої модуляції високої кратності.
2. Поясніть принцип формування і обробки фазомодульованих сигналів ФРМ-1 і ФРМ-2.
3. Поясніть, чому перехід до ФРМ-2 забезпечить інваріантність і до початкової фаз і до зсуву частоти?

4. Які алгоритми прийому можна застосовувати при ФРМ-1 і ФРМ-2?
5. Яка унікальна властивість сигналів з ФРМ-2?
6. Поясніть особливості алгоритмів прийому сигналів з ФРМ-1 і ФРМ-2?
7. Поясніть особливості автокореляційного прийому сигналів із ФРМ-2?

РОЗДІЛ 5. ПОРІВНЯННЯ МЕТОДІВ ОПТИМАЛЬНОЇ ОБРОБКИ БАГАТОПОЗИЦІЙНИХ СИГНАЛІВ

Постановка задачі

У даному розділі проведемо дослідження оптимальних методів обробки багатопозиційних сигналів, коли початкова фаза прийнятого сигналу вважається невідомою. Ці дослідження необхідні при реалізації багатоканальних модемів на місцевих (міських, сільських) мережах та мережах цільового призначення. Як відомо, при цифровизації каналів низької якості за допомогою модемів необхідно передбачити також випадки з невідомою початковою фазою надходжуваного сигналу. В загальній теорії зв'язку під оптимальним некогерентним прийомом розуміють некогерентний метод (алгоритм) обробки, що забезпечує мінімум імовірності помилки в прийомі елементу сигналу в каналі з гаусовським білим шумом при випадковій і рівномірно розподіленій початковій фазі прийнятого елемента, (про яку нічого більше не відомо). Всі інші неінформаційні параметри сигналу, крім початкової фази, у першу чергу частота та інтервал обробки, при оптимальному некогерентному прийомі повинні бути відомі точно. Що стосується амплітуди сигналу, то при використанні сигналів з рівною енергією вона може бути невідомою [55, 58]. Стосовно до ФМ сигналів оптимальний некогерентний метод прийому може бути використаний тільки для визначення переданої різниці фаз (а не абсолютної фази), котра при цьому методі обробки принципово є невідомою.

Якщо початкова фаза прийнятих елементів сигналу невідома або може бути оцінена за передісторією процесу з високою точністю, то когерентний (або квазікогерентний) демодулятор просто непрацездатний, у той час як некогерентний модулятор забезпечує досить високу завадостійкість і знаходиться поза конкуренцією.

Якщо ж початкова фаза прийнятих елементів сигналу відома або може бути оцінена за передісторією процесу з високою точністю, то придатні як когерентний, так і некогерентний демодулятори. Причому когерентний має більш високу завадостійкість. Наскільки має переваги в цих умовах когерентний демодулятор залежить від кратності модуляції, розмірності (числа елементів) оброблюваного відрізка сигналу та інших факторів. Програш у завадостійкості оптимального некогерентного прийому когерентному зростає в міру збільшення позиційності і розмірності оброблюваного сигналу. Зокрема, при чотирьохпозиційній ФРМ програш більший, ніж при двопозиційній ФРМ; при прийомі сигнально-кодової конструкції в цілому програш більший, ніж при поелементному прийомі.

Між відомою або зовсім невідомою початковою фазою, є великий спектр проміжних ситуацій. З одного боку, початкова фаза точно невідома і змінюється за часом, але, з іншого боку, може бути оцінена з деяким ступенем точності за передісторією процесу. Співвідношення між завадостійкістю квазікогерентного й оптимального некогерентного методів обробки в проміжних ситуаціях залежить від динаміки зміни фази і частоти сигналу та апіорних даних про ці зміни, від співвідношення між тривалістю елемента, тривалістю сеансу зв'язку й інтервалом кореляції фази в каналі і, нарешті, від наявних алгоритмічних і технічних засобів спостереження за частотою і фазою.

Невідома, невизначена початкова фаза сигналу має місце на початку будь-якого сеансу зв'язку, на протязі всього інтервалу прийому при короткочасних сеансах, при передачі інформації окремими імпульсами або пакетами, розділеними пасивними паузами, та у всіх інших випадках, коли до моменту обробки передісторія сигналу або відсутня, або занадто короткочасна. Аналогічні ситуації може бути у системах неперервної передачі інформації, що працюють з каналами зі змінними параметрами. Взагалі в каналах з порівняно швидкими змінами параметрів, коли на інтервалі однієї – двох посилок початкова фаза незначно змінюється, а інтервал когерентності занадто швидкий, щоб при даному відношенні сигнал-шум дати досить точну її оцінку, некогерентний прийом може дати кращі результати по завадостійкості, ніж квазікогерентний. Реалізаційні втрати квазікогерентного прийому через похибку установки початкової фази опорного коливання можуть виявитися більшими, ніж теоретичний програш оптимального некогерентного прийому когерентному. У зв'язку

з цим у названих ситуаціях віддають перевагу некогерентному, зокрема оптимальному некогерентному прийому.

Іноді до некогерентного прийому звертаються і в інших випадках. Наприклад, у каналі з постійними параметрами когерентний прийом сигналів з однократної ФРМ має більшу завадостійкість, ніж оптимальний некогерентний, однак виграш цей незначний. Адже некогерентний демодулятор не вимагає формування когерентних із прийнятими сигналами опорних коливань і, відповідно, простіший когерентного. Остання обставина в ряді випадків є вирішальною при виборі між когерентними й оптимальним некогерентним методами прийому. Крім оптимального некогерентного методу існують різні субоптимальні, близькі до нього по завадостійкості, але все-таки уступають некогерентного методу прийому, котрий реалізуються ще простіше.

Як відомо, при когерентному прийому сигналів без надлишковості перехід до обробки в цілому декількох елементів сигналу не приводить до збільшення завадостійкості порівняно з поелементним прийомом []. Виявляється, при оптимальному некогерентному прийомі ФРМ сигналів це не так, і збільшення інтервалу обробки, декількох посилок сигналу з ФРМ, приводить до зменшення ймовірності помилки, котра наближається до ймовірності помилки при строго когерентному прийомі цих же сигналів. Зрозуміло, цей висновок зроблено, коли початкова фаза сигналу залишається постійною на всьому розглянутому інтервалі обробки.

Таким чином, оптимальний некогерентний прийом сигналів із ФРМ дозволяє шляхом збільшення інтервалу обробки наблизитися до потенційної завадостійкості даної системи сигналів без оцінювання початкової фази. Тому оптимальні некогерентні демодулятори сигналів із ФРМ одержали широке застосування в багаточастотних (багатоканальних) системах з ортогональними канальними сигналами, котрі використовуються для передачі інформації в цифровій формі.

В другому розділі проведено дослідження ситуацій, при яких складаються умови, що відповідають оптимальному некогерентному та автокореляційному методам обробки прийому багатопозиційних сигналів і визначено завадостійкість демодуляторів, які реалізують ці методи прийому.

5.1. Оптимальний некогерентний прийом сигналів із ФРМ першого порядку

Згідно загальної теорії оптимального прийому розглянемо одержання алгоритмів оптимального некогерентного прийому сигналів із ФРМ [105, 142].

Нехай на вхід демодулятора надходить один з m рівномірних і рівнопотужних сигналів тривалістю τ :

$$S_{j\varphi}(t) = S_j(t) \cos \varphi + S_j^*(t) \sin \varphi, \quad j = 1, 2, \dots, m,$$

з випадкової, рівномірно розподіленої на інтервалі від 0 до 2π початковою фазою φ і білим гауссовським шумом $n(t)$, тобто

$$x(t) = S_{j\varphi}(t) + n(t).$$

При некогерентному прийомі ймовірність помилки буде мінімальною, коли демодулятор прийме рішення про передачу такого сигналу S_i , для якого при усіх $j \neq i$ виконується нерівність

$$\begin{aligned} & \left[\int_0^\tau x(t) S_i(t) dt \right]^2 + \left[\int_0^\tau x(t) S_i^*(t) dt \right]^2 > \\ & > \left[\int_0^\tau x(t) S_j(t) dt \right]^2 + \left[\int_0^\tau x(t) S_j^*(t) dt \right]^2. \end{aligned} \quad (5.1)$$

Тут S_j^* — перетворений по Гільберту сигнал S_j .

При ФРМ першого порядку величина τ дорівнює тривалості двох посилок сигналу, тобто $\tau = 2T$, тому алгоритм оптимального некогерентного демодулятора такий: на інтервалі n -ої послідовності варто прийняти рішення на користь варіанта сигналу S_i з різницею фаз $\Delta\varphi_i$

$$S_i(t) = \begin{cases} a \sin \omega t, & 0 < t \leq T, \\ a \sin(\omega t + \Delta\varphi_i), & T < t \leq 2T, \end{cases}$$

якщо при всіх $j \neq i$

$$\begin{aligned} & \left[\int_{(n-1)T}^{nT} x(t) \sin \omega t dt + \int_{nT}^{(n+1)T} x(t) \sin(\omega t + \Delta\varphi_i) dt \right]^2 + \\ & + \left[\int_{(n-1)T}^{nT} x(t) \cos \omega t dt + \int_{nT}^{(n+1)T} x(t) \cos(\omega t + \Delta\varphi_i) dt \right]^2 > \\ & > \left[\int_{(n-1)T}^{nT} x(t) \sin \omega t dt + \int_{nT}^{(n+1)T} x(t) \sin(\omega t + \Delta\varphi_j) dt \right]^2 + \\ & + \left[\int_{(n-1)T}^{nT} x(t) \cos \omega t dt + \int_{nT}^{(n+1)T} x(t) \cos(\omega t + \Delta\varphi_j) dt \right]^2. \end{aligned} \quad (5.2)$$

Проведемо в (5.2) прості перетворення та скорочення і скористаємося позначеннями:

$$\left. \begin{aligned} X_{n-1} &= \int_{(n-1)T}^{nT} x(t) \sin \omega t dt, & X_n &= \int_{nT}^{(n+1)T} x(t) \sin \omega t dt, \\ Y_{n-1} &= \int_{(n-1)T}^{nT} x(t) \cos \omega t dt, & Y_n &= \int_{nT}^{(n+1)T} x(t) \cos \omega t dt, \end{aligned} \right\} \quad (5.3)$$

одержимо наступний алгоритм:

$$\begin{aligned} & \cos \Delta \varphi_i (X_{n-1} X_n + Y_{n-1} Y_n) + \sin \Delta \varphi_i (X_{n-1} Y_n - Y_{n-1} X_n) > \\ & > \cos \Delta \varphi_j (X_{n-1} X_n + Y_{n-1} Y_n) + \sin \Delta \varphi_j (X_{n-1} Y_n - Y_{n-1} X_n), \\ & j = 1, 2, \dots, m; \quad j \neq i. \end{aligned} \quad (5.4)$$

Отримана система нерівностей (5.4) є найбільш загальний вид алгоритму оптимального некогерентного прийому сигналів із ФРМ-1, придатний для реалізації при будь-якому наборі інформаційних різниць фази: $\Delta \varphi_1, \Delta \varphi_2, \dots, \Delta \varphi_m$

Зміст одержаного алгоритму (5.4) полягає в наступному: у демодуляторі обчислюються косинуси кутів, котрі представляють собою різниці між: 1) кутом між векторами двох сусідніх посилок сигналу і 2) усіма дозволеними кутами між векторами двох сусідніх переданих посилок сигналу. При цьому перевага віддається тому із допустимих кутів (альтернатив), для котрого вказаний косинус максимальний. Особливість оптимального некогерентного прийому проявляється у тому, що кут між векторами сусідніх посилок сигналу обчислюється через їх проекції на два взаємно ортогональних коливання $\sin \omega t$ и $\cos \omega t$ (квадратурний прийом). Тоді величини X і Y , що обчислюються за (5.3), є не що інше, як проекції посилок прийнятого сигналу на координатні осі $\sin \omega t$ и $\cos \omega t$. Через ці проекції можна обчислити тригонометричні функції кута $\Delta \varphi_{n\xi}$ між векторами n і $(n-1)$ -й посилок прийнятого сигналу:

$$\left. \begin{aligned} \cos \Delta \varphi_{n\xi} &= \frac{X_{n-1}X_n + Y_{n-1}Y_n}{\sqrt{X_{n-1}^2 + Y_{n-1}^2} \sqrt{X_n^2 + Y_n^2}}, \\ \sin \Delta \varphi_{n\xi} &= \frac{X_{n-1}Y_n - Y_{n-1}X_n}{\sqrt{X_{n-1}^2 + Y_{n-1}^2} \sqrt{X_n^2 + Y_n^2}}. \end{aligned} \right\}$$

(5.5)

Таким чином, вираз в круглих дужках в (5.4) пропорційний косинусу і синусу різниці фази між n і $(n-1)$ -й послідовними прийнятим сигналами. Тому (5.4) можна представити в наступному еквівалентному виді:

$$\begin{aligned} \cos \Delta \varphi_i \cos \Delta \varphi_{n\xi} + \sin \Delta \varphi_i \sin \Delta \varphi_{n\xi} &> \cos \Delta \varphi_j \cos \Delta \varphi_{n\xi} + \\ + \sin \Delta \varphi_j \sin \Delta \varphi_{n\xi}, \quad j &= 1, 2, \dots, m; j \neq i, \end{aligned}$$

(5.6)

або ще простіше

$$\cos(\Delta \varphi_{n\xi} - \Delta \varphi_i) > \cos(\Delta \varphi_{n\xi} - \Delta \varphi_j), \quad j = 1, 2, \dots, m; j \neq i.$$

(5.7)

Користуючись (5.4), маємо більш прості алгоритми для конкретних систем сигналів. При однократній ФРМ-1 з різницями фази

$\Delta \varphi_1 = 0$ и $\Delta \varphi_2 = \pi$ одержуємо

$$X_{n-1}X_n + Y_{n-1}Y_n > -(X_{n-1}X_n + Y_{n-1}Y_n),$$

(5.8)

тобто переданий двійковий символ

$$J_n = \operatorname{sgn}(X_{n-1}X_n + Y_{n-1}Y_n).$$

(5.9)

Отриманий алгоритм (5.9) відіграє важливу роль у теорії прийому ФМ сигналів [94]. Він реалізується в демодуляторі, схема якого показана на рис. 2.1. У цій схемі, на відміну від найбільш застосованих схем квазікогерентних демодуляторів, генератор опорних коливаний (Γ) - некерований, автономний і з довільно невідомою початковою фазою. Це істотно спрощує реалізацію демодулятора. Разом з тим частота задавального генератора повинна співпадати з частотою несучого коливання ФМ сигналу; незначні відхилення частоти припустимі, проте вони приводять до зниження завадостійкості. Для роботи таких складових демодулятора, як інтегратор зі скиданням, елемент пам'яті на тривалість посылки T , реєстратор знака sgn , необхідний ще пристрій тактової синхронізації.

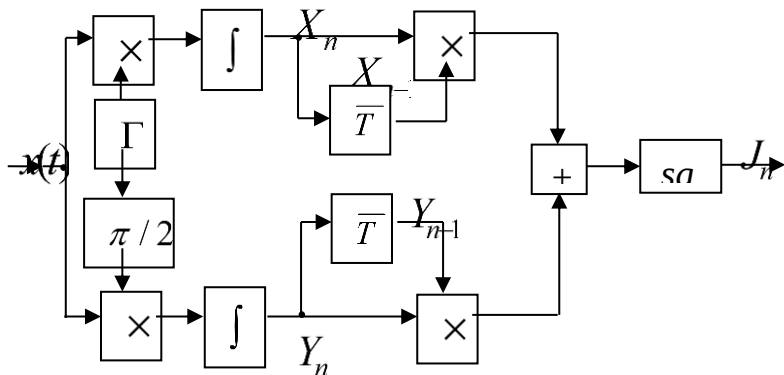


Рис. 5.1. Оптимальний некогерентний демодулятор сигналів з однократною ФМ-1

Зазначимо важливу ознаку одержаного алгоритму і схеми розглянутого демодулятора: демодулятор незначно (менш 1 дБ) уступає по завадостійкості ідеальному когерентному демодулятору, що не вимагає при цьому відомостей про початкову фазу. Але він простіший когерентних демодуляторів однократних сигналів.

В принципі можна реалізувати оптимальний некогерентний прийом без пристрою пам'яті про результати обробки на попередній посилю, як у схемі, показаної на рис. 2.1. Для цього потрібно скористатися вихідним алгоритмом (5.1), що для випадку однократної ФРМ-1 з варіантами різниці фази $\Delta\varphi_1 = 0$ и $\Delta\varphi_2 = \pi$ неважко перетворити до наступної форми:

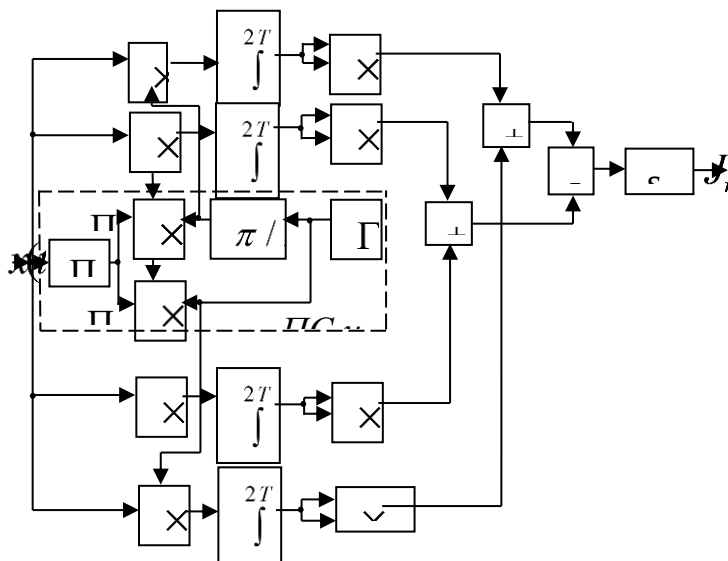
$$J = \operatorname{sgn} \left\{ \left[\int_0^{2T} x(t) \sin \omega t dt \right]^2 + \left[\int_0^{2T} x(t) \cos \omega t dt \right]^2 - \right. \\ \left. - \left[\int_0^{2T} x(t) \sin \omega t \left(\operatorname{sgn} \sin \frac{\pi}{T} t \right) dt \right]^2 - \right. \\ \left. - \left[\int_0^{2T} x(t) \cos \omega t \left(\operatorname{sgn} \sin \frac{\pi}{T} t \right) dt \right]^2 \right\}. \quad (5.10)$$

Функція $\operatorname{sgn} \sin(\pi/T)$ описує зміну фази опорного коливання на 180° на межі двох сусідніх посилок (0, T) і (T, 2T). У схемі демодулятора (рис. 5.2, а) ця функція формується за допомогою пристрою синхронізації (ПС), котрий створює періодичне прямокутне коливання $\operatorname{sgn} \sin(\pi/T)$, синфазне з границями посилок сигналу. Опорні коливання $f_1(t) = \sin \omega t$, $f_2(t) = \cos \omega t$, $f_3(t) = \sin \omega t \operatorname{sgn} \sin(\pi/T)$, $f_4(t) = \cos \omega t \operatorname{sgn} \sin(\pi/T)$, що подаються на входи чотирьох кореляторів, формуються за допомогою місцевого генератора (Γ) і фазообертача на $\pi/2$, а також за допомогою двох перемножувачів (Π_1 и Π_2) вихідного коливання ПС і прямого зсунутого на $\pi/2$ коливаний місцевого генератора. Демодулятор (рис. 5.2, а) оброблює послідовні пари посилок: 1 і 2, 3 і 4, 5 і 6 і т.д. Для обробки суміжних посилок цих пар, тобто 2 і 3, 4 і 5, 6 і 7 і т.д., необхідний ще один демодулятор. Таким чином, прийомний пристрій, що реалізує некогерентний оптимальний прийом сигналів з однократної ФРМ-1 на основі обробки двох посилок в цілому, складається, рис. 2.2, б, із двох ідентичних демодуляторів, що працюють зі зсувом на одну послітку (рис. 2.2, в), і пристрою синхронізації і формування

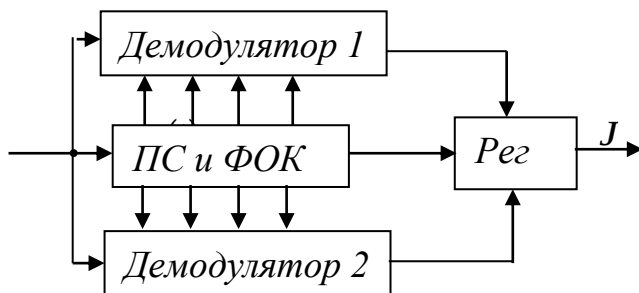
опорних коливань (ПС і ФОК). У вихідному регістрі (Рег) об'єднуються одержувані з виходів демодуляторів через послідовні двійкові символи. Варто зазначити, що за числом елементів даний демодулятор складніший рис. 5.1. При цифровій реалізації алгоритм (5.9) і схема рис. 5.1, безсумнівно, має перевагу, ніж алгоритм (5.10) і схема рис. 5.2, тому що в цьому випадку затримка вихідних величин кореляторів і їх перемноження не представляє складності.

Порівняння складності реалізації розглянутих алгоритмів зручно робити, якщо записати алгоритм (5.10) через величини (5.3) у наступному виді:

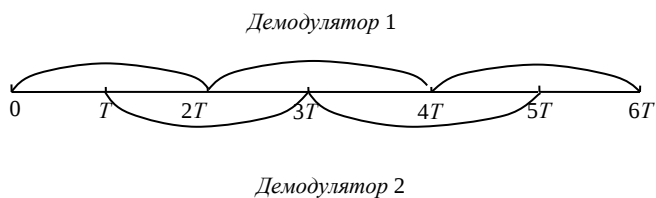
$$J_n = \text{sgn}[(X_{n-1} + X_n)^2 + (Y_{n-1} + Y_n)^2 - (X_n - X_{n-1})^2 - (Y_n - Y_{n-1})^2]. \quad (5.11)$$



а)



б)



в)

Рис. 5.2. Оптимальний некогерентний прийом сигналів з однократною ФРМ-1 при обробці двох посилок в цілому

Неважко установити, що алгоритми (5.9) і (5.11) еквівалентні.

Перейдемо тепер до багатопозиційні систем із ФРМ-1. У двократних моделах з ФРМ застосовуються в основному дві системи сигналів, показаних в табл. 2.1. В обох системах використовується двійковий маніпуляційний код Грея, котрий показаний, у чотирьохпозиційній системі є єдиним оптимальним кодом.

Таблиця 5.1

Системи сигналів у модемах з двократною ФРМ

Позначення	Різниця фаз системи, що передається		Маніпуляційний код Грея в підканалі	
	першої	другої	першому	другому
$\Delta\varphi_1$	$\pi / 4$	0	+	+
$\Delta\varphi_2$	$3\pi / 4$	$\pi / 2$	+	-
$\Delta\varphi_3$	$5\pi / 4$	π	-	-
$\Delta\varphi_4$	$7\pi / 4$	$3\pi / 2$	-	+

Знайдемо алгоритм оптимального некогерентного демодулятора спочатку для першої системи [82, 87]. Підставивши відповідні варіанти різниць фази

$\Delta\varphi_i = (2i - 1)\pi / 4, i = 1, \dots, 4$, у загальний алгоритм (5.6), одержимо

наступні чотири суми (загальний множник $1/\sqrt{2}$ опущений):

$$\left. \begin{aligned} a &= \cos \Delta\varphi_{n\xi} + \sin \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = \pi/4, \\ b &= -\cos \Delta\varphi_{n\xi} + \sin \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = 3\pi/4, \\ c &= -\cos \Delta\varphi_{n\xi} - \sin \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = 5\pi/4, \\ d &= \cos \Delta\varphi_{n\xi} - \sin \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = 7\pi/4. \end{aligned} \right\} \quad (5.12)$$

Кожний з чотирьох сум відповідає рішення про передану різницю фази, котрий повинний винести демодулятор, якщо ця сума більше трьох інших. Таким чином, в шуканому демодуляторі спочатку обчислюються a, b, c, d потім визначається найбільша з них, і їй ставиться у відповідність передана різниця

$\Delta\varphi_n$ і двійкова комбінація відповідно табл. 5.1. Проте, виразити передані двійкові символи безпосередньо через прийняті сигнали. Це зручніше зробити, співставивши (2.12) з даними табл. 5.1. Нехай, наприклад, з чотирьох сум найбільшою є a і, відповідно, по двом двійкових підканалах передано символ «+». Виходячи з того, що $a > b, \cos \Delta\varphi_{n\xi} > 0$, а з того, що

$a > d, \sin \Delta\varphi_{n\xi} > 0$. Як видно, у даному випадку знаки косинуса і синуса прийнятої різниці фаз співпадають зі знаками в двійкових підканалах. Неважко перевірити, що це має місце і в інших трьох ситуаціях.

Таким чином, двійкові символи J_{1n} и J_{2n} в двійкових підканалах можуть бути представлені у виді

$$J_{1n} = \text{sgn} \sin \Delta\varphi_{n\xi}; \quad J_{2n} = \text{sgn} \cos \Delta\varphi_{n\xi} \quad (5.13)$$

або через проєкції (2.3) прийнятого сигналу

$$\begin{aligned}
 J_{1n} &= \text{sgn}(X_{n-1}Y_n - X_nY_{n-1}); \\
 J_{2n} &= \text{sgn}(X_{n-1}X_n + Y_{n-1}Y_n).
 \end{aligned}
 \tag{5.14}$$

Схема відповідного демодулятора представлена на рис. 5.3.

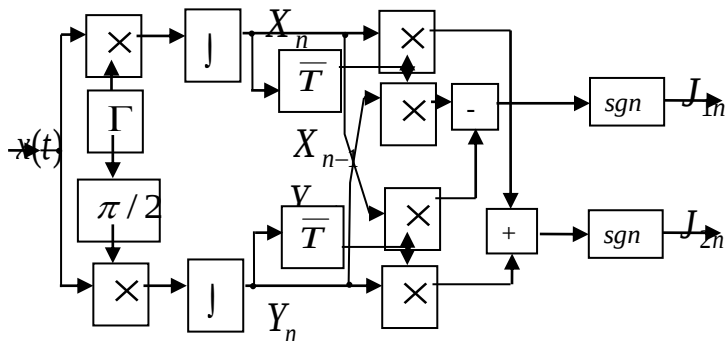


Рис. 5.3. Оптимальний некогерентний модулятор сигналів з двократною ФРМ-1 при різницях фази $\pi/4, 3\pi/4, 5\pi/4$ та $7\pi/4$

Алгоритм (5.14) зручний для реалізації у багаточастотній системі (БС) з ортогональними сигналами і двократною ФРМ, призначеної для передачі цифрової інформації.

Знайдемо аналогічний (5.14) алгоритм для другої системи сигналів (табл. 2.1.). Підставивши варіанти різниці фази $\Delta\varphi_i = i\pi/2$ ($i=0,1,2,3$) в загальний алгоритм (2.6), одержимо, що в демодуляторі обчислюються наступні величини:

$$\left. \begin{aligned}
 a &= \cos \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = 0, \\
 b &= \sin \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = \pi/2, \\
 c &= \cos \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = \pi, \\
 d &= -\sin \Delta\varphi_{n\xi}, \rightarrow \Delta\varphi_n = 3\pi/2.
 \end{aligned} \right\}
 \tag{5.15}$$

У (2.15) вказані рішення, котрі можуть видати демодулятор, якщо відповідна величина більше всіх інших. Зіставимо (5.15) з даними табл. 5.1. Нехай найбільшою є b , і, отже, по першому двійковому підканалу передається символ «+», а по другому «-». З того, що $b > c$, випливає $\sin \Delta \varphi_{n\xi} > -\cos \Delta \varphi_{n\xi}$, а з того що $b > a$, $\sin \Delta \varphi_{n\xi} > \cos \Delta \varphi_{n\xi}$, тобто

$$\sin \Delta \varphi_{n\xi} + \cos \Delta \varphi_{n\xi} > 0; \cos \Delta \varphi_{n\xi} - \sin \Delta \varphi_{n\xi} < 0. \quad (5.16)$$

Таким чином, у даному випадку символи в двійкових підканалах, співпадають зі знаками вказаних у (5.16) величин. Неважко перевірити, що й в інших трьох випадках, коли максимальною є a , c або d , передані двійкові символи визначаються знаками виразів (5.16).

Отже, шуканий алгоритм оптимального некогерентного прийому сигналів із двократною ФРМ-1 згідно (5.16) і (5.15) має вид:

$$\left. \begin{aligned} J_{1n} &= \text{sgn}[(X_{n-1}Y_n - X_nY_{n-1}) + (X_{n-1}X_n + Y_{n-1}Y_n)]; \\ J_{2n} &= \text{sgn}[(X_{n-1}X_n + Y_{n-1}Y_n) - (X_{n-1}Y_n - X_nY_{n-1})]. \end{aligned} \right\} \quad (5.17)$$

Алгоритм (5.17) відповідає схемі рис. 5.4.

Оптимальний алгоритм (5.17) можна представити в наступному компактному вигляді:

$$J_{1n} = \text{sgn} \sin(\Delta \varphi_{n\xi} + \pi / 4); \quad J_{2n} = \text{sgn} \cos(\Delta \varphi_{n\xi} + \pi / 4). \quad (5.18)$$

Як видно, переданий двійковий символ визначаються знаками синуса і косинуса прийнятої різниці фази $\Delta\varphi_{n\xi}$, зсунутої на кут $\pi/4$. Перехід від (5.18) до (5.17) можна здійснити, розкривши синус і косинус суми кутів у (5.18) і підставивши в одержані вирази значення $\cos \Delta\varphi_{n\xi}$ и $\sin \Delta\varphi_{n\xi}$ з (5.5).

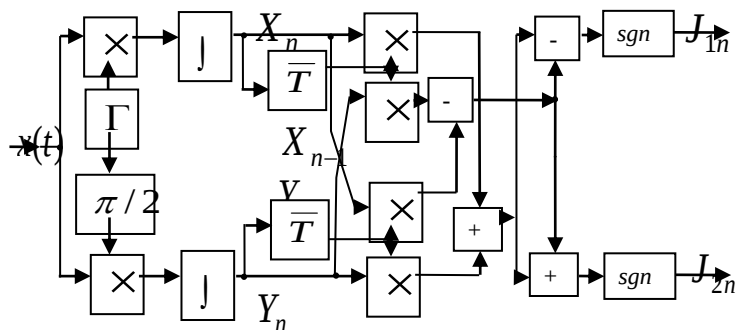


Рис. 5.4. Оптимальний некогерентний демодулятор сигналів з двократною ФРМ-1 при різницях фази $0, \pi/2, \pi$ та $3\pi/2$

З загальних алгоритмів декодування ФМ сигналів при використанні маніпуляційного коду Грея можна знайти алгоритми оптимального некогерентного прийому сигналів із ФРМ-1 довільної кратності і з довільними, але рівновіддаленими різницями фази. Нехай, наприклад, у двократній системі передаються різниці фаз:

$$\left. \begin{aligned} \Delta\varphi_1 &= \Delta\varphi_0, & \Delta\varphi_2 &= \Delta\varphi_0 + \pi/2; \\ \Delta\varphi_3 &= \Delta\varphi_0 + \pi, & \Delta\varphi_4 &= \Delta\varphi_0 + 3\pi/2, \end{aligned} \right\} \quad (5.19)$$

де $\Delta\varphi_0$ — довільна мінімальна різниця фаз. Тоді, підставивши в загальний алгоритм вираз для $\cos \Delta\varphi_{n\xi}$ і $\sin \Delta\varphi_{n\xi}$ з (5.5), одержимо наступний

алгоритм оптимального некогерентного прийому чотирьохпозиційних сигналів (5.19) при використанні коду Грея:

$$\left. \begin{aligned} J_{1n} &= \text{sgn}[(X_{n-1}X_n + Y_{n-1}Y_n)\sin(\pi/4 - \Delta\varphi_0) + \\ &+ (X_{n-1}Y_n - Y_{n-1}X_n)\cos(\pi/4 - \Delta\varphi_0)]; \\ J_{2n} &= \text{sgn}[(X_{n-1}X_n + Y_{n-1}Y_n)\cos(\pi/4 - \Delta\varphi_0) - \\ &- (X_{n-1}Y_n - X_nY_{n-1})\sin(\pi/4 - \Delta\varphi_0)]; \end{aligned} \right\} \quad (5.20)$$

Аналогічним чином можна одержати оптимальний некогерентний алгоритм обробки для трикратної системи з ФРМ при довільній мінімальній різниці фаз $\Delta\varphi_0$. Приведемо відповідні результати для двох найбільш використовуваних восьмипозиційних ФМ сигналів, представлених у табл. 5.2.

Для першої системи сигналів ($\Delta\varphi_0 = \pi/8$) одержуємо:

$$\left. \begin{aligned} J_{1n} &= \text{sgn}(X_{n-1}Y_n - X_nY_{n-1}); \\ J_{2n} &= \text{sgn}(X_{n-1}X_n + Y_{n-1}Y_n); \\ J_{3n} &= \text{sgn}[(X_{n-1}X_n + Y_{n-1}Y_n) + (X_{n-1}Y_n - X_nY_{n-1})] \times \\ &\times \text{sgn}[(X_{n-1}X_n + Y_{n-1}Y_n) - (X_{n-1}Y_n - X_nY_{n-1})]. \end{aligned} \right\} \quad (5.21)$$

Взагалі для знаходження алгоритму оптимального некогерентного прийому i — го двійкового підканалу N — кратної системи з ФРМ при використанні коду Грея потрібно представити загальний алгоритм декодування через $\cos \Delta\varphi_\xi$ і $\sin \Delta\varphi_\xi$, а їх замінити по (5.5).

Досліджено основну методику синтезу алгоритмів оптимального некогерентного прийому сигналів із ФРМ-1, котра розглядає всі можливі ситуації при обробці сигналів на двох посилках.

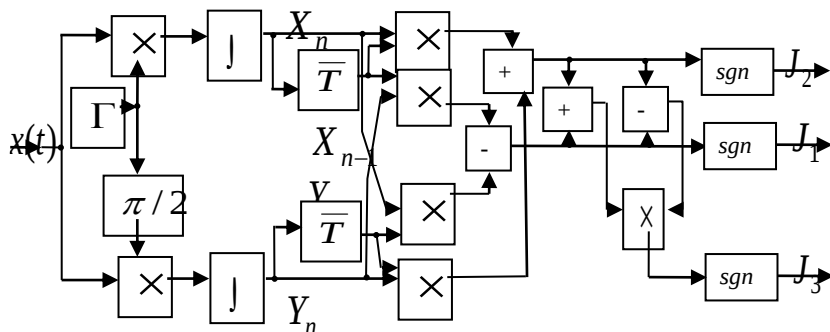


Рис. 5.5. Оптимальний некогерентний демодулятор сигналів з трьохкратною ФРМ

Одержані алгоритми можуть бути реалізовані різними способами.

5.2. Оптимальний некогерентний прийом сигналів із ФРМ другого порядку

Синтез алгоритмів оптимального некогерентного прийому сигналів із ФРМ-2 проведемо за загальним алгоритмом оптимальної некогерентної обробки (5.1), що був використаний при синтезі алгоритмів прийому сигналів із ФРМ-1. У даному випадку тривалість інтервалу обробки сигналу дорівнює трьом послідовним ($\tau = 3T$), так як передана друга різниця фаз в загальному випадку визначається не менш ніж трьома послідовними сигналами. Таким чином, необхідно записати всі можливі варіанти сигналу з ФРМ-2 на трьох послідовних і підставити ці варіанти в (5.1). При цьому перша з трьох послідовних є відліковою і має довільну, але однакову для усіх варіантів початкову фазу [50].

При однократній ФРМ-2 з різницями фази 0 або π на трьох послідовних мають місце чотири варіанти сигналу, зображені на рис. 5.6. Ці варіанти можна записати в наступному компактному виді:

$$\left. \begin{aligned} S_1(t) &= a \sin \omega t & \Delta^2 \varphi &= 0; \\ S_2(t) &= a \sin \omega t \operatorname{sgn} \sin(\pi t / T) & \Delta^2 \varphi &= 0; \\ S_3(t) &= a \sin \omega t \operatorname{sgn} \sin(\pi t / 2T + \pi / 2) & \Delta^2 \varphi &= \pi; \\ S_4(t) &= a \sin \omega t \operatorname{sgn} \sin \pi t / 2T & \Delta^2 \varphi &= \pi. \end{aligned} \right\} \quad (5.22)$$

Кожному з чотирьох сигналів у (5.22) відповідає зазначене значення різниці фази. Підставивши (5.22) у (5.1) і скористаємося позначеннями:

$$X_n = \frac{(n+1)T}{nT} \int x(t) \sin \omega t dt, \quad Y_n = \frac{(n+1)T}{nT} \int x(t) \cos \omega t dt, \quad (5.23)$$

одержимо, що в шуканому демодуляторі повинні обчислюватися наступні чотири величини:

$$\left. \begin{aligned} V_1 &= (X_{n-2} + X_{n-1} + X_n)^2 + (Y_{n-2} + Y_{n-1} + Y_n)^2; \\ V_2 &= (X_{n-2} - X_{n-1} + X_n)^2 + (Y_{n-2} - Y_{n-1} + Y_n)^2; \\ V_3 &= (X_{n-2} - X_{n-1} - X_n)^2 + (Y_{n-2} - Y_{n-1} - Y_n)^2; \\ V_4 &= (X_{n-2} + X_{n-1} - X_n)^2 + (Y_{n-2} + Y_{n-1} - Y_n)^2. \end{aligned} \right\} \quad (5.24)$$

Рішення приймається наступним чином: якщо найбільшою є V_1 або V_2 , то переданою вважається різниця фаз $\Delta^2 \varphi_n = 0$; якщо ж найбільшою виявляється V_3 або V_4 , то переданою вважається $\Delta^2 \varphi_n = \pi$.

Алгоритм (5.24) можна представити в іншому, еквівалентному виді, якщо обчислити квадрати сум у (5.24) і вилучати з отриманих виразів однакові члени –

квадрати величин X і Y . Тоді одержимо, що в демодуляторі повинні порівнюватися величини.

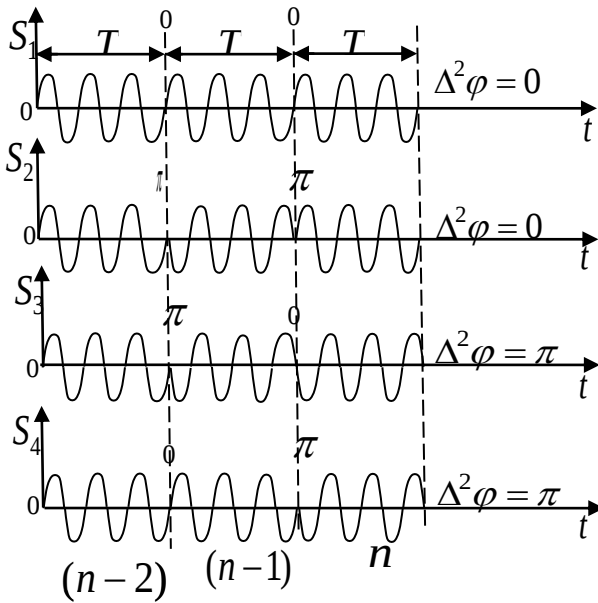


Рис. 5.6. Варіанти сигналу з однократною ФРМ-2

$$V_1 = (X_n X_{n-2} + Y_n Y_{n-2}) + (X_n X_{n-1} + Y_n Y_{n-1}) + \\ + (X_{n-1} X_{n-2} + Y_{n-1} Y_{n-2});$$

$$V_2 = (X_n X_{n-2} + Y_n Y_{n-2}) - (X_n X_{n-1} + Y_n Y_{n-1}) - \\ - (X_{n-1} X_{n-2} + Y_{n-1} Y_{n-2});$$

(5.25)

$$V_3 = -(X_n X_{n-2} + Y_n Y_{n-2}) - (X_n X_{n-1} + Y_n Y_{n-1}) + \\ + (X_{n-1} X_{n-2} + Y_{n-1} Y_{n-2});$$

$$V_4 = -(X_n X_{n-2} + Y_n Y_{n-2}) + (X_n X_{n-1} + Y_n Y_{n-1}) - \\ - (X_{n-1} X_{n-2} + Y_{n-1} Y_{n-2}).$$

Схема демодулятора, що відповідає синтезованому алгоритму, представлена на рис. 5.7. Ця схема крім звичайних для когерентних демодуляторів коректорів, генератора опорних коливань з довільною фазою і ліній затримки містить формувач величин V_1 , V_2 , V_3 і V_4 , що задовольняє алгоритму (5.23) або (5.24), і схему порівняння. Остання виносить рішення про переданий двійковий символ. Для здійснення строго оптимального некогерентного прийому в даному демодуляторі частота опорного коливання повинна співпадати з несучою частотою прийнятого сигналу, тобто цей демодулятор, також як оптимальний некогерентний демодулятор сигналів із ФРМ-1 (рис. 5.1.), неінваріантний до частоти сигналу і його завадостійкість зменшується при зсувах частоти.

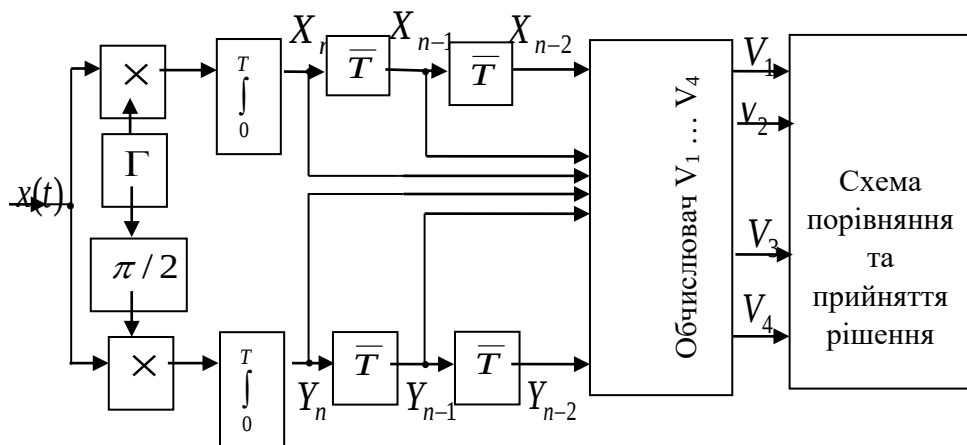


Рис. 5.7. Оптимальний некогерентний демодулятор сигналів
з однократною ФРМ-2

Чудова властивість алгоритмів (5.24) і (5.25) і демодулятора за схемою, рис. 5.7, полягає в тому, що при відсутності розстройки частоти завадостійкість демодулятора вища завадостійкості оптимального некогерентного демодулятора сигналів із ФРМ-1,

побудованого за схемою, рис. 5.1., і майже не відрізняється від завадостійкості когерентного демодулятора сигналів з однократною ФРМ-1. Це дозволяє зробити висновок про збільшення завадостійкості при подовженні інтервалу оптимальної некогерентної обробки безнадлишкових сигналів.

При двократній ФРМ-2 з варіантами різниці фази 0 , $\pi/2$, π і $3\pi/2$ на трьох послідовних, з яких перша має довільну, але однакову для усіх варіантів початкову фазу, виходить всього 16 варіантів сигналу – по чотири на кожну передану різницю. Не будемо розглядати всі варіанти, приведемо остаточний алгоритм, отриманий у результаті підстановки всіх 16 варіантів у загальний алгоритм (5.1). У шуканому демодуляторі обчислюємо величини:

$$\left. \begin{aligned} V_1 &= (X_{n-2} + X_{n-1} + X_n)^2 + (Y_{n-2} + Y_{n-1} + Y_n)^2, \\ V_2 &= (X_{n-2} + Y_{n-1} - X_n)^2 + (Y_{n-2} - X_{n-1} - Y_n)^2, \\ V_3 &= (X_{n-2} - X_{n-1} + X_n)^2 + (Y_{n-2} - Y_{n-1} + Y_n)^2, \\ V_4 &= (X_{n-2} + Y_{n-1} - X_n)^2 + (Y_{n-2} - X_{n-1} - Y_n)^2, \end{aligned} \right\} \Delta^2 \varphi_n = 0 \quad (5.26)$$

$$\left. \begin{aligned} V_5 &= (X_{n-2} + X_{n-1} + Y_n)^2 + (Y_{n-2} + Y_{n-1} - X_n)^2, \\ V_6 &= (X_{n-2} + Y_{n-1} - Y_n)^2 + (Y_{n-2} - X_{n-1} + X_n)^2, \\ V_7 &= (X_{n-2} - X_{n-1} + Y_n)^2 + (Y_{n-2} - Y_{n-1} - X_n)^2, \\ V_8 &= (X_{n-2} - Y_{n-1} - Y_n)^2 + (Y_{n-2} + X_{n-1} + X_n)^2, \end{aligned} \right\} \Delta^2 \varphi_n = \pi/2$$

$$\left. \begin{aligned} V_9 &= (X_{n-2} + X_{n-1} - X_n)^2 + (Y_{n-2} + Y_{n-1} + Y_n)^2, \\ V_{10} &= (X_{n-2} + Y_{n-1} + X_n)^2 + (Y_{n-2} - X_{n-1} - Y_n)^2, \\ V_{11} &= (X_{n-2} - X_{n-1} - X_n)^2 + (Y_{n-2} - Y_{n-1} + Y_n)^2, \\ V_{12} &= (X_{n-2} - Y_{n-1} + X_n)^2 + (Y_{n-2} + X_{n-1} - Y_n)^2, \end{aligned} \right\} \Delta^2 \varphi_n = \pi$$

$$\left. \begin{aligned} V_{13} &= (X_{n-2} + X_{n-1} - Y_n)^2 + (Y_{n-2} + Y_{n-1} + X_n)^2, \\ V_{14} &= (X_{n-2} + Y_{n-1} + Y_n)^2 + (Y_{n-2} - X_{n-1} - X_n)^2, \\ V_{15} &= (X_{n-2} - X_{n-1} - Y_n)^2 + (Y_{n-2} - Y_{n-1} + X_n)^2, \\ V_{16} &= (X_{n-2} - Y_{n-1} + Y_n)^2 + (Y_{n-2} + X_{n-1} - X_n)^2 \end{aligned} \right\} \Delta^2 \varphi_n = 3\pi / 2$$

У (5.26) проти четвірок V_j поставлені відповідні їм варіанти різниць фази.

Якщо хоча б одна величина V_i з четвірки виявляється більше всіх інших 15 величин V_j , приймається рішення на користь відповідної різниці. При побудові аналогічного (5.26) алгоритму для системи з різницями фаз $\Delta\varphi_n = \pi/4, 3\pi/4, 5\pi/4, 7\pi/4$ число порівнюваних у модуляторі величин значно збільшується.

Схема оптимального некогерентного модулятора сигналів із двократною ФРМ-2 при використанні різниць фази $0, \pi/2$ та $3\pi/2$ відрізняється від схеми, рис. 5.7, тим, що обчислювач і схема порівняння в даному випадку формують і порівнюють 16 величин (5.26) і визначають одну з чотирьох переданих різниць фази. Як видно з порівняння алгоритмів (5.24) і (5.26), перехід від однократної до двократної ФРМ-2 при оптимальній некогерентній обробці призводить до значного ускладнення демодулятора. Проте, при цифровій реалізації з використанням швидкодіючого процесора й оперативної пам'яті величини (5.26) можуть бути обчислені послідовно за час однієї n -ї посилки сигналу й ускладнення апаратури виявиться незначним.

Синтез оптимальних некогерентних алгоритмів обробки сигналів із трьохкратною ФРМ-2 і сигналів більш високої кратності розглядати не будемо. Ці алгоритми досить громіздкі і схеми демодуляторів відповідно ускладнюються [41, 42, 48].

У зв'язку з цим інтерес представляють субоптимальні алгоритми некогерентного прийому сигналів із ФРМ-2, деякі з яких мало уступають оптимальним за завадостійкістю, але в реалізації значно простіші. Один з методів побудови

субоптимальних некогерентних алгоритмів сигналів із ФРМ-2 будь-якої кратності, заснований на тому, що множина рішень на виході оптимального некогерентного демодулятора сигналів із ФРМ-1 (тобто дозволена безліч перших різниць фази створює алгебраїчне кільце порядку R , яке можна зрівняти з ізоморфним кільцем цілих чисел від 0 до $R - 1$). Отже, визначення переданого варіанту різниці фази можна здійснити шляхом обчислення різниці (по модулю R) двох R -них чисел, що послідовно з'являються на виході вказаного демодулятора ФРМ-1.

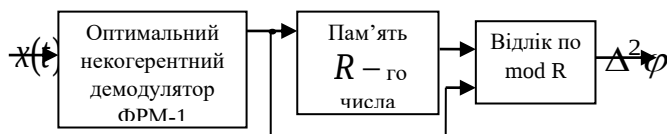


Рис. 5.8. Некогерентна обробка сигналів з ФРМ-2

Відповідна схема демодулятора представлена на рис. 5.8. Вона включає оптимальний некогерентний демодулятор сигналів із ФРМ першого порядку, пам'ять на час однієї посліжки і пристрій відліку по модулю R . Число R залежить від кратності модуляції N і мінімального значення допустимої різниці фази $\Delta^2\varphi_0$: якщо $\Delta^2\varphi_0 = 0$, то $R = 2^N$, якщо $\Delta^2\varphi_0 \neq 0$, то $R = 2\pi / \Delta^2\varphi_0$.

5.3. Автокореляційна обробка фазомодульованих сигналів

5.3.1. Застосування автокореляційної обробки сигналів

Найкращим при будь-яких умовах алгоритм, що базується на фундаментальному положенні теорії систем, відповідно деякій оптимальності завжди відносний, в тому змісті, що в залежності від обмежень, зумовлених характеристиками каналу і сигналу та особливостями режиму роботи системи зв'язку в цілому, оптимальними по деякому наперед заданому критерію виявляються різні алгоритми [47, 66].

Нехай є два варіанти переданого сигналу $S_1(t)$ і $S_2(t)$ (сигнали невідомої форми) та відомий інтервал їхнього існування $(0, \tau)$. Ці сигнали можна представити у виді розкладу по системах ортонормованих функцій $\{\varphi_i\}$ і $\{\psi_i\}$:

$$S_1(t) = \sum_{i=1}^{2B} \alpha_i \varphi_i(t), \quad S_2(t) = \sum_{i=1}^{2B} \beta_i \psi_i(t) \quad (5.27)$$

де $2B$ – база очікуваних сигналів;

$$\alpha_i = \int_0^{\tau} S_1(t) \varphi_i(t) dt; \quad \beta_i = \int_0^{\tau} S_2(t) \psi_i(t) dt \quad (5.28)$$

Сигнали $S_1(t)$ і $S_2(t)$ називаються сигналами невідомої форми, якщо відсутня будь-яка інформація про коефіцієнти φ_i та ψ_i . При цьому відомі базисні функції φ_i та ψ_i система функцій $\{\varphi_i\}$ визначає підпростір можливих форм сигналу $S_1(t)$, а система функцій $\{\psi_i\}$ – підпростір можливих форм сигналу $S_2(t)$. Можна вважати, що в розглянутому випадку відомі підпростори очікуваних сигналів, визначаються відповідним набором базисних функцій і часовим інтервалом їхнього існування.

Якщо відомі розподіли коефіцієнтів α_i і β_i , то для синтезу оптимального алгоритму прийому сигналів $S_1(t)$ і $S_2(t)$ можна скористатися простим правилом максимальної правдоподібності; якщо ж розподіли α_i і β_i невідомі, то варто скористатися правилом загальної максимальної

правдоподібності, відповідно до якого з двох можливих гіпотез ($S_1(t)$ або $S_2(t)$) вибирається та, для якої більше максимум функції правдоподібності.

Для випадку сигналів з однаковою енергією, коли відомо, що

$$\sum_{i=1}^{2B} \alpha_i^2 = \sum_{i=1}^{2B} \beta_i^2 \quad (5.29)$$

варто вважати переданим сигнал $S_1(t)$, якщо

$$\sum_{i=1}^{2B} \left[\int_0^{\tau} x(t) \varphi_i(t) dt \right]^2 > \sum_{i=1}^{2B} \left[\int_0^{\tau} x(t) \psi_i(t) dt \right]^2 \quad (5.30)$$

і сигнал $S_2(t)$ у протилежному випадку.

Неважко помітити, що в квадратних дужках у (5.30) стоять коефіцієнти розкладу прийнятого сигналу за базисними функціями φ_i і ψ_i , а суми квадратів цих коефіцієнтів (ліва і права частини (5.30)) дорівнюють енергії прийнятого сигналу в підпросторах $\{\varphi_i\}$ і $\{\psi_i\}$ відповідно. Тому алгоритм (5.30) можна представити в виді:

$$\int_0^{\tau} x_1^2(t) dt > \int_0^{\tau} x_2^2(t) dt, \quad (5.31)$$

де $x_1(t)$ і $x_2(t)$ — частини прийнятого сигналу, що знаходяться в підпросторах $\{\varphi_i\}$ і $\{\psi_i\}$.

Тому зміст отриманого алгоритму заключається в наступному; необхідно обчислити енергію сигналу, прийнятого в підпросторах першого і другого варіантів

сигналу, і вважати переданим той варіант, у підпросторі якого обчислена енергія більше. У зв'язку з цим алгоритми прийому виду (5.30) і (5.31) називають енергетичними або автокореляційними, маючи на увазі в останньому випадку, що при їхньому використанні в приймачі обчислюється згортка прийнятого сигналу з самим собою при тому або іншому часовому зсуві (в (5.31) цей зсув дорівнює нулю). Ці алгоритми справедливі при різних видах модуляції сигналу, причому для одержання відповідного часткового алгоритму варто підставити в (5.30) базисні функції, що відповідають даному виду модуляції.

Пояснимо на прикладі прийому сигналів з частотною модуляцією. Нехай частоти несучих коливань сигналів $S_1(t)$ і $S_2(t)$ невідомі, проте відомі підпростори цих сигналів: $S_1(t)$ існує в смузі частот Δf_1 , а $S_2(t)$ у смузі Δf_1 , причому Δf_1 і Δf_2 не перекриваються. Такі сигнали є частковим випадком сигналів невідомої форми. Їх прийом відповідно до приведеного алгоритму енергетичної (автокореляційної) обробки проводиться в такий спосіб. На інтервалі посліжки T обчислюється енергія прийнятого сигналу в смугах частот Δf_1 і Δf_2 . Переданим вважається сигнал S_1 , якщо енергія виявилася більшою в смузі Δf_1 і сигнал S_2 у протилежному випадку. При цьому інтеграл у лівій частині (2.31) є енергія сигналу, прийнятого в смузі частот Δf_1 , а інтеграл у правій частині (2.31) - енергія сигналу, прийнятого в смузі частот Δf_2 . Відповідно базисні функції φ_i в (2.30) є гармоніки частоти $2\pi/T$, що потрапляють в смугу $2\pi\Delta f_1$, а функції ψ_i - гармоніки, що попадають в смугу $2\pi\Delta f_2$.

Зазначимо, що розглянуті алгоритми працюють не тільки при невідомій, але і при змінній формі сигналів проте, строго кажучи, допустимі лише такі зміни, при яких кожний з варіантів переданого сигналу не виходить за рамки певного для нього підпростору. При ФРМ-1 це означає, що сигнали з

точністю до знаку повинні повторюватися на інтервалі двох посилок. Звідси, зокрема, випливають досить жорсткі вимоги до стабільності частоти несучого коливання при автокореляційному прийомі сигналів із ФРМ-1.)

Автокореляційні демодулятори сигналів із ФРМ-1 привертають увагу надзвичайною простотою, тому що для їх реалізації не потрібно ні пристроїв виділення когерентних опорних коливань, як у когерентних демодуляторах, ні кореляторів з квадратурними опорними коливаннями, як в оптимальних некогерентних демодуляторах [1, 52]. Але потрібно враховувати, що потенційно автокореляційні демодулятори уступають по завадостійкості когерентним і оптимальним некогерентним. На відміну від них завадостійкість автокореляційних (енергетичних) демодуляторів залежать не тільки від відношення h^2 енергії сигналу до спектральної щільності потужності шуму, але й до бази прийнятої суміші сигналу з шумом, що залежить, у свою чергу, від ширини смуги пропускання F вхідного фільтра приймача. Чим більше база $2FT$, тим нижче завадостійкість при тій же h^2 ; чим складніше сигнал, тим більше його розмірність, тим більше за інших рівних умов імовірність помилки - такі особливості автокореляційного прийому. Разом з тим при прийомі вузькосмугових сигналів з базою $B \approx 2$, автокореляційні демодулятори уступають оптимальним некогерентним незначно. При оцінці порівняльної завадостійкості когерентних і максимальних некогерентних демодуляторів, з одного боку, і автокореляційних, з іншого - не слід забувати, що перші мають перевагу лише при виконанні умов їх оптимальності і працездатності. Якщо ці умови не виконуються і має місце прийом сигналу апіорно невідомої форми, то автокореляційний демодулятор може забезпечити меншу ймовірність помилки. Тому можна зробити висновок, що строгий мінімум імовірності помилки автокореляційний прийом (як метод, що випливає з правила узагальненої максимальної правдоподібності) забезпечує при рівномірно розподілених невідомих параметрах (коефіцієнтах розкладу сигналу невідомої форми).

Алгоритми автокореляційної обробки сигналів із ФРМ-1 безпосередньо випливають із правила загальної максимальної правдоподібності (ЗМП) і в цьому змісті є оптимальними для сигналів невідомої форми. Аналогічний оптимальний (у змісті ЗМП) алгоритм обробки сигналів із ФРМ-2 приводить до схеми енергетичного

демодулятора. Як і у випадку оптимального некогерентного прийому, енергетичний демодулятор сигналів із ФРМ-2 перевершує за завадостійкістю еквівалентний йому автокореляційний демодулятор сигналів із ФРМ-1. Разом з тим завадостійкість демодуляторів цих типів залежить від стабільності частоти несучого коливання. Уникнути цю залежність вдається в автокореляційному демодуляторі ФРМ-2.

Якщо при ФРМ-1 автокореляційний прийом привабливий, головним чином, своєю простотою, то при ФРМ-2 автокореляційний прийом, що є в цьому випадку однією із субмаксимальних модифікацій алгоритму прийому відповідних сигналів невідомої форми, має унікальну властивість інваріантності до частоти несучого коливання, властивістю, що не має місця ні при когерентному, ні при максимальному некогерентному методах обробки. Внаслідок цього автокореляційні алгоритми прийому сигналів із ФРМ-2 мають принципове значення в теорії і техніці зв'язку. Їх доцільно застосовувати в каналах з невизначеною частотою сигналу.

Розглянемо ситуації при яких складаються умови, що відповідають каналу з невизначеною частотою, внаслідок нестабільності частоти задавальних генераторів, у всіх ланках системи зв'язку невизначена частота несучого коливання має місце на початку будь-якого сеансу зв'язку, протягом всього інтервалу прийому при короткочасних сеансах, при передачі інформації в імпульсному або пакетному режимі, особливо якщо пакети формуються різними передавачами або в різних лініях зв'язку, а також у всіх інших випадках, коли до моменту початку обробки передісторія сигналу або відсутня, або дуже короткочасна. Невизначеність частоти сигналу, крім того, є наслідком ефекту Доплера, що виникає при зв'язку з швидко рухомими об'єктами або ретрансляції сигналів через рухомий ретранслятор. При різних випадках руху об'єкта зв'язку можуть виникати непередбачені зміни частоти несучого коливання, що важко компенсувати за короткий час за допомогою пристроїв частотного автопідстроювання. В ряді випадків частотне автопідстроювання не реальне і у системах безперервної передачі інформації, якщо останні працюють по каналах зі змінними параметрами.

Таким чином, при передачі цифрової інформації різними каналами зв'язку виникає великий ряд ситуацій, при яких приймач повинний обробляти сигнал з невідомою або неточно відомою частотою несучого коливання - канали з невизначеною частотою сигналу. Таким каналам і ситуаціям адекватний

автокореляційний метод прийому при використанні сигналів із ФРМ другого порядку.

Автокореляційні модеми з ФРМ-2 володіють властивістю відносної або абсолютної інваріантності до зміни частоти несучого коливання, використовуються в системах передачі цифрової інформації при зв'язку з швидко рухомими об'єктами, в системах супутникового зв'язку, а також в системах волоконно-оптичного зв'язку.

5.3.2. Оптимальні автокореляційні (енергетичні) демодулятори сигналів із ФРМ першого порядку

Скористаємося алгоритмом прийому сигналів невідомої форми (5.31) для синтезу відповідного алгоритму прийому сигналів з однократною ФРМ-1. У цьому випадку варіанти сигналу невідомої форми на інтервалі двох посилок можуть бути представлені:

$$\left. \begin{aligned} S_1(t) &= \begin{cases} f(t), & 0 < t < T, \\ f(t-T), & T < t < 2T; \end{cases} \\ S_2(t) &= \begin{cases} f(t), & 0 < t < T, \\ -f(t-T), & T < t < 2T, \end{cases} \end{aligned} \right\} \quad (5.32)$$

де $f(t)$ – невідома функція. Для представлення сигналів S_1 і S_2 можна вибрати базисні функції наступного виду:

$$\text{для } S_1(t)$$

$$\begin{aligned}\varphi_i(t) &= (1/T) \sin i\omega_0 t, & 0 < t < T, \\ \varphi_i^*(t) &= (1/T) \cos i\omega_0 t, & T < t < 2T;\end{aligned}$$

для $S_2(t)$

$$\begin{aligned}\psi_i(t) &= \begin{cases} (1/T) \sin i\omega_0 t, & 0 < t < T, \\ (-1/T) \sin i\omega_0 t, & T < t < 2T; \end{cases} \\ \psi_i^*(t) &= \begin{cases} (1/T) \cos i\omega_0 t, & 0 < t < T, \\ (-1/T) \cos i\omega_0 t, & T < t < 2T; \end{cases}\end{aligned}$$

(5.33)

де $\omega_0 = 2\pi/T$. Функції (5.33) при $i = 1$ представлені на рис. 5.9.

Базисні функції (5.33) визначають два непересічні підпростори, що відповідають двом варіантам переданого сигналу, один із яких має стрибок фаз усіх частотних складових на 180° , а другий не має стрибка фаз.

Поряд із зазначеною розбивкою переданих при ФРМ сигналів на два непересічні підпростори можна запропонувати й іншу розбивку, при якій ці підпростори мають різні частотні складові. Дійсно, оскільки сигнал $S_1(t)$ на інтервалі $(T, 2T)$ такий же, як на інтервалі $(0, T)$, то в розкладанні сигналу по гармоніках частоти π/T відповідають всі непарні гармоніки, а в аналогічному розкладі сигналу $S_2(t)$, що на інтервалі $(T, 2T)$ повторює сигнал на інтервалі $(0, T)$ зі зворотним знаком, відповідають всі парні гармоніки.

Таким чином, при ФРМ базис підпростору сигналу $S_1(t)$ складають парні гармоніки частоти π/T , а базис підпростору сигналу $S_2(t)$ - непарні:

$$\left. \begin{aligned} S_1(t) &= \sum_{i=1}^B a_{2i} \sin 2i \frac{\pi}{T} t + b_{2i} \cos 2i \frac{\pi}{T} t; \\ S_2(t) &= \sum_{i=1}^B a_{2i-1} \sin(2i-1) \frac{\pi}{T} t + b_{2i-1} \cos(2i-1) \frac{\pi}{T} t. \end{aligned} \right\} \quad (5.34)$$

Схематичне зображення відповідних (5.34) підпросторів сигналів S_1 і S_2

представлений на рис. 5.10, де $\omega_0 = \pi / T$.

Знайдемо оптимальний алгоритм прийому сигналів невідомої форми з ФРМ. Для цього перепишемо загальний алгоритм (5.30) у наступному виді:

$$\begin{aligned} & \sum_{i=1}^B \left\{ \left[\int_0^{2T} x(t) \varphi_i(t) dt \right]^2 + \left[\int_0^{2T} x(t) \varphi_i^*(t) dt \right]^2 \right\} > \\ & > \sum_{i=1}^B \left\{ \left[\int_0^{2T} x(t) \psi_i(t) dt \right]^2 + \left[\int_0^{2T} x(t) \psi_i^*(t) dt \right]^2 \right\} \end{aligned} \quad (5.35)$$

У порівнянні з (5.30) у (5.35) прийнятий інтервал інтегрування $\tau = 2T$, так як при ФРМ кожен інформаційний символ визначається двома послідовними сигналами. Підставивши (5.33) у (5.35), отримаємо

$$\begin{aligned} & \sum_{i=1}^B \left\{ \left[\int_0^T x(t) \sin i \omega_0 t dt + \int_T^{2T} x(t) \sin i \omega_0 t dt \right]^2 + \right. \\ & \left. + \left[\int_0^T x(t) \cos i \omega_0 t dt + \int_T^{2T} x(t) \cos i \omega_0 t dt \right]^2 \right\} > \\ & > \sum_{i=1}^B \left\{ \left[\int_0^T x(t) \sin i \omega_0 t dt - \int_T^{2T} x(t) \sin i \omega_0 t dt \right]^2 + \right. \\ & \left. + \left[\int_0^T x(t) \cos i \omega_0 t dt - \int_T^{2T} x(t) \cos i \omega_0 t dt \right]^2 \right\} \end{aligned} \quad (5.36)$$

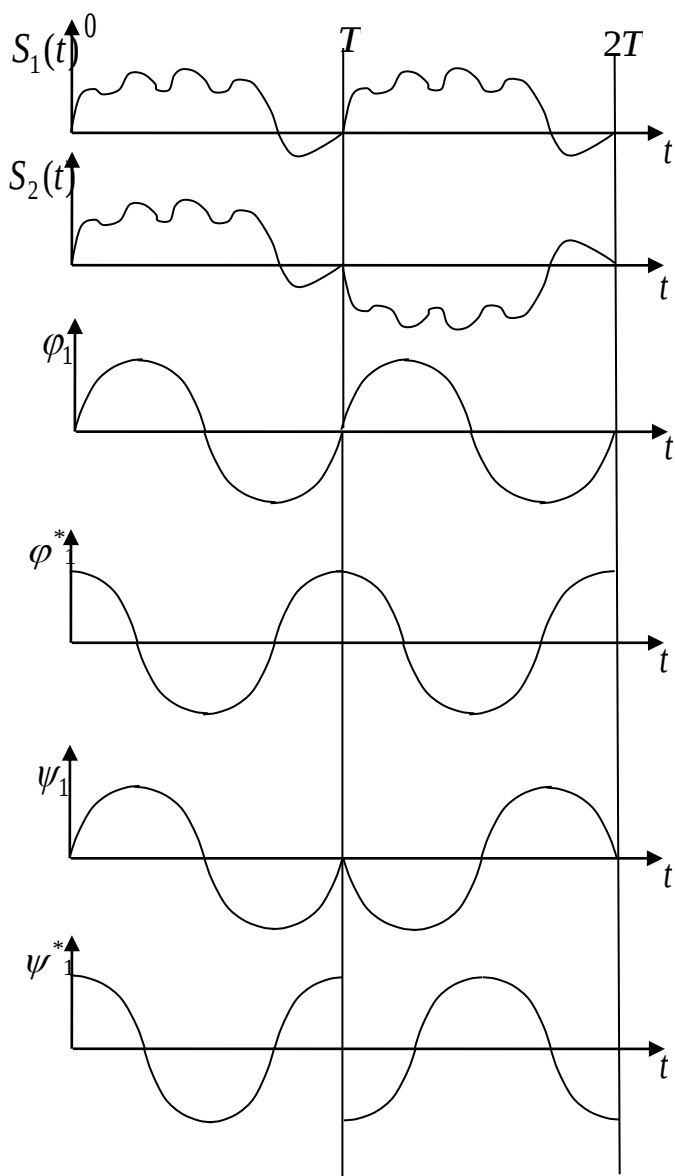


Рис. 5.9. Варіанти сигналу невідомої форми з ФРМ
та їх базисні функції на двох посилках

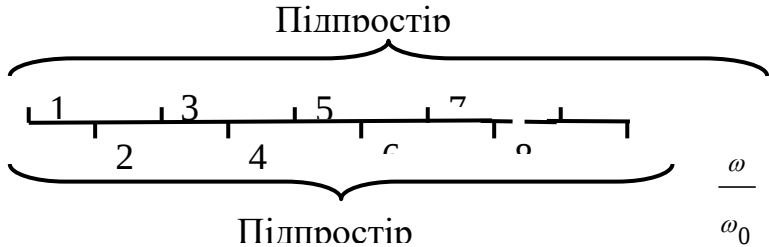


Рис. 5.10. Підпростори варіантів сигналу з ФРМ невідомої форми

Звівши до квадрату і скоротивши однакові члени в лівій і правій частинах (5.36), одержимо:

$$J = \operatorname{sgn} \sum_{i=1}^B \left[\int_0^T x(t) \sin i \omega_0 t dt \int_T^{2T} x(t) \sin i \omega_0 t dt + \right. \\ \left. + \int_0^T x(t) \cos i \omega_0 t dt \int_T^{2T} x(t) \cos i \omega_0 t dt \right]. \quad (5.37)$$

Інтеграли в (5.37) в межах $0, T$ пропорційні коефіцієнтам розкладу прийнятого сигналу на інтервалі от 0 до T , а інтеграли в межах $T, 2T$ - коефіцієнтам розкладу прийнятого сигналу на інтервалі от T до $2T$. Отже, сума добутоків цих коефіцієнтів пропорційна скалярному добуткові сигналів на двох сусідніх посилках, тобто (5.37) можна записати у виді:

$$J = \operatorname{sgn} \int_0^T x(t)x(t-T)dt \quad (5.38)$$

що співпадає з алгоритмом автокореляційного прийому сигналів з однократною ФРМ.

Таким чином, використання оптимального алгоритму прийому сигналів невідомої форми, що реалізує правило загальної максимальної правдоподібності, у випадку ФРМ-1 приводить до автокореляційної схеми обробки сигналу, представленої на рис. 5.11.

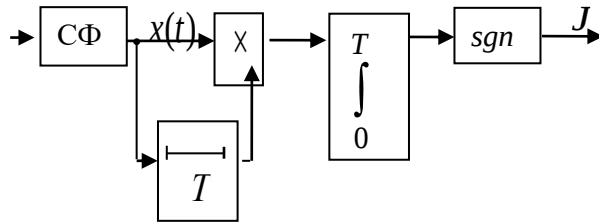


Рис. 5.11. Оптимальний автокореляційний демодулятор сигналів з однократною ФРМ-1

Вхідний сигнал надходить на смуговий фільтр (СФ), що відіграє в автокореляційному демодуляторі двояку роль.

По-перше, він виконує звичайну функцію селекції корисного сигналу по частоті. Така селекція необхідна в тих випадках, коли на вхід демодулятора надходить груповий сигнал системи передачі з ЧРК та в інших випадках. У цій функції аналогічний вхідний смуговий фільтр застосовується і на вході когерентних і оптимальних некогерентних демодуляторів. По-друге, вхідний смуговий фільтр служить для обмеження спектра (а, отже, і потужності) флуктуаційного шуму, що надходить на вхід демодулятора, оскільки при автокореляційному методі прийому, на відміну від кореляційних методів - когерентного й оптимального некогерентного, завадостійкість залежить від ширини смуги частот (потужності) шуму, а не тільки від його спектральної щільності [59, 62].

На перший погляд, може здаватися, що в алгоритмі автокореляційної обробки (5.38) не передбачається смугова фільтрація сигналу. Однак, якщо звернутися до алгоритму (5.35), з якого отриманий алгоритм (5.38), то видно, що сигнал $\chi(t)$ представлений у виді розкладання по сукупності B гармонійних складових і, отже, обмежений по частоті смугою $F=B/T$.

Принципово важливим елементом автокореляційного демодулятора сигналів із ФРМ-1 є елемент пам'яті або лінія затримки сигналу на тривалість посилки T . До тривалості затримки тут пред'являються досить жорсткі вимоги. Якщо в системах з автокореляційним прийомом сигналів ФРМ першого порядку співвідношення між частотою несучого коливання і тривалістю посилки не може бути довільним: в посилці тривалістю T повинно укладатися ціле число періодів несучого коливання частоти ω , тобто $\omega T = 2\pi k$, де k - ціле, тільки в цьому випадку форма сигналу повторюється на двох посилках, що є необхідною умовою оптимальності алгоритму прийому сигналів з ФРМ-1 невідомої форми. Відхилення від співвідношення $\omega T = 2\pi k$ викликані нестабільністю частоти ω , або ж затримки T , приводять до зниження завадостійкості, а потім і до втрати працездатності демодулятора. Допустимі відхилення частоти $\Delta\omega$ і затримки ΔT можна визначити із співвідношення

$$\Delta\varphi_{\text{доп}} = \Delta\omega T + \omega\Delta T + \Delta\omega\Delta T \quad (5.39)$$

де $\Delta\varphi_{\text{доп}}$ — допустимий паразитний фазовий зсув між коливаннями сусідніх елементів сигналу з ФРМ-1. Для збереження досить високої завадостійкості значення $\Delta\varphi_{\text{доп}}$ не повинно перевершувати частки (5..10%) мінімального дозволеного стрибка фази, що складає π при однократній ФРМ, $\pi/2$ — при двократній, $\pi/4$ — при трикратній і т.д. Варто звернути увагу на те, що, як видно з (2.39), відхилення частоти і затримки можуть компенсувати один одного, якщо вони мають різні знаки. На цьому, зокрема, базуються методи

адаптивної корекції нестабільності затримки. Звичайно допустимі відхилення частоти і затримки визначають при фіксованому і рівному номінальному значенні одного з цих параметрів, наприклад визначають $\Delta\omega_{\text{доп}}$ при $\Delta T = 0$ або $\Delta O_{\text{доп}}$ при $\Delta\omega = 0$. При точній частоті несучого коливання ($\Delta\omega = 0$) припустима нестабільність тривалості затримки $\Delta T_{\text{доп}}$ випливає з простого співвідношення

$$\Delta T_{\text{доп}} = \Delta\varphi_{\text{доп}} / \omega. \quad (5.40)$$

Звідси видно, що краще зменшувати частоту ω несучої, щоб послабить вимоги до стабільності лінії затримки. У зв'язку з цим, в автокореляційних демодуляторах іноді звертаються до переносу спектра прийнятого сигналу на більш низьку проміжну частоту. Граничним випадком такого перетворення є «перенесення на нульову частоту», тобто виділення квадратурних огинаючих сигналів, за допомогою згортки з опорним коливанням, що співпадає за частотою з несучим коливанням, котре приводить до схеми, еквівалентної оптимальному некогерентному демодулятору.

Поряд з алгоритмом (5.38) і схемою, приведеної на рис. 5.11, існують і інші еквівалентні їм алгоритми та схеми автокореляційної обробки сигналів із ФРМ-1, котрі не потребують безпосереднього обчислення коефіцієнта кореляції прийнятого сигналу.

При цифровій реалізації зручно користуватися алгоритмом (5.38) і схемою рис. 5.11, причому, питання реалізації вирішується в значній мірі в залежності від наявної елементної бази для конкретного діапазону частот і швидкостей передачі.

Дослідження алгоритмів роботи демодуляторів сигналів із ФРМ-2 проводяться аналогічно демодуляторів сигналів із ФРМ-1 за загальним оптимальним алгоритмом прийому сигналів невідомої форми. Особливість синтезу оптимального автокореляційного алгоритму прийому сигналів із ФРМ-2 у порівнянні з ФРМ-1 полягає в тому, що варіанти переданого сигналу при ФРМ-2 не ортогональні. Внаслідок цього підпростори варіантів сигналу при ФРМ-2 пересікаються, і енергія одного з варіантів сигналу, що попадає в підпростір

іншого, не дорівнює нулеві. Неважко, переконатися, що пересікання підпросторів варіантів сигналу при однократній ФРМ-2 симетричні, тобто частка енергії першого варіанта, що попадає в підпростір іншого, дорівнює частці енергії іншого, що попадає в підпростір першого. У даному випадку, як і при ортогональних сигналах, загальний алгоритм (5.30) може бути використаний для синтезу алгоритму прийому, що відповідає правилу узагальненої максимальної правдоподібності. На рис. 5.12 представлені чотири можливих варіанти довільного двопозиційного сигналу на трьох посилах при фіксованому сигналі на першій посиці: S_1, S_2, S_3 та S_4 . На кожній з посилок тривалістю T сигнал може приймати одне з двох значень, що відрізняються знаком. Якщо передаються варіанти S_1 або S_2 , це відповідає другій різниці фаз $\Delta^2 \varphi = 0$, якщо ж S_3 або S_4 , то $\Delta^2 \varphi = \pi$.

Знайдемо згортки варіантів сигналу

$$(S_i, S_j) = \int_0^{3T} S_i(t) S_j(t) dt = \begin{cases} 3E & \text{при } i = j \\ \pm E & \text{при } i \neq j \end{cases}, \quad (5.44)$$

де E - енергія сигналу на одній посиці; $i, j = 1, 2, 3, 4$. Сигнали $S_1 \dots S_4$ неортогональні, проте їх взаємні скалярні добутки рівні по модулі. Для того щоб скористатися загальним алгоритмом прийому сигналів невідомої форми, варто підставити в нього базисні функції варіантів сигналу:

$$\left. \begin{aligned}
&\text{для } S_1(t) \\
&\varphi_{1i} = \sin i \omega_0 t, \quad 0 < t < T; \\
&\text{для } S_2(t) \\
&\varphi_{2i} = \sin i \omega_0 t v, \quad \begin{cases} 0 < t < T, \\ 2T < t < 3T, \end{cases} \\
&\varphi_{2i} = -\sin i \omega_0 t, \quad T < t < 2T; \\
&\text{для } S_3(t) \\
&\varphi_{3i} = \sin i \omega_0 t v, \quad 0 < t < 2T, \\
&\varphi_{3i} = -\sin i \omega_0 t, \quad 2T < t < 3T; \\
&\text{для } S_4(t) \\
&\varphi_{4i} = \sin i \omega_0 t v, \quad 0 < t < 2T, \\
&\varphi_{43i} = -\sin i \omega_0 t, \quad T < t < 3T; \end{aligned} \right\}
\end{aligned}
\tag{5.45}$$

де $\omega_0 = 2\pi/T$. На рис. 5.12 як приклад показані базисні функції

$\varphi_{11}, \varphi_{21}, \varphi_{31}$ и φ_{41} .

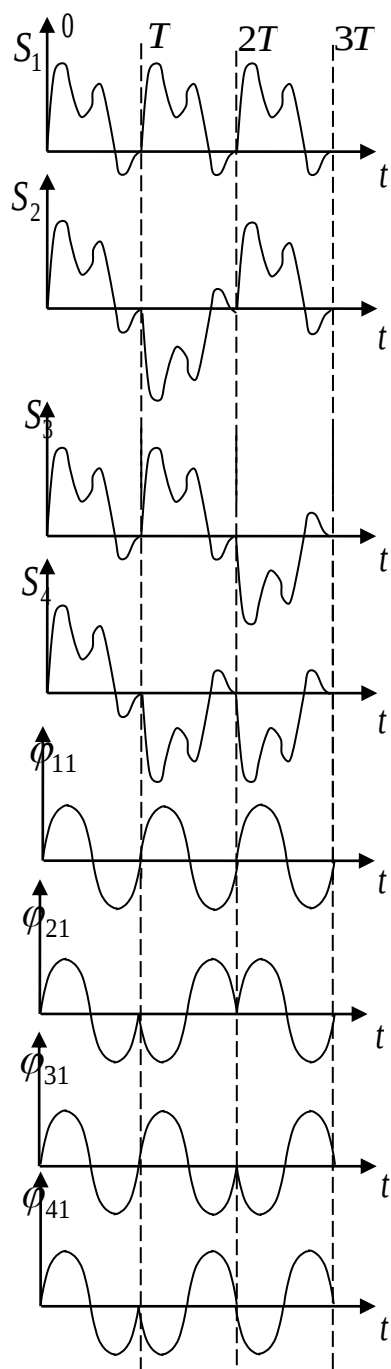


Рис. 5.12. Варіанти сигналу невідомої форми на трьох послідовних
та їх базисні функції

Вимоги до стабільності частоти сигналу і тривалості затримки демодуляторів сигналів із ФРМ-2 такі ж, як для автокореляційних демодуляторів сигналів із ФРМ-1.

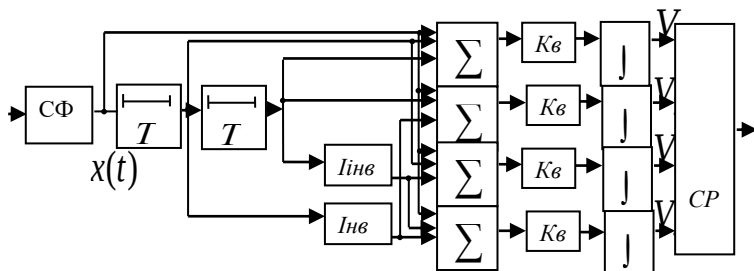


Рис. 5.13. Оптимальний енергетичний (автокореляційний)
демодулятор сигналів з однократною ФРМ-2

При дослідженні автокореляційного прийому сигналу з багатократною ФРМ користуються тими ж методами обробки, що мають загальні закономірності: при збільшенні числа позицій різниці фаз і відповідному збільшенні швидкості передачі завадостійкість помітно зменшується. При переході від однократної ФРМ до двократної з незмінною швидкістю передачі завадостійкість зменшується незначно. Найчастіше використовуються на практиці автокореляційні модеми з двократною ФРМ-1, котрі дозволяють при тій же швидкості передачі, що і модеми з однократною ФРМ-1 майже вдвічі скоротити смугу частот при незначних енергетичних втратах.

5.4. Розробка і дослідження алгоритму обробки сигналу в багатоканальних модемах з ортогональним розподілом каналів

Постановка задачі

У випадку, коли фаза сигналу, що надходить, змінюється повільно, є можливість шляхом аналізу попередніх елементів сигналу, що надходить визначити з певним ступенем точності початкову фазу очікуваних сигналів і реалізувати близький до когерентного, так званий квазікогерентний прийом. Ступінь точності залежить, з

одного боку, від того, з якою швидкістю флукує фаза сигналу, що надходить, а з іншого боку, від відношення енергії до спектральної щільності шуму, оскільки наявність шумів не дозволяє точно вимірити значення фази попередніх елементів [7].

Вплив шуму можна зменшити, проводячи усереднення результатів виміру за великий проміжок часу. Але збільшення часу приводить до появи погрішності в оцінці очікуваної фази внаслідок того, що враховуються відрізки сигналу, прийняті набагато раніше за часом від прийнятого в даний момент елемента сигналу, тобто, менш кориговані з ним по фазі. Таким чином, існує деякий оптимальний інтервал усереднення, тим більше тривалий, чим повільніше змінюється фаза і чим більший рівень шуму в каналі. Застосування когерентного прийому для двійкових сигналів дозволяє одержати енергетичний виграш $\sim 1\text{дБ}$ у порівнянні з оптимальним некогерентним прийомом. У перспективних багатоканальних модемах ТЧ (тональної частоти) каналу для досягнення високих питомих швидкостей передбачається використовувати багатопозиційний сигнал. Для прийому таких сигналів, як відомо, доцільно використовувати когерентний прийом, оскільки при збільшенні числа варіантів сигналу ефективність когерентного прийому в порівнянні з некогерентним складає 3-5 дБ.

До нинішнього часу у багатоканальних модемах за рідкісними винятками використовувався некогерентний прийом. Це викликано, в основному, двома причинами.

Перша – відомі методи реалізації квазікогерентного прийому базувалися, головним чином, на безпосередній обробці фазоманіпульованого сигналу у той час як у багатоканальних модемах з ортогональними сигналами після розподілу каналів мають місце тільки значення проекції кожного початкового сигналу на взаємно ортогональні опорні коливання з довільною початковою фазою. Наприклад, одним з методів формування опорних коливань (ФОК) є метод множення частоти чи фази, що полягає в нелінійному перетворенні ФМ сигналу. Відсутність розподілених каналних сигналів на вході багатоканального модему з ортогональним розподілом вилучає можливість реалізації когерентного прийому на основі подібних методів. Відомі також методи ФОК, засновані на виведенні проекції прийнятого сигналу, підстроювання частоти і фази керованого генератора. При застосуванні таких методів необхідно було б у кожному каналі багатоканального модему використовувати автономне кільце фазового автопідстроювання частоти (ФАП) з керованим генератором, що звичайно ускладнює побудову системи.

Друга причина – реалізація багатоканального модему з когерентним методом прийому на базі елементної бази середньої інтеграції представляла деяке ускладнення і не знайшла поширення. З появою мікропроцесорної техніки з'явилася можливість реалізувати практично будь-які алгоритми, що лежать в основі таких систем [29].

Відомі окремі алгоритми когерентної обробки багаточастотних сигналів призначені для конкретної 16-позиційної конструкції сигналу з квадратурною амплітудною модуляцією QAM-16. Когерентне підстроювання проводиться в такий спосіб: аналізуються проекції прийнятого сигналу на некогерентні осі, потім знаходиться величина розбіжності від положення пропонованих когерентних осей. Ця величина (кут) протягом деякого часу усереднюється, потім підкоректовуються осі і знаходяться нові проекції, після цього приймаються рішення. Розглянемо особливості цих алгоритмів. По-перше, в них усереднена величина піддається нелінійним перетворенням, що погіршує характеристики оцінок у порівнянні з максимально правдоподібним оцінюванням. По-друге, застосування цих алгоритмів має труднощі, тому що вони сконструйовані тільки для однієї визначеної конструкції сигналів. Нарешті, дослідження числа операцій показало, що програмна реалізація демодулятора, яка використовує такого типу алгоритми обробки сигналу, складна.

Таким чином, становить інтерес розробка універсального алгоритму когерентної обробки багатопозиційних сигналів у багатоканальних моделах. Один з можливих підходів до рішення цієї задачі полягає в наступному.

5.4.1. Особливості і шляхи рішення задачі когерентного прийому багаточастотних групових сигналів

Як відомо, розподіл каналних сигналів у демодуляторі когерентного модему здійснюється за допомогою активних фільтрів (кореляторів), у яких обчислюються проекції сигналів на координатній осі приймача. Алгоритм обчислення синфазної квадратурної проекції (X_n, Y_n) k -го каналного сигналу на n -й посилює має вид:

$$\begin{aligned} x_n &= \int_{(n-1)T}^{nT} S(t) \cos \omega k t dt, \\ y_n &= \int_{(n-1)T}^{nT} S(t) \sin \omega k t dt, \end{aligned} \quad (5.4.1)$$

де $S(t)$ – прийнятий груповий сигнал.

Як відомо, алгоритм оптимального демодулятора полягає в обчисленні відстані між прийнятим і варіантами переданого сигналу й у виборі варіанта з найменшою відстанню. Якщо позначити через x_j^*, y_j^* проекції j-го варіанта переданого сигналу ($j = 1, 2, \dots, m$) на ті ж опорні коливання (вираз 5.4.1)

$$\begin{aligned} x_j^* &= \int_{(n-1)T}^{nT} S_j(t) \cos \omega k t dt, \\ y_j^* &= \int_{(n-1)T}^{nT} S_j(t) \sin \omega k t dt, \end{aligned} \quad (5.4.2)$$

то вище зазначену відстань можна представити як:

$$d_n = \sqrt{(x_n - x_j^*)^2 + (y_n - y_j^*)^2} \quad (5.4.3)$$

і, отже, номер переданого сигналу є:

$$i = \arg \min_j \left[(x_n - x_j^*)^2 + (y_n - y_j^*)^2 \right] \quad (5.4.4)$$

Для реалізації алгоритму (5.4.4) необхідно за проекцією прийнятого сигналу X_n, Y_n на деякій множині посилок оцінити значення x_j^*, y_j^* для усіх варіантів сигналу. Алгоритми оцінювання засновані на усередненні тих чи інших параметрів сигналу. При цьому в якості усереднених величин варто вибирати деякі перетворення вихідних параметрів X_n, Y_n , які інваріантні до номера переданого сигналу. Таким інваріантним перетворенням є, зокрема, кут між прийнятим і найближчим до нього переданим варіантами сигналу [118].

Розглянемо метод оцінювання варіантів переданого сигналу шляхом усереднення кута між прийнятим і найближчим до нього переданим вектором сигналу.

Обчислюємо на N посилках середню амплітуду сигналу

$$A_{cp} = \frac{1}{N} \sum_{n=1}^N \sqrt{x_n^2 + y_n^2} . \quad (5.4.5)$$

Визначаємо усереднені значення амплітуд усіх варіантів сигналу

$$A_j = \alpha_j A_{cp}; \quad j = 1, 2, \dots, m, \quad (5.4.6)$$

де α_j – коефіцієнт, що відповідає j -у варіанту сигналу, що залежить від типу обраної конструкції сигналу.

Знаходимо $\cos \Delta\varphi$ і $\sin \Delta\varphi$, де $\Delta\varphi$ – кут між прийнятим вектором X_n, Y_n і дозовленим вектором (оцінкою), найближчим до прийнятого (X_{jn}^*, Y_{jn}^*) , тобто

$$\cos \Delta\varphi_n = \frac{X_n \cdot X_{jn}^* + Y_n \cdot Y_{jn}^*}{\sqrt{X_n^2 + Y_n^2} \cdot \sqrt{X_{jn}^{*2} + Y_{jn}^{*2}}} . \quad (5.4.7)$$

В даному випадку передбачається, що є деякі оцінки проекцій X_j^*, Y_j^* варіантів переданого сигналу, на підставі яких демодулятор прийме рішення, який із сигналів переданий. В якості початкових оцінок можуть бути використані проекції першої з прийнятих посилок.

Усереднюємо отримані $\cos \Delta\varphi_n$ і $\sin \Delta\varphi_n$:

$$\begin{aligned}\cos \Delta \tilde{\varphi}_n &= \frac{1}{N} \sum_{n=1}^N \cos \Delta \varphi_n \\ \sin \Delta \tilde{\varphi}_n &= \frac{1}{N} \sum_{n=1}^N \sin \Delta \varphi_n\end{aligned}$$

(5.4.8)

Тепер варто змінити оцінки проекцій сигналу таким чином, щоб фаза відповідних векторів змінилася на $\Delta \varphi_n$, а амплітуда стала рівної A_j . Уточнені оцінки проекції j-го варіанта x_j^{**} і y_j^{**} дорівнюють:

$$\begin{aligned}x_j^{**} &= A_j \cos(\varphi_j^* + \Delta \tilde{\varphi}_n), \\ y_j^{**} &= A_j \sin(\varphi_j^* + \Delta \tilde{\varphi}_n),\end{aligned}$$

(5.4.9)

чи після тригонометричних перетворень:

$$\begin{aligned}y_j^{**} &= A_j (\sin \varphi_j^* \cos \Delta \tilde{\varphi}_n + \cos \varphi_j^* \sin \Delta \tilde{\varphi}_n), \\ x_j^{**} &= A_j (\cos \varphi_j^* \cos \Delta \tilde{\varphi}_n - \sin \varphi_j^* \sin \Delta \tilde{\varphi}_n).\end{aligned}$$

(5.4.10)

Розкривши значення $\sin \varphi_j^*$ й $\cos \varphi_j^*$, одержимо:

$$\begin{aligned} y_j^{**} &= \alpha_j \left(y_j^* \cos \Delta \tilde{\varphi}_n + x_j^* \sin \Delta \tilde{\varphi}_n \right), \\ x_j^{**} &= \alpha_j \left(x_j^* \cos \Delta \tilde{\varphi}_n - y_j^* \sin \Delta \tilde{\varphi}_n \right), \end{aligned} \quad (5.4.11)$$

де α_j - коефіцієнт, що приймає визначене значення у залежності від номера вектора й вибраної конструкції сигналів. Цей алгоритм когерентного прийому багатопозиційних сигналів у багатоканальних модемах не є алгоритмом максимально правдоподібного оцінювання. Розглянутий алгоритм, на відміну від відомих, є універсальним, тобто його можна застосовувати в системах багатопозиційних сигналів.

При зміні системи сигналів змінюються тільки коефіцієнти α_j у формулі (5.4.11).

Цей алгоритм орієнтований на реалізацію демодулятора за допомогою мікропроцесорів. Він, проте, володіє тим недоліком, що внаслідок нелінійного перетворення (див. формулу 5.4.7) оцінки проекції варіантів сигналу не є ефективними.

Для одержання ефективних оцінок варіантів сигналу x_j і y_j варто скористатися правилом максимальної правдоподібності, що у сполученні з правилом вибору (5.4.4) дасть оптимальний алгоритм когерентної обробки розглянутих групових сигналів багатоканальних модемів.

5.4.2. Синтез оптимального алгоритму когерентної обробки багаточастотних групових сигналів

Сигнали з амплітудно-фазовою й амплітудно-фазорізницевою модуляцією (АФРМ) звичайно застосовуються для досягнення високих питомих швидкостей цифрової передачі (3 біт/сек. Гц і більш) у каналах з жорстким обмеженням смуги частот і порівняно доброю енергетикою. Найбільш відомими і часто використовуваними є 16-ти позиційні системи АФМ і АФРМ сигналів, що відповідає 4-и кратній маніпуляції. Відомі випадки практичного використання сигналів і більш високої кратності.

Розглянемо загальний випадок цифрової передачі за допомогою m – позиційного сигналу з довільними амплітудами $a_1, a_2 \dots a$ і початковими фазами

$\varphi_1, \varphi_2, \dots, \varphi_m$, не роблячи поки що різниці між АФМ і АФРМ. При такій

постановці задачі i -й варіант переданого сигналу можна представити в наступному вигляді:

$$S_i(t) = a_i \sin(\omega t + \varphi_i), \quad i = 1, 2, \dots, m. \quad (5.4.12)$$

У каналі з гауссовським некорельованим шумом оптимальний алгоритм прийому сигналів (5.4.12) може бути сформульований таким чином: фіксується переданий i -й варіант сигналу, якщо при усіх $j \neq i$ має місце нерівність

$$\int_0^T [x(t) - S_i(t)]^2 dt < \int_0^T [x(t) - S_j(t)]^2 dt, \quad (5.4.13)$$

де $x(t)$ – прийнятий сигнал, T - тривалість посилки.

$$i = \arg \min_0^T \int [x(t) - S_i(t)]^2 dt. \quad (5.4.14)$$

При цифровій обробці зручно переходити від високочастотного сигналу (5.4.13) до його відображення через координати в двомірному просторі, що на практиці відповідає, наприклад, операціям перенесення чи спектру поділу ортогональних каналних сигналів у багатоканальній системі.

Отже, нехай відомі обчислені на інтервалі однієї посилки проєкції прийнятого сигналу $x(t)$ і сигналів (5.4.12) на опорні коливання з довільною фазою Φ_0 :

$$\left. \begin{aligned} x_0 &= \int_0^T x(t) \cos(\omega t + \varphi_0) dt, \\ y_0 &= \int_0^T x(t) \sin(\omega t + \varphi_0) dt, \end{aligned} \right\} \quad (5.4.15)$$

$$\left. \begin{aligned} x_i &= \int_0^T s_i(t) \cos(\omega t + \varphi_i) dt, \\ y_i &= \int_0^T s_i(t) \sin(\omega t + \varphi_i) dt, \end{aligned} \right\} \quad (5.4.16)$$

$i = 1, 2, \dots, m.$

Тоді оптимальний алгоритм (5.4.14) можна представити в наступному виді

$$i = \arg \min \int_0^T \left[(x_0 - x_j)^2 + (y_0 - y_j)^2 \right] dt, \quad (5.4.17)$$

Вхідні в (9.17) $i = 1, 2, \dots, m$ величини X_0 і Y_0 визначаються, як видно з (5.4.15), у результаті обробки поточної прийнятої посилки сигналу, а величини X_j і

y_j , число яких дорівнює $2m$, повинне бути відомим апріорно чи в обчисленні (оціненні) у процесі прийому попередніх посилок сигналу.

Для обчислення оцінок проекції варіантом сигналу X_j і Y_j скористаємося методом приведення й усереднення проекції прийнятого сигналу. На роль усереднених величин виберемо для визначеності проекції першого варіанту (5.4.12), тобто величини X_1 і Y_1 (див. 5.4.16), до них же будемо перетворювати (приводити) інші варіанти прийнятого сигналу в процесі підстроювання по інформаційному сигналі. Якщо прийнятий сигнал $x(t)$ містить на інтервалі N посилок сигналу $S_1(t)$ в суміші з гауссовським шумом, то, як відомо, максимально правдоподібні оцінки цих величин $\tilde{X}_1$ і $\tilde{Y}_1$ рівні:

$$\left. \begin{aligned} \tilde{X}_1 &= \frac{1}{N} \sum_{n=1}^N x_{0n}, \\ \tilde{Y}_1 &= \frac{1}{N} \sum_{n=1}^N y_{0n}, \end{aligned} \right\} \quad (5.4.18)$$

де x_{0n} і y_{0n} - величини проекції (5.4.16) на інтервалі n -ої послідовності, а хвилястою лінією над X_1 і Y_1 відзначається, що це оцінки. Оцінки (5.4.18) є незміщеними й ефективними. Їх можна сформувати у незміщені й ефективні оцінки проекції всіх інших варіантів сигналу, що входять в оптимальний алгоритм (5.4.17). Для цього введемо позначення $\varphi_j = \varphi_1 + \Delta\varphi_j$, утворимо проекцію в такий спосіб:

$$\begin{aligned}
x_j &= \int_0^T a_j \sin(\omega t + \varphi_j) a_0 \sin(\omega t + \varphi_0) dt = \frac{a_j}{a_1} \int_0^T a_1 \sin(\omega t + \varphi_1 + \Delta\varphi_j) a_0 \sin(\omega t + \varphi_0) dt = \\
&= \frac{a_j}{a_1} [\cos \Delta\varphi_j \int_0^T a_1 \sin(\omega t + \varphi_1) Q_0 \sin(\omega t + \varphi_0) dt + \sin \Delta\varphi_j \int_0^T Q_1 \sin(\omega t + \varphi_1) a_0 \sin(\omega t + \varphi_0) dt] = \\
&= \frac{a_j}{a_1} [\cos \Delta\varphi_j \int_0^T a_1 \sin(\omega t + \varphi_1) a_0 \sin(\omega t + \varphi_0) dt - \sin \Delta\varphi_j \int_0^T a_1 \sin(\omega t + \varphi_1) a_0 \cos(\omega t + \varphi_0) dt] = \\
&= \frac{a_j}{a_1} (x_1 \cos \Delta\varphi_j - y_1 \sin \Delta\varphi_j).
\end{aligned}$$

Аналогічним чином одержуємо проекцію y_j . Замінивши тепер величини x_1 і y_1 їхніми оцінками, отримаємо:

$$\left. \begin{aligned} \tilde{x}_j &= \frac{a_j}{a_1} (\tilde{x}_1 \cos \Delta\varphi_j - \tilde{y}_1 \sin \Delta\varphi_j), \\ \tilde{y}_j &= \frac{a_j}{a_1} (\tilde{x}_1 \sin \Delta\varphi_j + \tilde{y}_1 \cos \Delta\varphi_j), \end{aligned} \right\} \quad (5.4.20)$$

де $\Delta\varphi_j$ - відома різниця фаз між сигналами $S_j(t)$ і $S_1(t)$.

Зазначимо, що при обчисленні оцінок по (5.4.20) немає необхідності мати інформацію про амплітуди варіантів сигналу a_j і a_1 , а досить знати відношення цих амплітуд a_j / a_1 .

Отримані алгоритми вирішують задачу когерентного прийому багатопозиційного АФМ сигналу при наявності спеціального синхросигналу, що, наприклад, передує передачі інформаційних посилок: по (5.4.16) обчислюються проекції синхросигналу на опорне коливання з довільною початковою фазою, потім по (5.4.18) визначаються оцінки проекцій першого варіанта сигналу $\tilde{x}_1$ і $\tilde{y}_1$, по (5.4.19) обчислюються оцінки проекцій усіх m варіантів сигналу і, нарешті, отримані оцінки $\tilde{x}_j$ і $\tilde{y}_j$; усі варіанти сигналу підставляються замість x_j і y_j в алгоритм (5.4.17), по якому проводиться демодуляція інформаційних посилок.

Підкреслимо, що розглянутий алгоритм орієнтований на прийом сигналу з абсолютною ФМ, тому що наявність синхросигналу виключає невизначеність початкової фази, що заважає застосуванню абсолютної фазової модуляції.

Повернемося тепер до розглянутого алгоритму і пристосуємо його до випадку – практично найбільш важливого – коли синхросигнал відсутній і “підстроювання” проекції зразків сигналу приходится вести безпосередньо по інформаційних посилках. У даному випадку необхідно усереднювати не проекції прийнятого сигналу, а зведені проекції. При цьому операція зведення полягає в перетворенні прийнятих проекцій до проекцій, наприклад, першого варіанта сигналу з використанням прийнятого рішення про переданий зразок сигналу.

Нехай, як і раніше, $\Delta\tilde{\varphi}_n$ - різниця фаз між варіантом сигналу, на користь якого прийняте рішення на n -ій посилці і першим варіантом сигналу, $\tilde{a}_n$ - амплітуда сигналу, на користь якого прийняте рішення на n -ій посилці. Тоді $\Delta\tilde{\varphi}_n$ приймає значення з дискретної множини дозволених фаз, визначених в (5.4.12). Що стосується $\tilde{a}_n$, то ця величина дорівнює фактичній амплітуді, прийнятій на n -ій посилці змішаних сигналу з шумом, однак надалі вона ототожнюється з амплітудою того варіанту сигналу, на користь якого винесене рішення на n -ій посилці. Хвилястою лінією в обох випадках підкреслюється, що ці оцінки фактично можуть бути помилковими. Тоді зведені проекції X_{1n} і Y_{1n} прийнятого сигналу на n -ій посилці обчислюються через прийняті проекції X_{on} і Y_{on} за формулами:

$$\begin{aligned} x_{1n} &= \frac{a_1}{\tilde{a}_n} (x_{0n} \cos \Delta\tilde{\varphi}_n + y_{0n} \sin \Delta\tilde{\varphi}_n) \\ y_{1n} &= \frac{a_1}{\tilde{a}_n} (y_{0n} \cos \Delta\tilde{\varphi}_n - x_{0n} \sin \Delta\tilde{\varphi}_n) \end{aligned} \quad (5.4.21)$$

Підкреслимо, що $\tilde{a}_n$ і $\Delta\tilde{\varphi}_n$ визначаються рішенням про переданий на n -ій посилці варіант сигналу, прийнятим за результатами обробки величини X_{on} і Y_{on} .

Далі, як і в алгоритмі прийому по синхросигналу, величини (5.4.21) усереднюються:

$$\left. \begin{aligned} \tilde{x}_1 &= \frac{1}{N} \sum_{n=1}^N \frac{a_1}{\tilde{a}_n} (x_{on} \cos \Delta \tilde{\varphi}_n + y_{on} \sin \Delta \tilde{\varphi}_n) \\ \tilde{y}_1 &= \frac{1}{N} \sum_{n=1}^N \frac{a_1}{\tilde{a}_n} (y_{on} \cos \Delta \tilde{\varphi}_n - x_{on} \sin \Delta \tilde{\varphi}_n) \end{aligned} \right\} \quad (5.4.22)$$

Відмінність алгоритму (5.4.22) від (5.4.18) полягає в тому, що в (5.4.18) усереднення проводиться на інтервалі синхросигналу, а в (5.4.22) – на “плинному” інтервалі в М посилок, що передують оброблюваній в даний момент.

Відзначимо, що при обчисленні оцінок по (5.4.22) немає необхідності в апіорних зведеннях про середню потужність сигналу, що знаходиться, тому що в цей алгоритм входять тільки відношення амплітуд.

Таким чином, співвідношення (5.4.21), (5.4.22) разом з (5.4.17) утворюють шуканий алгоритм когерентної обробки.

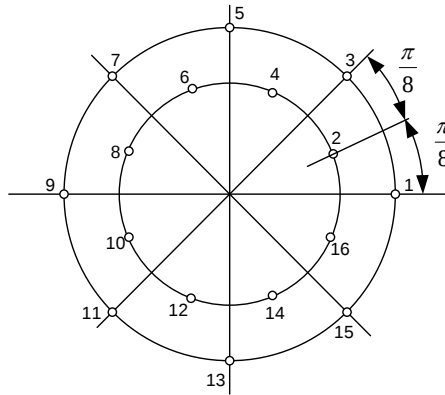
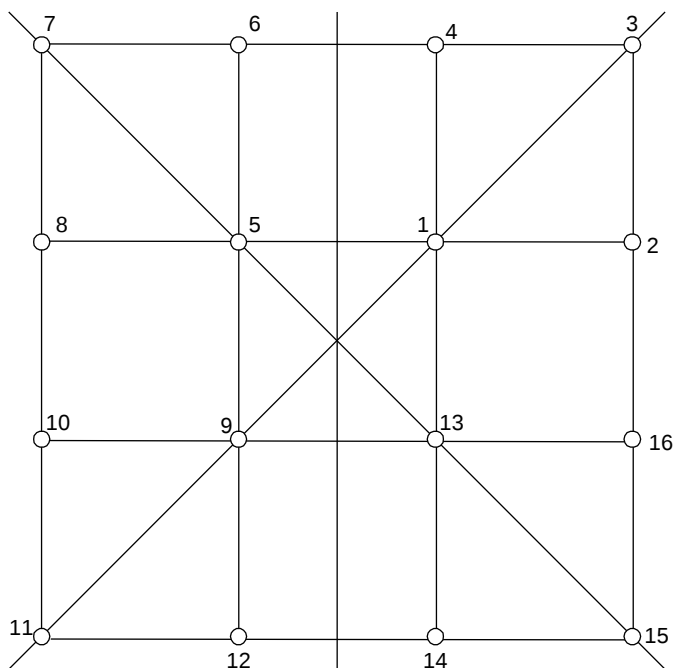


Рис.5.1

Варіанти різниць фаз $\Delta \varphi_j$, необхідні для обчислення по формулах (5.4.20) і (5.4.21), неважко визначити за рис.5.2.

Рис.5.2. Шістнадцятипозиційний сигнал з квадратурно-амплітудною модуляцією (QAM)



Для деяких варіантів зручно обчислити безпосередньо $\cos \Delta \varphi_j$ і $\sin \Delta \varphi_j$.

Наприклад, для сигналу з номером 4 маємо:

$$\sin \Delta \varphi_j = \sin(\varphi - \frac{\pi}{4}) = \frac{1}{2}(\sin \varphi_4 - \cos \varphi_4) = \frac{1}{\sqrt{2}} \frac{y_4}{\sqrt{x_4^2 + y_4^2}} - \frac{x_4}{\sqrt{x_4^2 + y_4^2}} = \frac{1}{\sqrt{2}} (\frac{3}{\sqrt{10}} - \frac{1}{\sqrt{10}}) = \frac{1}{\sqrt{5}}$$

$$\cos \Delta \varphi_j = \cos(\varphi_4 - \frac{\pi}{4}) = \frac{1}{2}(\sin \varphi_4 + \cos \varphi_4) = \frac{2}{\sqrt{5}}.$$

(5.4.24)

Таким чином, при прийомі 4-го варіанту сигналу зведені проекції згідно (5.4.21), (5.4.22) і (5.4.23), обчислюються за формулою:

$$\begin{aligned}
 x_{1n} &= \frac{a_1}{a_4} \left(x_{on} \frac{2}{\sqrt{5}} - y_{on} \frac{1}{\sqrt{5}} \right) = \frac{1}{5} (2x_{on} - y_{on}), \\
 y_{1n} &= \frac{a_1}{a_4} \left(y_{on} \frac{2}{\sqrt{5}} - x_{on} \frac{1}{\sqrt{5}} \right) = \frac{1}{5} (2y_{on} - x_{on}).
 \end{aligned}
 \tag{5.4.25}$$

Аналогічним чином, для 2-го варіанта сигналу одержуємо $\sin \Delta \varphi_2 = -\frac{1}{\sqrt{5}}$, $\cos \Delta \varphi_2 = \frac{2}{\sqrt{5}}$, а зведені проекції при прийомі 2-го варіанта сигналу дорівнюють:

$$\begin{aligned}
 x_{1n} &= \frac{1}{5} (2x_{on} - y_{on}) \\
 y_{1n} &= \frac{1}{5} (2y_{on} - x_{on})
 \end{aligned}
 \tag{5.4.26}$$

Тригонометричні функції інших фазових вузлів системи сигналів рис.4.21 одержуємо з фазових кутів сигналів 1, 2, 3, 4.

Дійсно, ця система сигналів формується шляхом подвійної фазової модуляції, так що сигнали 5, 6, 7, 8 будуються поворотом сигналів 1, 2, 3 і 4 - на 90° , сигнали 9, 10, 11 і 12 – поворотом сигналів на 180° , сигнали 13, 14, 15, 16 – поворотом на - 90° .

Ці зведені проекції усереднюються, у результаті чого одержуємо оцінки проекцій першого варіанта сигналу $\tilde{x}_1$ і $\tilde{y}_1$. По формулах (5.4.20) далі обчислюються оцінки всіх проекцій $\tilde{x}_j$ і $\tilde{y}_j$.

Слід зазначити, що розглянуті тут алгоритми когерентної обробки багатопозиційних АФМ чи АФРМ сигналів особливо зручні для багатоканальних (багаточастотних) систем з ортогональними канальними сигналами, тому що в цих системах для поділу ортогональних сигналів використовуються ті ж, що й у

розглянутих алгоритмах, процедура обчислення проекцій прийнятого сигналу на два взаємоортогональних опорних коливань з довільною початковою фазою .

У багаточастотних демодуляторах опорні коливання всіх каналів звичайно формуються від однієї довільної частини. Разом з тим початкові фази каналних сигналів мають різні зсуви, а в ряду випадків вони взагалі слабо зв'язані (наприклад, у радіоканалах із селективними згасаннями – некорельовані). Внаслідок цього в багатоканальних демодуляторах важко використовувати методи формування опорних коливань, заснованих на підстроюванні фази керованого генератора.

Описані алгоритми когерентної обробки не потребують підстроювання фази опорних коливань каналних сигналів і дозволяють досить просто реалізувати як ортогональний поділ сигналів, так і когерентний прийом. Дійсно, для поділу каналів використовуються ті ж опорні коливання, що і при оптимальному некогерентному прийомі, а когерентний прийом здійснюється на підставі підстроювання варіантів сигналу в кожному каналі багатоканального модему [30].

5.4.3. Моделювання когерентного багаточастотного модему.

Задачі, що можуть бути вирішені шляхом моделювання, формуються таким чином. Зауважимо, що запропоновані алгоритми когерентної обробки багатопозиційних АФМ чи АФРМ сигналів особливо зручні для багатоканальних (багаточастотних) систем з ортогональними каналними сигналами, тому що в цих системах для розділу ортогональних сигналів використовуються ті ж процедури обчислення проекцій прийнятого сигналу на два взаємоортогональних опорних коливання з довільною початковою фазою.

У багаточастотних демодуляторах опорні коливання всіх каналів звичайно формуються від загального задаючого генератора. Разом з тим початкові фази каналних сигналів мають різні зсуви, а в ряду випадків вони взагалі слабо зв'язані (наприклад, у радіоканалах із селективними згасаннями – некорельовані). Унаслідок цього в багатоканальних демодуляторах важко використовувати методи формування опорних коливань, заснованих на підстроюванні фази керованого генератора.

Описані алгоритми когерентної обробки не потребують підстроювання фази опорних коливань каналних сигналів і дозволяють досить просто реалізувати як ортогональний поділ сигналів, так і когерентний прийом. Дійсно, для поділу каналів використовуються ті ж опорні коливання, що і при оптимальному некогерентному прийомі. Когерентний прийом здійснюється на підставі підстроювання варіантів сигналу в кожному каналі багатоканального модему.

1. Визначення ймовірності помилки при ідеальних опорних коливань для різних відношень сигнал/завада. Це дозволить визначити потенційну

завадостійкість демодулятора при використанні різних груп багатопозиційних сигналів.

2. Визначення ймовірності помилки при формуванні варіантів сигналів для різних інтервалів усереднення. Визначення мінімального інтервалу усереднення, при якому завадостійкість близька до потенційної.

3. Порівняння різних систем сигналів по завадостійкості при розглянутому методі обробки.

4. Визначення тривалості початкового входження в зв'язок, а також тривалість входження після різних за величиною стрибків фази сигналу.

При дослідженні був використаний метод статистичного моделювання, у якому всі процеси задаються у виді алгоритмів і орієнтованих на використанні для розрахунків на ЕОМ.

Як джерело дискретної інформації для моделювання була використана програма формування псевдовипадкової послідовності чисел від 0 до 15. Моделювання діяння адитивного флуктуаційного шуму заснованого на формуванні чисел з нормальним законом розподілу, нульовою середньою і заданою дисперсією. Ці числа, у свою чергу, можуть бути отримані з послідовності псевдовипадкових чисел, рівномірно розподілених на інтервалі $(0, 1)$, на підставі центральної граничної теореми теорії ймовірностей.

Моделююча програма складається з головної і дев'яти підпрограм.

У задачу головної моделюючої програми входить: введення робочих масивів; завдання початкових значень окремим змінним і масивам, введення вихідних даних сигналів, що характеризують задану систему, і відношення сигнал/шум; звертання до різних підпрограм. Головною програмою також здійснюється підрахунок ймовірності помилки, керування роботою всіх блоків моделі і вивід на друк (принтер) необхідних вихідних даних і результатів.

Для зручності опису позначимо підпрограми відповідно номерами 1, 2,...,9.

Підпрограма 1 установлює початкові параметри системи і задає необхідне число посилок усереднення.

Підпрограма 2 забезпечує підрахунок витраченого часу при розрахунку.

Підпрограма 3 здійснює видачу випадкових чисел від 0 до 15, а також при необхідності завдання різних стрибків фаз і амплітуди вхідного сигналу.

Підпрограма 4 виконує об'єднання вхідного сигналу з шумом; відносно кодування сигналу; обчислення по формулі 3.17.

Підпрограма 5 проводить обчислення за формулами 3.21 і 3.22.

Підпрограма 6 здійснює обчислення дисперсії завади у залежності від заданої величини відношень сигнал/шум.

Підпрограми 7-8 задають систему сигналів. Значимо, що при переході в іншу систему сигналів потрібні лише невеликі зміни в підпрограмах 7 і 8.

Розрахунок проводився окремо для кожного сигнального сузір'я за допомогою системи „MathCAD ® v.7.0 Professional Edition”. Текст програми наведено в додатку. В результаті зроблених розрахунків визначені ймовірності похибки для різних відношень сигнал/завада. Ймовірність похибки визначалася таким чином. Порівнювалися номери переданих і прийнятих векторів сигналів. Потім визначалося відношення помилкове прийнятих векторів сигналу до загального числа переданих векторів. Причому, у залежності від очікуваної ймовірності похибки був заданий необхідний обсяг вибірки Q .

При визначенні обсягу вибірки доцільно скористатися результатами. Як відомо [10] для кожного Q і заданої ймовірності “довірення” β побудувати область, всередині якої значення ймовірності помилки ($P_{ном}$) сумісно зі знайденим у досвіді значенням частоти ($P_{ном}^*$). Зображені криві, що обмежують також області для різних Q при ймовірності “довірення” $\beta = 0.9$. Наприклад, при $P_{ном}^* = 0.1$ для забезпечення інтервалу “довірення” [0.09 - 0.11], необхідний обсяг вибірки складе ≈ 200 , при $P_{ном}^* = 0.07$ для забезпечення інтервалу “довірення” [0.065 - 0.075] необхідний обсяг складе ≈ 1000 . Таким чином, обсяг вибірки необхідно вибирати не менш $\frac{20}{P_{ном}^*}$.

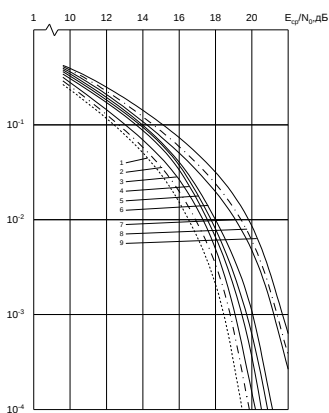


Рис. 5.3 Залежність ймовірності помилок від відношення сигнал-шум

Графіки, приведені на рис.5.3 ілюструють залежність імовірності помилки $P_{\text{пом}}$ від відношення енергії сигналу до спектральної щільності потужності шуму для системи сигналів АФМ, зображеної на рис.5.1. Криві на рисунку позначені цифрами 1, 2...9. Крива 1 характеризує потенційну завадостійкість даної системи сигналів при строго когерентному прийомі. Крива 2 отримана в результаті моделювання при точних опорних коливаннях і характеризує потенційну завадостійкість розробленого алгоритму обробки. Криві 3...7 отримані в результаті моделювання і характеризують завадостійкість демодулятора при когерентному прийомі сигналів для різних інтервалів усереднення M , що відповідно вибиралися рівними 1, 5, 10, 20 і 100.

- $M = 100$ (крива 3)

- $M = 20$ (крива 4)

- $M = 10$ (крива 5)

- $M = 5$ (крива 6)

- $M = 1$ (крива 7)

Крива 8 характеризує теоретичну завадостійкість для 4 - кратної ФРМ при когерентному методі прийому. Крива 9 отримана експериментально в результаті лабораторних досліджень 48 - канального модему за допомогою імітатора каналу Т4 типу "канал 2". Ця крива визначає реальну завадостійкість демодулятора 48 - канального модему з оптимальним некогерентним методом прийому, у якому використаний 16 - позиційний сигнал, приведений на рис.9/3.7, (χ_2 у два рази більше ніж χ_1).

Аналіз цих графіків (рис.5.3) підтверджує висновок про перевагу використання когерентних методів прийому для багатопозиційних сигналів. Енергетичний виграш можна одержати як від застосування більш ефективної системи сигналів (наприклад, QAM чи АФМ), так і від використання розробленого алгоритму когерентного прийому.

У даному випадку, порівнявши криву 3 і криву 9, видно, що сумарний виграш складає ≈ 3 дБ.

Мінімальний інтервал усереднення, обумовлений числом посилок M , при якому реальна завадостійкість близька до потенційного (рис.5.2), складає ≥ 20 посилок.

На рис.5.4 зображені криві, що характеризують завадостійкість когерентного прийому сигналів для різних M при використанні конструкції сигналу QAM-16, представленої на рис.5.2. Криві позначені цифрами 1, 2...7, де:

- крива 1 - теоретична завадостійкість для цієї системи сигналів;
- крива 2 - отримана в результаті моделювання для точних опорних коливань;
- крива 3 - для $M = 100$;
- крива 4 - для $M = 20$;
- крива 5 - для $M = 10$;
- крива 6 - для $M = 5$;
- крива 7 - для $M = 1$;

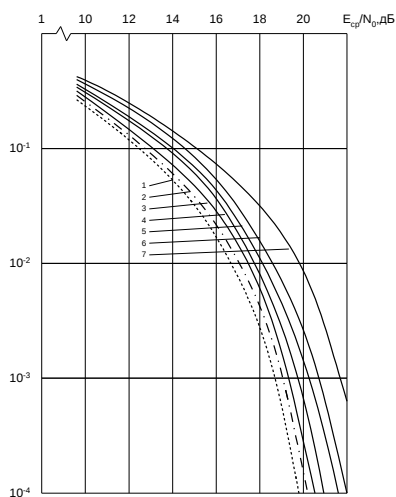


Рис.5.4. Залежність імовірності помилок від відношення сигнал-шум

На графіку представлена кількісна залежність наближення до ідеального варіанта (ідеальним є варіант, при якому опорні коливання - точні; на графіках рис.5.3 і рис.5.4 - це криві 2) при різних інтервалах усереднення M для систем сигналів QAM і

АФМ, причому $\Delta = h_M^2 - h_{n0}^2$. Для обох систем сигналів число усереднених посилок, при якому завадостійкість близька до потенційної, повинне бути не менш 20.

На графіку рис.5.5 представлені криві, що характеризують завадостійкість двох систем сигналів QAM і АФМ при $M=100$. Порівняння цих кривих показує, що застосування системи сигналів, позначеної АФМ дає якийсь (незначний) енергетичний виграш у порівнянні з QAM, однак, іноді на практиці доцільно використовувати і систему QAM, якщо це спрощує реалізацію демодулятора. Потім порівнювалася послідовність векторів переданого і прийнятого сигналів, і фіксувався номер посилки, починаючи з якої ці послідовності збігалися. Це дозволило визначити кількість неправильно прийнятих посилок сигналу, а, отже, і час входження в синхронізм.

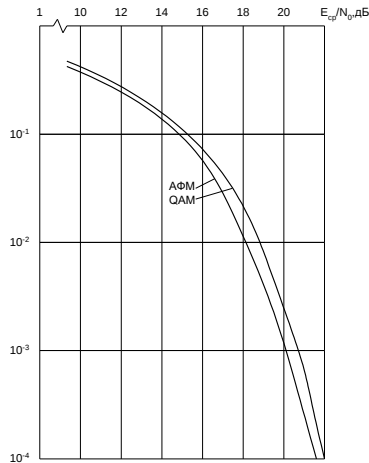
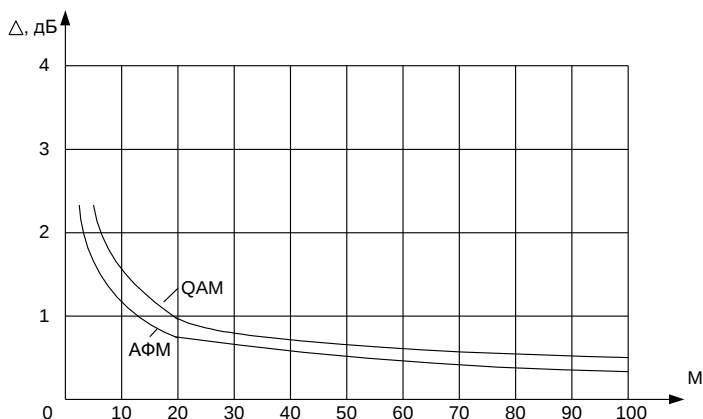


Рис.5.5. Залежність імовірності помилок від відношення сигнал-шум для системи сигналів АФМ і QAM

Відзначимо, що в процесі входження в синхронізм усереднення проводиться не на плинному інтервалі в M посилок, а на інтервалі прийнятих посилок сигналу доти, поки кількість прийнятих посилок не досягне M . Тому, визначаючи, що система входить у синхронізм через інтервал часу до k посилок, необхідно врахувати, що задана завадостійкість, буде забезпечуватися через інтервал часу M посилок. На рис.5.6 представлена кількісна залежність завадостійкості демодулятора до ідеального варіанту при різних M .

Рис.5.6. Кількісна залежність наближена до ідеального варіанту при різних M



Графіки, представлені на рис.5.7 і рис.5.8, ілюструють залежність початкового входження в синхронізм (рис.5.7), а також залежність тривалості входження в синхронізм після стрибка фази (рис.5.8) від відношення сигнал/шум і величини розбіжності переданого сигналу від оцінок варіантів сигналу в демодуляторі і величини стрибка. По осі абсцис відкладені різні кути ($\Delta\varphi$, град), що характеризують розбіжність переданого сигналу від оцінок варіантів сигналу в демодуляторі (рис.5.8) і величину стрибка фази переданого сигналу. По осі ординат - відповідно тривалість початкового входження і тривалість входження в синхронізм після стрибка фази. Кожна крива отримана для визначення відношенням сигнал/шум при числі усереднених посилок рівному 10 ($M=10$).

Рис.5.7. Залежність тривалості входження у зв'язок від величини стрибка фази при різних відношеннях сигнал-шум

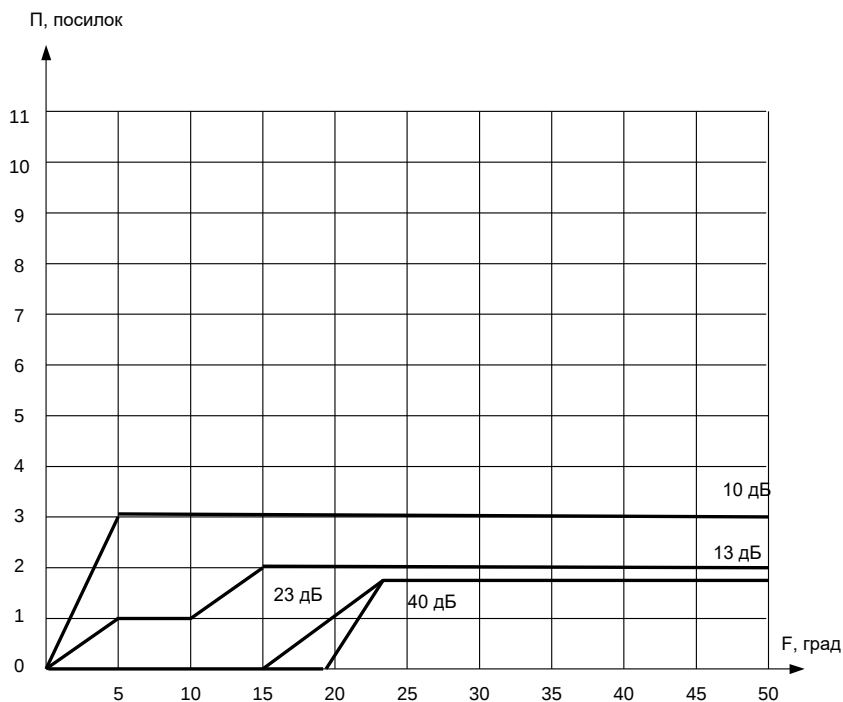
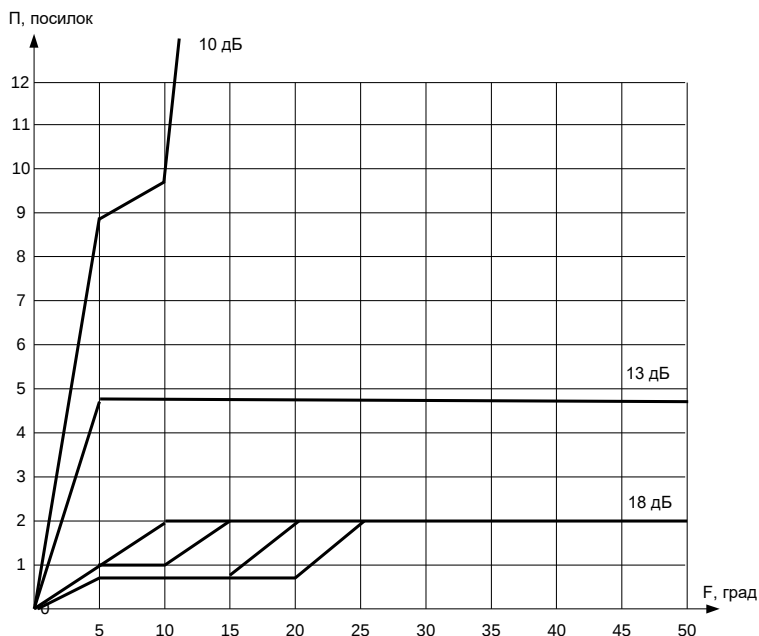


Рис.5.8. Залежність тривалості початкового входження при різних відношеннях сигнал - шум (для $M=10$)



Можна зробити висновок, що при відношенні сигнал/шум не менше 20дБ тривалість початкового входження в синхронізм, а також після стрибка фази складає ≈ 2 посилки.

Отже:

У дослідженні і представленні універсальні алгоритми оптимального когерентного прийому багатопозиційних АФМ і АФРМ сигналів багатоканальних модемів з ортогональними сигналами.

1. Розроблений універсальний алгоритм оптимального прийому багатопозиційних АФМ сигналів в багатоканальних модемах з взаємно ортогональними сигналами, оснований на максимально правдоподібному оцінюванні варіантів сигналу.

2. Розроблені конкретні алгоритми когерентної обробки ефективних систем багатопозиційних сигналів.

3. За допомогою статичного моделювання на ЕОМ знайдені характеристики завадостійкості розроблених алгоритмів в залежності від тривалості інтервалу усереднення оцінюваних параметрів.

4. Визначено, що завадостійкість розробленого алгоритму когерентного прийому сигналів з плинно змінними параметрами наближена до потенційного при тривалості усереднення більше 20 посилок.

Контрольні питання

1. Дайте характеристику некогерентного методу прийому.
2. Які особливості автокореляційного прийому?
3. В яких каналах передавання інформації доцільно використовувати автокореляційний прийом сигналів із ФРМ-2?
4. Який висновок доцільно зробити при порівняльному аналізі оптимальних некогерентних та автокореляційних методів обробки сигналів?

РОЗДІЛ 6. РОЗРАХУНОК ОБСЯГУ КЕРУЮЧОЇ ІНФОРМАЦІЇ ДЛЯ ЗАБЕЗПЕЧЕННЯ ТОЧНОСТІ ПОКАЗНИКІВ ЯКОСТІ СИСТЕМ І МЕРЕЖ ЗВ'ЯЗКУ

6.1. Кількісна міра інформації

Призначення будь-якої системи зв'язку – передати у заданий проміжок часу якомога більше достовірних відомостей від джерела до приймача. Проблема достовірності розглядалась в теорії завадостійкості, а в теорії інформації вивчаються кількісні закономірності передачі, збереження та обробки інформації. Зокрема, найбільша увага приділяється визначенню середньої швидкості передачі інформації та вирішенню задачі максимізації цієї швидкості шляхом застосування відповідного кодування. Граничні співвідношення теорії інформації дозволяють оцінити ефективність різних систем зв'язку і встановити умови узгодження в інформаційному відношенні джерела з каналом і канала зі споживачем.

Для дослідження цих питань з загальних позицій необхідно насамперед встановити універсальну кількісну міру інформації, що не залежить від конкретної фізичної природи повідомлень, що передаються.

При визначенні одиниці кількості інформації слід враховувати, що кількість інформації в повідомленні про деякі події істотно залежать від ймовірності цієї події.

$$I(a) = \log_k \frac{1}{P(a)} = -\log_k P(a),$$

де $P(a)$ – ймовірність події a , k – основа логарифму.

Тому що $0 < P(a) < 1$, то $I(a)$ – величина завжди додатна та скінченна.

При $P(a) = 1$ кількість інформації дорівнює нулю, тобто повідомлення про відому подію ніякої інформації не несе.

Логарифмічна міра має природню в даному випадку властивість адитивності, згідно якої кількість інформації, яка міститься в декількох незалежних повідомленнях, дорівнює сумі кількості інформації в кожному з них.

Дійсно, якщо спільна ймовірність n незалежних повідомлень $P(a_1 \text{ у } a_2 \text{ у } \dots \text{ у } a_n) = P(a_1)P(a_2) \dots P(a_n)$, то кількість інформації в цих повідомленнях дорівнює

$$I(a_1 \text{ у } a_2 \dots a_n) = -\log_k P(a_1 \text{ у } a_2 \dots a_n) = -\sum_{i=1}^n \log_k P(a_i) = \sum_{i=1}^n I(a_i).$$

При $k = 2$ кількість інформації записується в двійкових одиницях.

$$I(a_i) = \log_2 P(a_i), \text{ дв.од.}$$

Двійкову одиницю іноді називають біт.

Двійковій одиниці інформації, тобто інформації в один біт, відповідає інформація, що міститься в повідомленні, яке прийнято по каналу зв'язку без завад, при апріорній ймовірності даного повідомлення $P(a) = 0,5$.

$$I(a) = \log_2 \frac{1}{P(a)} = -\log_2 P(a) = \log_2 2 = 1.$$

Формула $I(a) = \log_k \frac{1}{P(a)} = -\log_k P(a)$ дозволяє обчислити

кількість інформації в повідомленнях, ймовірність яких відрізняється від нуля. Це в свою чергу передбачує, що повідомлення дискретні та кількість їх обмежена. В такому випадку прийнято говорити про ансамбль повідомлень, який описується сукупністю можливих повідомлень і їх ймовірностей.

$$\left\{ \begin{array}{c} a_1, \quad a_2, \quad \dots, \quad a_m \\ P(a_1), P(a_2), \dots, P(a_m) \end{array} \right\}$$

До визначення інформації можна наблизитись і з іншого боку.

Будемо розглядати як повідомлення не окрему літеру, а ціле слово. Якщо всі літери рівноімовірні і

$$P_{сл} = 1/N,$$

де $N = m^n$ – кількість слів.

В цьому випадку можна записати:

$$I_{сл} = -\log P_{сл} = -\log(1/m^n) = n \log m.$$

Для двійкового каналу $m = 2$. Тоді $n \log 2 = n$ дв.од.

В загальному випадку при передаванні повідомлень невизначеність знімається не повністю. Так в каналі з шумами можливі помилки. По прийнятому повідомленню тільки з певною ймовірністю

$$P(u/v) < 1,$$

де U – передане повідомлення, V – прийняте повідомлення,

можна судити про те, що було передане повідомлення U . Тому після отримання повідомлення залишиться певна невизначеність, що характеризується величиною апостеріорної ймовірності $P(u/v)$, а кількість інформації, яка міститься в сигналі V , визначається ступенем зменшення невизначеності при його прийомі.

Якщо $P(u)$ – апіорна ймовірність, то кількість інформації в прийнятому повідомленні V відносно переданого повідомлення U , очевидно, буде дорівнювати

$$I(u,v) = \log \frac{1}{P(u)} - \log \frac{1}{P(u/v)} = -\log P(u) + \log P(u/v) = \log \frac{P(u/v)}{P(u)}.$$

Цей вираз можна розглядати також як різницю між кількістю інформації, яка надійшла від джерела повідомлення, і тою кількістю інформації, яка втрачена в каналі за рахунок дії шумів.

Наведена вище кількісна міра інформації враховує лише статистичні властивості повідомлень, які передаються.

Окрім статистичних властивостей повідомлень, які передаються, кожне повідомлення має свій зміст (семантичний зміст), який визначає цінність того чи іншого повідомлення для людини. Наведена вище оцінка кількості інформації не

враховує цей бік повідомлень, тобто не враховує суб'єктивну оцінку інформації людиною.

6.2. Ентропія джерела дискретних повідомлень

В багатьох випадках, коли потрібно узгодити канал з джерелом повідомлення, знання про кількість інформації, яка міститься в окремих повідомленнях, виявляється недостатньою. Виникає необхідність в характеристиках, які б дозволяли оцінювати інформаційні властивості джерела повідомлень в цілому. Однією з важливих характеристик такого роду є середня кількість інформації, що приходить на одне повідомлення.

В найпростішому випадку, коли всі повідомлення рівноймовірні $P(a_i) = 1/m$, середня кількість інформації дорівнює $\log m$. Тому, при рівноймовірностних незалежних повідомленнях інформаційні властивості джерела залежать тільки від числа повідомлень в ансамблі m .

Однак в реальних умовах повідомлення, за звичай, мають різну ймовірність. Так, наприклад, одні літери алфавіту зустрічаються частіше – інші рідше.

Тому знання числа повідомлень m в ансамблі є недостатнім. Необхідно мати дані про ймовірності кожного повідомлення:

$$P(a_1), \dots, P(a_m).$$

Так як ймовірності повідомлень не однакові, то вони несуть різну кількість інформації.

Менш ймовірні повідомлення несуть більшу кількість інформації і навпаки.

Середня кількість інформації, що приходить на одне повідомлення, визначається як математичне очікування

$$H(a) = \overline{I(a_i)} = \sum_{i=1}^m P(a_i) I(a_i) = - \sum_{i=1}^m P(a_i) \log P(a_i) \text{ біт/повідомлення} \quad (6.1)$$

Величина $H(a)$ називається ентропією. Цей термін позичений з термодинаміки, де мається аналогічний по своїй формі вираз, який характеризує невизначеність стану фізичної системи.

В теорії інформації ентропія $H(a)$ також характеризує невизначеність ситуації до передачі повідомлення, оскільки наперед не відомо, яке з повідомлень ансамбля джерела буде передано. Чим більша ентропія, тим сильніша невизначеність і тим більшу інформацію в середньому несе одне повідомлення джерела.

6.3. Ентропія неперервних повідомлень

При передаванні неперервних повідомлень передані сигнали $S(t)$ є функціями часу, які належать деякій множині, а отримані сигнали $x(t)$ будуть їх реалізаціями.

Усі дійсні сигнали мають спектри, загальна енергія яких зосереджена в обмеженій смузі частот F – згідно теоремі В.А.Котельникова такі сигнали визначаються своїми значеннями в точках відліку, які вибирають через інтервали

$$\Delta t = \frac{1}{2F}.$$

В каналі на сигнал накладається завади, в наслідок чого кількість різних рівнів сигналу в точках відліку буде скінченною. Отже, сукупність значень неперервного сигналу еквівалентна деякій дискретній скінченній сукупності. Це дозволяє визначити необхідну кількість інформації і пропускну спроможність при передаванні

неперервних повідомлень, що базуються на результатах, які отримано для дискретних повідомлень.

Визначимо спочатку кількість інформації, яке міститься у одному відліку сигналу P_{x_i} відносно переданого сигналу S_i :

$$I(s_i, x_j) = \lim_{\Delta s \rightarrow 0} \log \frac{p(s_i / x_j)}{p(s_i) \Delta s} = \log \frac{p(s_i / x_j)}{p(s_i)},$$

де $p(s_i), p(s_i / x_j)$ – щільність імовірності.

Середня кількість інформації

$$I(s, x) = \int \int_{S, X} p(s, x) I(s, x) dS dx = \int \int_{S, X} p(s, x) \log \frac{p(s/x)}{p(s)} dx ds,$$

де $p(s, x)$ – сумісна щільність імовірності; S, X – області можливих значень S_i та x_j .

Вираз для $\overline{I(s, x)}$ можна представити в двох виглядах

$$\overline{I(s, x)} = H(s) - H(s/x);$$

$$I(s, x) = H(x) - H(x/s),$$

де $H(s) = - \int_S p(s) \log p(s) ds$ – диференціальна ентропія сигналу S ;

$H(x) = - \int_x p(x) \log p(x) dx$ – диференціальна ентропія сигналу X ;

$H(s/x) = - \int_S \int_x p(x, s) \log p(s/x) ds dx$ – умовна диференціальна ентропія сигналу S ;

$H(x/s) = - \int_S \int_x p(s, x) \log p(x/s) ds dx$ – умовна диференціальна ентропія сигналу або ентропія шуму.

Багато властивостей ентропії неперервного розподілення аналогічні властивостям ентропії дискретного розподілення.

Якщо неперервний сигнал S обмежено інтервалом $S_1 \leq S < S_2$, то ентропія $H(s)$ максимальна і дорівнює $\log(s_2 - s_1)$, коли сигнал S має рівномірний розподіл

$$p(s) = 1/(s_2 - s_1).$$

Якщо обмежено середньоквадратичне значення неперервного сигналу

$$\int_{-\infty}^{\infty} s^2 P(s) ds = \sigma^2,$$

то ентропія максимальна при нормальному розподіленні щільності ймовірності

$$p(s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{s^2}{2\sigma^2}};$$

$$H(s)_{\max} = \log \sqrt{2\pi e} \sigma,$$

де $e = 2,71$.

Величина диференціальної ентропії (на відміну від ентропії дискретних сигналів) залежить від розмірності неперервного сигналу. Завдяки чому вона не є мірою кількості інформації, хоч і характеризує ступінь невизначеності, яка властива джерелу.

Тільки різниця ентропій кількісно визначає середню інформацію. $P(u_i/v_j) < 1$.

Кількість інформації, яка необхідна для усунення невизначеності, яка залишилася, дорівнює тій частині інформації, яку знищено в результаті дії завад. Кількість отриманої інформації визначається як різниця

$$I(u_i, v_j) = \log \frac{1}{P(u_i)} - \log \frac{1}{P(u_i/v_j)} = \log \frac{P(u_i/v_j)}{P(u_i)}.$$

Для оцінки середньої кількості отриманої інформації при передаванні одного повідомлення значення $I(u_i, v_j)$ необхідно усереднити по всьому ансамблю $\mathbf{u}$ та

$\mathbf{v}$

$$I(u, v) = \bar{I}(u, v) = \sum_{i=1}^{m_u} \sum_{j=1}^{m_v} P(u_i, v_j) \log \frac{P(u_i/v_j)}{P(u_i)} = \sum_{j=1}^{m_v} P(v_j) \sum_{i=1}^{m_u} P(u_i/v_j) \log \frac{P(u_i/v_j)}{P(u_i)},$$

де $P(u_i, v_j) = P(v_j)P(u_i/v_j) = P(u_i)P(v_j/u_i)$ – суцільна ймовірність переданого та отриманого сигналів;

m_u – кількість сигналів u в ансамблі на вході; m_v – кількість сигналів в ансамблі на виході каналу.

У загальному випадку $m_v \neq m_u$.

Величина $I(u, v)$ характеризує у середньому кількість інформації, яка містить отриманий сигнал v відносно переданого сигналу u і тому $I(u, v)$ називають також середньою взаємною інформацією між u та v .

Зазвичай вираз для $I(u, v)$ надається у вигляді формули

$$I(u, v) = H(u) - H(u/v).$$

6.4. Ефективність систем передавання інформації

Найбільш загальна оцінка – коефіцієнт використання каналу

$$\eta = R/C.$$

Для дискретних систем

$$\eta = \eta_1 \eta_2,$$

де η_1 – ефективність системи кодування;

η_2 – ефективність системи модуляції. В

Вводячи надлишковість повідомлення $\chi_1 = 1 - \eta_1$ і надлишковість сигналу

$$\chi_2 = 1 - \eta_2, \text{ отримаємо}$$

$$\eta = 1 - \chi,$$

де $\chi = \chi_1 + \chi_2 - \chi_1 \chi_2$ – повна надлишковість системи.

При передаванні неперервних повідомлень

$$\eta = \frac{F_m \log(P_c^* / P_n^* + 1)}{F \log(P_c / P_n + 1)}.$$

Коефіцієнт використання потужності сигналу

$$\eta_p = R / (P_c / N_c),$$

де N_0 – інтенсивність завади.

Коефіцієнт використання смуги частот каналу

$$\eta_n = \frac{R}{F}.$$

6.5. Ентропія розподілу ймовірностей керованого процесу

Сучасний стан інформаційних мереж дозволяє виконати впровадження глобальної інформаційної інфраструктури (ГІ), яка надає можливість користувачам отримати усі види комунікаційних послуг з високою якістю, належною вартістю у будь-якому місті та у будь-який час. Щоб забезпечити такі вимоги необхідна більш гнучка архітектура системи управління, яка легко могла б підтримати швидке введення в дію нових послуг та їх підтримку з належною імовірністю на всій глобальній мережі.

Системам управління властиві такі ознаки:

- велика складність;
- розвинені функціональні можливості;
- жорсткі вимоги до захисту інформації;
- висока надійність та точність;
- розосередженість на великій території.

За прогнозом, кількість управляючої інформації в таких системах буде різко зростати з наданням послуг. Внаслідок чого система управління може поглинути основну мережу. Одним з головних завдань для системи управління такими мережами є визначення мінімально необхідної кількості управляючої інформації, яка повинна повністю забезпечити управляючу мережу з належною точністю параметрів. Методики розрахунку пропускної спроможності каналів системи управління за обсягом управляючої інформації для телекомунікаційних мереж немає. На експлуатації користуються тільки практичними методами, які дозволяють в окремих випадках зробити наближений розрахунок тих чи інших параметрів мережі.

В дисертаційній роботі розроблено інформаційно-ентропійний підхід до процесу управління мережею.

У розділі запропоновано інформаційно-ентропійний метод розрахунку кількості інформації, а як наслідок, і необхідної пропускної спроможності каналів системи управління.

Представлено комбіновану структуру системи управління (СУ), яка забезпечує її функціонування одночасно в двох режимах. Перший режим контролює параметри інформаційної мережі і на основі цього СУ приймає ефективне рішення. Другий – дозволяє попередити аварійні ситуації шляхом проведення СУ аналізу збурюючих факторів та коректировки параметрів мережі.

Розроблений інформаційно-ентропійний метод дозволяє знайти необхідний мінімум управляючої інформації, при якому параметри мережі будуть забезпечувати задану точність, володіти як якістю адаптивності до плинних спрогнозованих збурювань, так і інваріантністю до випадкових збурюючих факторів.

Розглянемо спосіб запропонованого інформаційно-ентропійного методу розрахунку.

Нехай процес функціонування мережі в кожний момент часу t характеризується вектором змінних станів $\{x_1(t), x_2(t), \dots, x_n(t)\}$ або функціоналом $\Phi\{x_1(t), x_2(t), \dots, x_n(t)\}$. Вказані змінні станів є випадковими величинами зі своїми законами розподілу $p(x_i)$. Управління мережею - процес приведення змінних стану мережі за заданий час з початкового стану до оптимального. Наприклад, x_i - час доставки інформації між двома вузлами мережі. Якість функціонування мережі характеризується затримкою повідомлень, яка обчислюється величиною затримки на всіх вузлах. Завдання управління мережею має на меті мінімізацію і підтримку на рівні, не більше за задане середнє очікування і зменшення дисперсії затримки повідомлень в мережі.

При цьому вимоги до точності управління мережею часто поєднуються з вимогами збільшення швидкості протікання процесу. У термінах теорії інформації є визначення: процес управління мережею є процес зменшення невизначеності стану мережі, що може бути чисельно представлено, як зміна ентропії мережі в процесі управління, оскільки ентропія визначає відхилення параметрів мережі за даний момент часу.

Відомо, що середня кількість інформації при перенесенні сигналу дорівнює різниці ентропії розподілу імовірностей цієї величини до і після отримання сигналу. Таким чином, визначивши зміну ентропії керованого процесу функціонування мережі в процесі управління, можна визначити кількість інформації, необхідну для отримання заданого відхилення ймовірності змінних стану мережі.

Вичерпною характеристикою координат процесу кожного заданого моменту часу - є розподіл імовірностей цих величин.

Для безпосередніх координат, що приймають довільні значення в деяких інтервалах, будемо розглядати щільність імовірності:

$$p(x_1, x_2, \dots, x_n). \quad (6.2)$$

Кожну сукупність значень параметрів інформаційної мережі можна розглядати, як певний стан мережі.

У теорії ймовірності і теорії інформації поняття ентропії визначається розподілом ймовірностей будь-яких змінних. Ентропією безперервного розподілу ймовірностей змінних $x_1, x_2, \dots, x_n$ прийнято величину -

$$H = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, x_2, \dots, x_n) \cdot \log p(x_1, x_2, \dots, x_n) dx_1, \dots, dx_n \quad (6.3)$$

Теорія інформації розглядає ентропію розподілу ймовірностей в системах отримання і передачі інформації. При цьому вважається, що після отримання інформації про яку-небудь величину, розподіл імовірності цієї величини і, отже, ентропія можуть істотним чином змінитися.

Тим самим поняттю ентропії в теорії інформації приписується значення «суб'єктивної», або відносної характеристики. Подібну ентропію будемо називати *інформаційною ентропією*. Крім інформаційної ентропії, для опису управління взаємодії елементів мережі, доцільно використати інше поняття ентропії, яке є, нібито додатком поняття фізичної ентропії до керованих процесів, що характеризуються не координатами безлічі елементів мережі, а обмеженим числом різних величин - координат. Цей вид ентропії будемо називати ентропією *розподілу* ймовірностей координат керованого процесу або просто *ентропією процесу*. Вказані поняття ентропії мають однакові математичні вирази.

Ентропія розподілу ймовірностей координат керованого процесу виражається тією ж формулою, що і інформаційна ентропія:

$$H = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, x_2, \dots, x_n) \cdot \log p(x_1, x_2, \dots, x_n) dx_1, \dots, dx_n.$$

Відмінність має місце в сутності вказаних понять і полягає в наступному. Інформаційна ентропія, як вище вказувалося, істотним чином залежить від інформації, що отримується даним спостерігачем, і є, внаслідок цього, відносною «суб'єктивною» характеристикою. Ентропія процесу і розподіл ймовірностей координат цього процесу - об'єктивні характеристики, які не залежать від інформації, отриманої окремим спостерігачем.

Іншими словами, вважається, що ентропію розподілу ймовірностей координат будь-якого процесу, не можна змінити, вимірявши з тією чи іншою точністю ці координати (параметри) інформаційної мережі (ІМ). Для зміни ентропії процесу необхідне втручання в цей процес, тобто управління або «природний» вплив. Фізична ентропія є окремий випадок ентропії процесу, а саме той випадок, коли координатами процесу є параметри інформаційної мережі.

Отже, є два загальних поняття ентропії, що мають однаковий математичний вираз: ентропія процесу і інформаційна ентропія.

Розглянемо деякі властивості ентропії. Оскільки математичні визначення інформаційної ентропії і ентропії процесу однакові, всі співвідношення в рівній мірі слушні для обох видів ентропії.

Якщо координати $x_1, x_2, \dots, x_n$ незалежні, то:

$$p(x_1, x_2, \dots, x_n) = p(x_1) p(x_2) \dots p(x_n) \quad i,$$

$$\log p(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \log p(x_i),$$

де $p(x_1), p(x_2), \dots, p(x_n)$ - щільності розподілу ймовірностей величин $x_1, x_2, \dots, x_n$.

Таким чином, для незалежних величин:

$$\begin{aligned}
H &= - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1)p(x_2)\dots p(x_n) \cdot \left(\sum_{i=1}^n \log p(x_i) \right) dx_1, dx_2, \dots, dx_n = \\
&= - \int_{-\infty}^{\infty} p(x_1) \log p(x_1) dx_1 \cdot \int_{-\infty}^{\infty} p(x_2) dx_2 \dots \int_{-\infty}^{\infty} p(x_n) dx_n - \\
&- \int_{-\infty}^{\infty} p(x_1) dx_1 \int_{-\infty}^{\infty} p(x_2) \log p(x_2) dx_2 \int_{-\infty}^{\infty} p(x_3) dx_3 \dots \int_{-\infty}^{\infty} p(x_n) dx_n - \\
&- \int_{-\infty}^{\infty} p(x_1) dx_1 \int_{-\infty}^{\infty} p(x_2) dx_2 \dots \int_{-\infty}^{\infty} p(x_{n-1}) dx_{n-1} \int_{-\infty}^{\infty} p(x_n) \log x_n dx_n.
\end{aligned}$$

Але за визначенням щільності ймовірності

$$\int_{-\infty}^{\infty} p(x_i) dx_i = 1.$$

Тому

$$H = \sum_{i=1}^n H_i, \quad (6.4)$$

де H_i - ентропія i -ї з незалежних величин.

Отже, ентропія незалежних параметрів процесу дорівнює сумі ентропії кожної координати.

Якщо процес характеризується лише одним параметром x , то ентропія:

$$H = - \int_{-\infty}^{\infty} p(x) \cdot \log p(x) dx. \quad (6.5)$$

Якщо параметр x може приймати лише дискретні значення, імовірності яких дорівнюють $P_1, P_2, \dots, P_m$, то ентропія такого параметру виражається формулою:

$$H = - \sum_{j=1}^m P_j \log P_j. \quad (6.6)$$

Ця формула може бути отримана простим граничним переходом з (6.5).

Якщо дискретна величина не має відхилення, тобто значення її строго фіксоване, то всі імовірності P_j , крім однієї, рівні нулю, а ця єдина імовірність дорівнює 1. Таким чином, ентропія дискретної величини при відсутності відхилення дорівнює нулю. Це безпосередньо витікає з (6.6), де $0 \leq P_j \leq 1$.

Ентропія величини, яка може приймати m значень, як відомо, максимальна, якщо всі ці значення незалежні і рівноімовірні.

Ентропія безперервної величини заданої абсолютно точно, дорівнює негативній нескінченності.

Це положення легко довести, здійснюючи граничний перехід від розподілу ймовірностей з кінцевою дисперсією до випадку, коли дисперсія дорівнює нулю. Щільність імовірності при такому переході перетворюється в δ -функцію і ентропія прагне до $-\infty$.

$$H = - \int_{-\infty}^{\infty} \delta(x) \cdot \log \delta(x) dx = -\infty.$$

Слід відзначити, що зростання абсолютної величини ентропії, зі зменшенням дисперсії (збільшенням точності), відбувається в логарифмічній пропорції, тобто «пливно». Перетворення в нескінченність $H(x)$, при точному завданні безперервної величини x , не можна вважати недоліком поняття ентропії, як характеристики невизначеності. Відзначене положення вказує лише на недоцільність використання

поняття ентропії в тих випадках, коли застосовується ідеалізація (модель) процесу з точним завданням безперервних величин.

У реальних процесах фізичні величини ніколи не можуть бути виміряні і задані абсолютно точно, і ентропія ніколи не перетворюється в нескінченність.

Ентропія безперервної величини, маючи заданий середній квадрат, максимальна при нормальному законі розподілу ймовірності.

Для підтвердження цього положення знайдемо екстремум ентропії.

$$H = - \int_{-\infty}^{\infty} p(x) \cdot \log p(x) dx, \quad (6.7)$$

при умовах

$$\left. \begin{aligned} \int_{-\infty}^{\infty} p(x) dx &= 1; \\ \int_{-\infty}^{\infty} x^2 p(x) dx &= \bar{x}^2 = \sigma^2. \end{aligned} \right\} \quad (6.8)$$

Умова екстремуму функціонала:

$$H + \lambda_1 \int_{-\infty}^{\infty} p(x) dx + \lambda_2 \int_{-\infty}^{\infty} x^2 p(x) dx = \int_{-\infty}^{\infty} (-p(x) \log p(x) + \lambda_1 p(x) + \lambda_2 x^2 p(x)) dx,$$

полягає в рівності

$$\frac{d}{dp(x)}(-p(x)\log p(x) + \lambda_1 p(x) + \lambda_2 x^2 p(x)) = -\log p(x) + \lambda_1 + \lambda_2 x^2 = 0.$$

Звідки:

$$p(x) = C e^{\lambda_2 x^2},$$

де $C = \text{const}$; $\lambda = \text{const}$.

Визначаючи C і λ_2 з умови (6.8), отримусмо формулу нормального розподілу ймовірностей (розподіл Гаусса):

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

Легко перевірити, що нормальний розподіл при заданих умовах відповідає саме максимуму ентропії і цей максимум дорівнює:

$$H = \log(\sigma \sqrt{2\pi e}) = \log(4,13\sigma). \quad (6.9)$$

Таким чином, при заданому середньому квадраті нормальний розподіл ймовірності відповідає найбільшій невизначеності величини x з всіх можливих розподілів.

Тут необхідно указати на важливе положення: практичне значення має не абсолютна величина ентропії, а її приріст.

Це положення справедливе, мабуть, як твердження, що значення має не абсолютна величина потенціалу, а різниця потенціалів. Воно знімає ускладнення, пов'язане з розмірністю величин. Дійсно, якщо x - розмірна фізична величина, то щільність ймовірностей $p(x)$ має розмірність зворотну розмірності x .

Отже, під знаком логарифма з'являється розмірна величина. Проте у виразу різниці ентропій завжди присутній логарифм відношення і фізична розмірність скорочується. Це наочно демонструє формула (6.9). Різниця ентропій двох станів, що мають нормальні розподіли і середні квадрати σ_1^2, σ_2^2 , дорівнює:

$$H_2 - H_1 = \log(\sigma_2 \sqrt{2\pi e}) - \log(\sigma_1 \sqrt{2\pi e}) = \log \frac{\sigma_2}{\sigma_1} = -\log \frac{\sigma_1}{\sigma_2}. \quad (6.10)$$

Використання поняття ентропії, зрозуміло, не замінює детального вивчення процесу управління взаємодією елементів інформаційної мережі. Однак розгляд ентропії дозволяє встановити деякі загальні умови організації управління процесами.

6.6. Ентропія в замкнених системах управління

Нехай стан процесу (функціонування мережі) в кожному даний момент часу характеризується сукупністю станів $x_1, x_2, \dots, x_n$, які назвемо координатами процесу. Координати процесу завжди мають відхилення відносно бажаних значень, тобто містять випадкові складові.

Одним з основних завдань управління функціонуванням IM є зменшення відхилень процесу від бажаного. При цьому, вимога до точності управління підвищується.

Необхідно зазначити, що в літературі з теорії інформації, часто отримання інформації про який-небудь процес подібне до поняття про зміну ентропії самого цього процесу. Нагадаємо, що поняття приросту ентропії стану процесу є величина відхилення відносно деякого бажаного проходження процесу. Отже, зміна ентропії стану процесу є об'єктивною характеристикою. Ця зміна не може бути досягнута тільки шляхом отримання інформації, тобто «відображенням» або «зображенням» процесу. Для зміни ентропії, у вказаному її розумінні, необхідна організована зміна процесу, тобто управління.

Отримання певної кількості інформації в процесі необхідне, але ще недостатнє для здійснення управління.

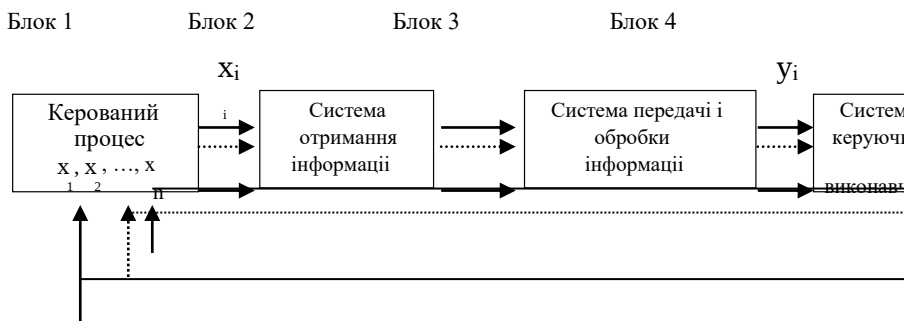


Рис. 6.1 Замкнена схема процесу управління

На рис. 6.1. зображена замкнута схема процесу управління. Інформація про керований процес (блок 1), (координати X_i) приймається за допомогою системи отримання інформації (блок 2). Отримана інформація обробляється в системі передачі і обробки інформації (блок 3). Оброблена інформація поступає в систему формування управляючих впливів Z_i - виконавчу систему (блок 4).

Управляючі впливи X_i впливають на керований процес в напрямку зменшення ентропії процесу.

При розгляді передачі і перетворення інформації величини $x_1, x_2, \dots, x_n$ потрібно трактувати як вхідні величини системи отримання і перетворення інформації, а $y_1, y_2, \dots, y_n$ - як вихідні величини цієї системи. Отже, кількість переданої і перетвореної інформації виражається логарифмом відношення щільності імовірності вхідних величин після отримання сигналу на виході.

Кількість інформації залежить від вхідних і вихідних величин. Для загального опису процесу передачі і перетворення інформації зручно розглядати математичне очікування або середню кількість інформації при передачі однієї величини ($x \rightarrow y$) і виражається формулою:

$$I = M[H(x)] - M_y[H_y(x)]. \quad (6.11)$$

Величина $M_y[H_y(x)]$ є усередненою характеристикою ефекту дії завад, а також запізнення при передаванні сигналу.

Отже, середня кількість інформації при передачі $x \rightarrow y$ дорівнює різниці ентропії розподілу ймовірностей вхідної величини до отримання сигналу і усередненої ентропії розподілу ймовірностей цієї величини після отримання сигналу.

Слід зауважити, що у разі нормального розподілу ймовірностей вхідної і вихідної величин середня кількість інформації при передаванні $x \rightarrow y$ можемо виразити через коефіцієнт кореляції z вхідної і вихідної величин.

Зазначимо, що величини $x_1, x_2, \dots, x_n$ так само як і величини $y_1, y_2, \dots, y_n$ розглядаємо як компоненти n -мірних векторів і передавання $x_1, x_2, \dots, x_n \rightarrow y_1, y_2, \dots, y_n$ можна називати передаванням n -ого вектора $x \rightarrow y$.

Раніше, ніж записати відповідні вирази, пояснимо поняття кореляційної матриці і кореляційного визначника випадкового вектора. Кореляційні моменти є математичні очікування відхилень випадкових величин:

$$\begin{aligned} R_{\gamma\mu}^{xx} &= M[(x_\gamma - m_{x\gamma})(x_\mu - m_{x\mu})]; \\ R_{\gamma\mu}^{xy} &= M[(x_\gamma - m_{x\gamma})(y_\mu - m_{y\mu})]; \\ R_{\gamma\mu}^{yy} &= M[(y_\gamma - m_{y\gamma})(y_\mu - m_{y\mu})]. \end{aligned} \quad (6.12)$$

Матриці, складені з кореляційних моментів, називаються кореляційними матрицями, а визначники, відповідні цим матрицям, іменуються кореляційними визначниками.

Можна показати, що ентропія випадкового вектора x , яка має нормальний закон розподілу, дорівнює:

$$H(x) = \log \sqrt{(2\pi e)^n |R^{xx}|}, \quad (6.13)$$

де $|R^{xx}|$ - кореляційний визначник даного вектора.

Математичне очікування кількості інформації при передачі вектора $x \rightarrow y$ при нормальному розподілі вхідних і вихідних величин:

$$I = -\frac{1}{2} \log \frac{|R|}{|R^{xx}| \cdot |R^{yy}|}. \quad (6.14)$$

Якщо вихідний вектор не залежить від вхідного, то $R_{\mu\mu}^{xy} = 0$, $|R| = |R^{xx}| \cdot |R^{yy}|$, і кількість інформації дорівнює нулю: $I=0$. Для одновимірного випадку:

$$|R^{xx}| = R_{11}^{xx} = \sigma_x^2; \quad |R^{yy}| = R_{11}^{yy} = \sigma_y^2;$$

$$|R| = \begin{vmatrix} R_{11}^{xx} & R_{11}^{xy} \\ R_{11}^{xy} & R_{11}^{yy} \end{vmatrix} = R_{11}^{xx} R_{11}^{yy} - (R_{11}^{xy})^2$$

і вираз (6.7) можна представити у вигляді:

$$I = -\frac{1}{2} \log(1 - r^2),$$

$$\text{де } r = \frac{R_{11}^{xy}}{\sigma_x \sigma_y}.$$

При розгляді понять передачі інформації в системах управління, необхідно проаналізувати інформаційні процеси при перетворенні координат і втрати інформації при проходженні сигналу через послідовно ввімкнені системи.

У системах отримання передачі і обробки інформації (рис. 6.1) часто здійснюється перетворення координат. Перетворення координат:

$$y_1 = f_1(x_1, x_2, \dots, x_n),$$

$$y_2 = f_2(x_1, x_2, \dots, x_n),$$

.....

$$y_n = f_n(x_1, x_2, \dots, x_n)$$

називається взаємно однозначним, якщо перетворення

$$\frac{D(y_1, y_2, \dots, y_n)}{D(x_1, x_2, \dots, x_n)} = \begin{vmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} & \dots & \frac{df_1}{dx_n} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} & \dots & \frac{df_2}{dx_n} \\ \dots & \dots & \dots & \dots \\ \frac{df_n}{dx_1} & \frac{df_n}{dx_2} & \dots & \frac{df_n}{dx_n} \end{vmatrix}$$

відмінне від нуля. Тут $f_1, f_2, \dots, f_n$ - деякі функції, що диференціюються.

Зазначимо, що середня кількість інформації, яку передають, не змінюється при взаємно однозначному перетворенні координат.

Інформація проходить послідовно від системи отримання інформації, до системи передачі, обробки інформації і виконавчу систему. Система передачі, обробки інформації і виконавча система не мають безпосереднього впливу з боку керованого процесу, бо передбачається лише однонаправлена дія управляючих координат z_i на керований процес.

Якщо ж керований процес надає безпосередній зворотний вплив на управляючі координати (виконавчу систему), то його можна привести до випадку однонаправленої дії додаткової інформації в систему отримання інформації.

Питання ставиться так: чи може кількість інформації при передаванні $x_i \rightarrow z_i$ перевищувати кількість інформації при передаванні $x_i \rightarrow y_i$. Негативна відповідь на це запитання приводиться без доказу, в зв'язку з достатньою ймовірністю.

Середня кількість інформації при проходженні послідовно сполучених систем може тільки убавати або (у граничному разі відсутності додаткових завад) залишатися незмінною:

$$I_{xz} \leq I_{xy},$$

де: $I_{xz} = M[H(x)] - M_z[H_z(x)]$ - середня кількість інформації при передачі $x \rightarrow y \rightarrow z$;

$I_{xy} = M[H(x)] - M_y[H_y(x)]$ - середня кількість інформації при передачі $x \rightarrow y$.

Звичайно, в теорії інформації розглядається передача інформації через розімкнені системи. Для процесів управління характерні замкнені контури циркуляції інформації.

Розглянемо загальну схему процесу управління, представлену на рис.6.1.

Відзначимо, що модель системи такого типу є адаптивною і відповідно, практично обробляє усі плинні прогнозовані збурення, які впливають на виконавчу систему.

Контур циркуляції в даній системі можна подати у вигляді:

$$x \rightarrow y \rightarrow z \rightarrow x \text{ або } x \rightarrow z \rightarrow x.$$

Для отримання загальних співвідношень, справедливих для будь-яких ланцюгів замкнутого контура (в які можуть входити і люди-оператори), необхідно використовувати досить загальне поняття передачі інформації. Розімкнемо контур в якій-небудь точці, наприклад, в точці з'єднання вимірювальної системи (системи отримання інформації) з керованим процесом (об'єктом). Іншими словами, припинимо надходження інформації від керованого процесу до вимірювальної системи. Подамо на вхід вимірювальної системи інформацію, тобто деяке значення x від стороннього джерела. Тоді після проходження сигналу через систему управління і керований об'єкт

на виході утвориться значення x' . Таким чином, в розімкненій системі має місце передача інформації типу $x \rightarrow x'$.

Як відомо з основних співвідношень теорії інформації, кількість інформації при передачі $x \rightarrow x'$ дорівнює:

$$I = M[H(x)] - M_x[H_x(x')], \quad (6.15)$$

де $H(x)$ - ентропія x , $M_x[H_x(x')]$ - усереднене за x значення ентропії після отримання сигналу x' .

Для скорочення запису використані позначення:

$H(x_1, x_2, \dots, x_n) = H(x)$, де $x_1, x_2, \dots, x_n$ - координати (параметри) керованого процесу.

Співвідношення (6.15) можна представити у вигляді

$$I = H(x) - H_x(x') + S, \quad (6.16)$$

де $S = M[H(x)] - M_x[H_x(x')]$.

Величина S називається приростом ентропії приведених шумів. Зміст цієї величини наступний. Величина $H_x(x')$ є ентропією координат на виході (ентропія процесу) при заданому сигналі x на вході. Ця ентропія має кінцеву величину внаслідок впливу різного роду похибок і завад, причому вона відображає сумарний ефект дії похибок і завад вимірювачів, системи передачі і перетворення інформації, виконавчої системи, а також збурюючих сил безпосередньо на керований процес. Дія цих завад, ніби приведена до виходу контуру - керованим координатам x' . Член $M_x[H_x(x')]$ – це та ж величина, усереднена за вхідною величиною x .

Таким чином, приріст ентропії приведених завад S є різниця ентропії завад, приведених до виходу, і тієї ж ентропії, усередненої за вхідними величинами.

Якщо замкнути контур дискретного управління, то на вхід сприймаючої системи будуть впливати координати керованого процесу в деякий момент часу $t-1$, а величиною x' будуть координати того ж керованого процесу в момент часу t , то величини x та x' будуть мати:

$$x = x(t-1); x' = x(t).$$

Величини $H(x')$ і $H_x(x')$ при замкненому контурі дискретного управління звертаються до значень ентропії керованого процесу відповідно в $(t-1)$ -й і t -й моменти часу.

Рівняння (6.16) приймає вигляд:

$$I(t) = H(t-1) - H(t) + S(t),$$

або

$$H(t) - H(t-1) = S(t) - I(t), \quad (6.17)$$

де $H(t-1)$ і $H(t)$ - значення ентропії керованого процесу відповідно в $(t-1)$ -й і t -й моменти часу;

$I(t)$ - кількість інформації, передана за відповідний інтервал часу від джерела до одержувача (яким є той же керований процес).

Отже, приріст ентропії, керованого дискретним чином процесу, за інтервал дискретності дорівнює різниці приросту ентропії приведених завод і кількості інформації, переданої в замкненому контурі ($x(t-1) \rightarrow x(t)$) за той же інтервал часу. Позначимо індексом 0 початковий момент часу і з співвідношення (6.14), знаходимо:

$$H(t) = H(0) - \sum_{i=0}^t I(i) + \sum_{i=0}^t S(i). \quad (6.18)$$

Ентропія керованого дискретним чином процесу дорівнює різниці початкової ентропії цього процесу, складеної з сумарним приростом ентропії завад, і сумарної кількості інформації, переданої через контур $(x(t-1) \rightarrow x(t))$.

Для стаціонарного процесу $H(t) = H(t-1)$ і $S(t) = I(t)$.

Тому, для підтримки стаціонарного керованого процесу необхідно передавати через контур управління протягом інтервалу дискретності кількість інформації, яка дорівнює відповідному приросту ентропії приведених завад. Зазначимо, що для некерованого процесу кількість інформації $I(t)$ дорівнює нулю, $H(t) - H(t-1) = S(t)$, тобто приріст ентропії приведених завад дорівнює приросту ентропії некерованого процесу за відповідний інтервал часу.

Отже, для управління процесом функціонування мережі необхідно передавати протягом інтервалу дискретності управління, управляючу інформацію, обсяг якої визначається як сума необхідного зменшення ентропії процесу функціонування мережі за цей інтервал часу і сумарного приросту ентропії некерованого процесу за цей же інтервал часу.

У загальному випадку, якщо процес характеризується вектором $\{X_1(t), X_2(t), \dots, X_n(t)\}$, то ентропія дорівнює:

$$H = - \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p(x_1, x_2, \dots, x_n) \cdot \log p(x_1, x_2, \dots, x_n) dx_1, \dots, dx_n.$$

Оскільки важко визначити заздалегідь реальний закон розподілу $p(x)$, зробимо припущення, що $p(x)$ має нормальний закон розподілу. У цьому випадку ентропія, як відомо, визначається як:

$$H = \log(4,13\sigma).$$

Таким чином, приріст (зменшення) ентропії в (24) буде визначатися як:

$$H(t) - H(t-1) = \log \frac{\sigma_1(t)}{\sigma_1(t-1)} + \dots \log \frac{\sigma_n(t)}{\sigma_n(t-1)}, \quad (6.19)$$

і для оцінки кількості інформації, яка передається в системі управління мережею, ми повинні оцінити відношення середньоквадратичних відхилень параметрів повідомлень, що досліджуються в мережі в різних ситуаціях, а така оцінка може бути зроблена за вимогами до параметрів основної мережі.

6.7. Ентропія та інформація в розімкнених системах управління

Як відомо, складні системи схильні до дії незакономірних збурень, що різко змінюються. Тому система управління повинна підтримати вимогу до точності параметрів мережі, незважаючи на стохастичний характер зовнішніх збурень. Системи управління, що володіють такими властивостями, називаються інваріантними (тобто не змінюють свої параметри).

Таким чином, блок 1 (рис.6.2) відповідно відображає джерело стохастичних збурень, блоки 2, 3 і 4 - безпосередньо систему управління (отримання інформації, передачі і обробки інформації, виконавча система), а блок 5 - це керований процес, тобто керована інформаційна мережа. Скорочено цей контур може бути представлений у вигляді:

$$X \rightarrow Y \rightarrow Z.$$

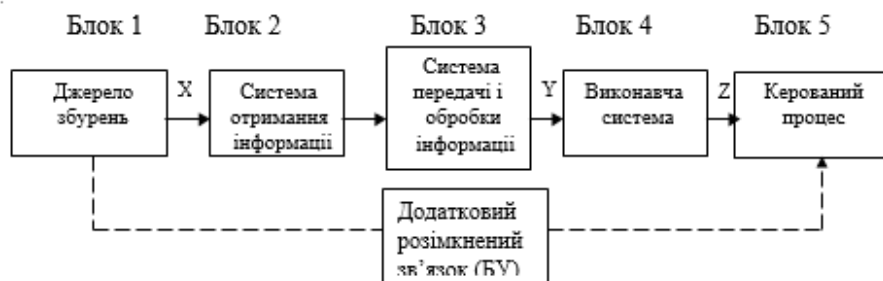


Рис. 6.2 Інваріантна система управління в розімкнутому вигляді

Рис. 6.2 Інваріантна система управління в розімкнутому вигляді

Як відомо, технічна реалізація абсолютно інваріантних систем складна. Їх потрібно розглядати як граничні характеристики, до яких прагнуть, але досягнути важко. Тому в ряді систем досягається інваріантність не повна, а часткова, або система інваріантна до ε , де під ε з розуміємо похибку системи. Системою інваріантною до ε назвемо таку систему, в якій зберігаються межі абсолютно інваріантної системи і існує динамічна похибка σ , що характеризує міру інваріантності. Користуючись основними визначеннями, згідно теорії інформації, можна указати таку абсолютно інваріантну систему, яка буде відрізнятися від вихідної наявністю динамічної похибки σ при максимально необхідній кількості інформації ΔH . Порівняння таких систем для однієї множини вхідних сигналів дозволить отримати оцінку інваріантності. При цьому ΔH і буде мірою ε -інваріантності для даного типу множини вхідних сигналів.

Використовуючи оцінки інваріантності до ε , отримаємо вираз, що визначає необхідну кількість інформації ΔH :

$$\Delta H = K \left(\log \frac{\sigma_1}{\sigma_1} + \log \frac{\sigma_2}{\sigma_2} + \dots + \log \frac{\sigma_n}{\sigma_n} \right), \quad (6.20)$$

де σ_i - можливі середньоквадратичні відхилення параметрів інформаційної мережі, що контролюються, від необхідних значень;

—
 σ_i - середньоквадратична допустима похибка параметрів інформаційної мережі, що контролюються, від необхідних значень; $i=1, 2, \dots, n$;

K - коефіцієнт, який залежить від типу каналу зв'язку, що використовується.

Необхідно зазначити, система управління інформаційною мережею повинна володіти як можливістю адаптації до стійких плинних змін режиму, так і властивістю інваріантності, що дозволяє системі бути нечутливою до випадкових збурень. Таким чином, система управління повинна мати комбіновану структуру, яка об'єднує описані вище адаптивну (замкнену) і інваріантну (розімкнену). Структура її представлена на рис. 6.3, де БУ – блок узгодження.

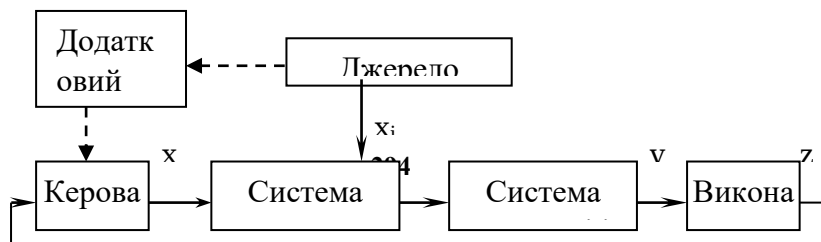


Рис. 6.3 Структура комбінованої системи управління

Таким чином, середня кількість інформації, яку необхідно передати в системі управління повинна дорівнювати сумі виразів (6.19) і (6.20). При цьому буде забезпечено як властивість адаптивності, так і інваріантності.

Описаний метод дозволяє знайти необхідний мінімум інформації, при якому параметри мережі матимуть задану точність. Якщо кількість інформації в системі управління буде меншою, то точність параметрів буде менше заданої.

Очевидно, з точки зору теорії інформації і динамічної точності, необхідно, щоб система управління забезпечувала задану точність параметрів мережі при мінімальному обсязі управляючої інформації.

Отже, з двох систем управління переваги будуть надані тій, яка вимагає меншу кількість управляючої інформації, це дозволяє знизити вимоги до якості каналів передавання даних, що використовуються системами управління.

Тому інформаційно-ентропійний метод дозволяє визначити нетрадиційним способом кількість керуючої інформації у системі управління будь-якої телекомунікаційної мережі. Таку оцінку за допомогою інших практичних методів майже неможливо отримати, так як обмежена кількість вихідних даних, особливо для мереж, на яких застосовуються різні технології. Практична цінність забезпечується можливістю врахування ємності управляючої інформації з переліком вимог до точності параметрів основної мережі. Така оцінка визначається деякою дисперсією відхилення від математичного очікування. Чим більші вимоги до точності параметрів управляючої мережі, тим більша потрібна ємність управляючої інформації.

Запропонована комбінована система управління (рис.6.3) дозволяє управляти телекомунікаційною мережею у двох режимах, одночасно забезпечуючи як властивості адаптивності, так і інваріантності параметрів мережі.

Зазначимо, що перший режим (назвемо його адаптивним) використовується у всіх системах управління. Він визначається тим, що виконує контроль параметрів телекомунікаційної мережі, на підставі якого система управління приймає рішення. Структура такої системи представлена на рис.6.1.

Другий режим (розімкнена система) дозволяє усунути аварійні ситуації, тому що аналізує спрогнозовані збурювальні фактори, враховуючи які, виконує коригування параметрів мережі. Структура такої системи представлена на рис.6.2.

І, нарешті, комбінована структура (рис.6.3) дозволяє реалізувати обидва режими. Кількість інформації, яка передається та обробляється системою управління з одного боку, визначає основні властивості керованої телекомунікаційної мережі, а з другого - зумовлює її структуру та основні властивості.

6.8. Розрахунок обсягу управляючої інформації для забезпечення необхідної точності параметрів мережі

Інформаційно-ентропійний метод полягає в тому, що кількість управляючої інформації визначається як різниця ентропій телекомунікаційної мережі перед ввімкненням СУ і після встановлення режиму нормальної роботи. При цьому враховується необхідна кількість управляючої інформації, яка необхідна для підтримки режиму нормальної роботи. Режим нормальної роботи мережі визначається середнє квадратичним відхиленням контрольованих параметрів від допустимих значень. Кількість каналів та їх параметри залежать від обсягу управляючої інформації.

Наведемо методику, яка реалізує інформаційно-ентропійний метод і дозволяє визначити необхідні параметри каналів зв'язку системи управління, забезпечує необхідну точність управління характеристиками основної телекомунікаційної мережі.

Як зазначено вище, одним з основних завдань проектування СУ є визначення мінімально необхідної кількості управляючої інформації, що забезпечує необхідну точність на керуючій мережі. У термінах теорії інформації «управління мережею» — це процес зменшення невизначенності її стану, що чисельно характеризує зміни відповідної ентропії, оскільки ентропія визначає відхилення параметрів мережі у кожному фіксованому моменті часу.

Вихідними даними для розрахунку є відомості про:

- об'єкт управління;
- параметри об'єктів управління;
- відхилення параметрів вказаних об'єктів;

У кожному конкретному випадку при виборі об'єктів управління та їх параметрів виходять з необхідної точності розрахунків та потреб системи управління.

Завдання розрахунку обсягу управляючої інформації доцільно реалізувати в інтерактивному режимі, завдяки якому користувач в зручній для себе формі вибирає об'єкти та параметри управління.

Для розрахунку обсягу управляючої інформації за вказаною методикою були розроблені алгоритми та програмне забезпечення. Як приклад спеціалізованого інструментального програмного забезпечення використовувалась система управління базами даних FoxPro for WINDOWS, заснована на dbf-файлах, що успішно об'єднує в собі простоту з великим набором сервісних послуг.

Програмне забезпечення, побудовано за модульним принципом, складається з таких процедур:

- головної, котра реалізує ініціалізацію змінних, підготовку файлів бази даних та виклик процедур меню;
- процедур меню, що забезпечують інтерактивну взаємодію користувача з даними (викликається необхідна процедура обробки, перегляду та корекції);
- підрахунку кількості інформації та числа каналів для системи управління;
- вивід на екран таблиць з початковими та вихідними даними;
- вивід на друкуючий пристрій (принтер) таблиць з початковими та вихідними даними;

Алгоритм головної процедури демонструється на рис.6.4.

Розрахунок за розробленою методикою виконано на прикладі з наступними об'єктами управління:

- лінійними трактами;
- системами передачі з частотним розподілом каналів;
- системами передачі з часовим розподілом каналів;
- мережними трактами (первинні, вторинні, третинні);
- радіорелейними трактами;
- лініями передачі.

Параметри оцінки стану вибраних об'єктів управління мережею наведені у таблиці 6.1. Вказані параметри отримані з аналізу нормативно-технічної документації експлуатації

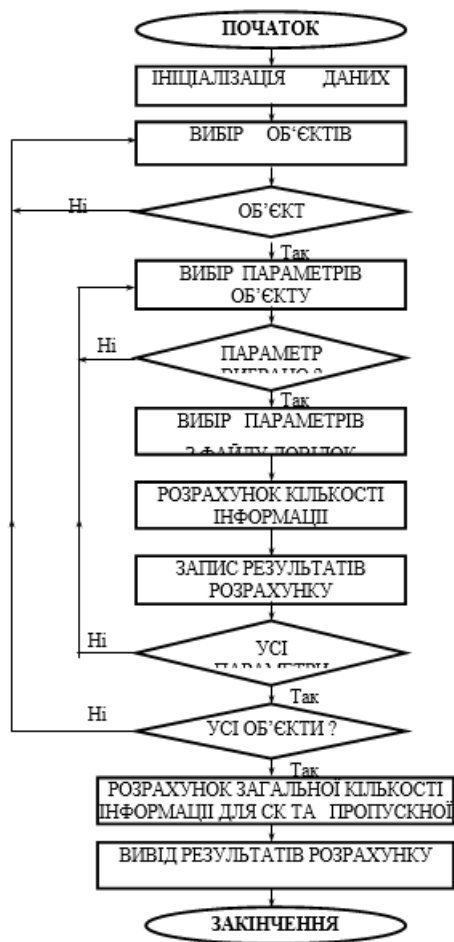


Рис. 6.4 Алгоритм головної процедури

Відхилення параметрів системи управління

Таблиця 6.1.

№ п/п	Параметри	Відхилення параметрів системи управління			
		До виконання	Після вмикання	При вмиканні	Під час виконання через інцидент, фактор
1	2	3	4	5	6
	Рівень контрольних частот трактів в системах передачі, дБ:				
1	Лінійних на виході.	13,5	1,35	4,5	6,75
2	Мережних з частотним розподілом каналів.	6,0	0,6	2,0	3,0
	Коефіцієнт трактів в мережах передачі з				
3	Лінійних	10^{-6}	10^{-10}	10^{-8}	10^{-7}
4	Мережних	10^{-6}	10^{-10}	10^{-8}	10^{-7}
5	Тиск повітря в кабелі, атм	0 15	0 015	0 05	0 075
6	Рівень КЧ ЛТ в НПП, дБ	12,9	1,3	4,3	6,45
7	Відхилення рівня струмів КЧ, дБ:				
	від номінальних значень на передачі	$\pm 1,0$	$\pm 0,1$	$\pm 0,3$	$\pm 0,5$
8	на виході лінійного тракту	$\pm 5,0$	$\pm 0,5$	$\pm 1,7$	$\pm 2,5$
9	Відхилення АЧХ в робочій смузі частот	$\pm 5,0$	$\pm 0,5$	$\pm 1,70$	$\pm 2,5$
1	Відхилення залишкового підсилення	$\pm 3,0$	$\pm 0,3$	$\pm 1,0$	$\pm 1,5$
0	лінійного тракту від його середнього значення				
1	Відхилення характеристик групового часу	$\pm 1,5$	$\pm 0,15$	$\pm 0,5$	$\pm 0,75$
1	проходження відносно частоти 2,5 МГц в				
1	Вихилення рівня КЧ 4,287 МГц ланцюга КК-	$\pm 4,0$	$\pm 0,4$	$\pm 1,33$	$\pm 2,0$
2	ОПП на виході приймальної частини стойки,				
1	Відхилення АЧХ у робочій смузі частот на				
3	частоті 4 МГц ланцюга КК-ОПП				
	КК-ОПП на виході приймальної частини	15,0	$\pm 1,5$	5,0	7,5
1	Різниця в рівні передачі (допуск на згасання	15,0	1,5	5,0	7,5
4	станційного кабелю) ¹ , дБ				
1	Різниця в рівнях передачі (допуск на згасання	15,0	1,5	5,0	7,5
5	станц. кабелю) ² , дБ				
1	Теж саме ³	5,0	0,5	1,7	2,5
1	Номінальний відносний рівень на вході тракту				
7	передачі ⁴ , дБ:				
	за потужністю	10,8	1,08	3,6	5,4
	за напругою	13,5	1,35	4,5	6,75
1	Теж саме ³ , дБ				
	за потужністю	10,8	1,08	3,6	5,4
	за напругою	12,6	1,26	4,2	6,3

№ № п/п	Параметри	Відхилення параметрів системи управління			
		До виконання	Після виплики	При випликан	Під час виконання через випадков. фактор
1	2	3	4	5	6
1	Номінальний відносний рівень на вході тракту				
9	прийому ⁴ , дБ:				
	за потужністю	6,9	0,69	2,3	3,45
	за напругою	9,6	0,96	3,2	4,8
2	Теж саме ³ , дБ:				
	за потужністю	6,9	0,69	2,3	3,45
	за напругою	8,7	0,87	2,9	4,35
2	Допустима середня потужність підсумкового	3,6	0,36	1,2	1,8
1	навант. в ЧНН ¹ , дБ/Вт				
2	Теж саме ³ , дБ/Вт	1,8	0,18	0,6	0,9
2	Теж саме ³ , дБ/Вт	0,9	0,09	0,3	0,45
2	Різниця між середнім і номінальним рівнем	±5,0	±0,5	±1,7	±2,5
4	прийому ⁴ , дБ				
2	Модифікація частоти передавального сигналу	2,0	0,2	0,7	1,0
5	для уніфікованої апаратури ¹ , Гц				
2	Точність підтримки діаграми рівня ⁴ , дБ	5,0	0,5	1,7	2,5
2	Допустиме відхилення рівня прийому ² , дБ	7,5	0,75	2,5	3,75
2	Точність встановленого рівня прийому ² , дБ	±3,0	±0,3	±1,0	±1,5
2	Різниця між середнім і номінальним рівнем	5,0	0,5	1,7	2,5
9	прийому ² , дБ				
3	Захищеність від підсумкових завад, не менше ⁴ ,	7,8	0,78	2,6	3,9
3	Точність підтримки діаграми рівня ³ , дБ	±10	±1,0	±3,3	±5,0
3	Допустиме відхилення рівня прийому ² , дБ	7,5	0,75	2,5	3,75
2					
3	Точність встановленого рівня	±3,0	±0,3	±1,0	±1,5
3	прийому ³ , дБ				
3	Різниця між середнім і номінальним	±5,0	±0,5	±1,7	±2,5
3	Захищеність від підсумкових завад, не	9,0	0,9	3,0	4,5
5	менше ² , дБ				
3	Точність підтримки діаграми рівня ³ , дБ	±10	±1	±3,3	±5
3	Рівень контрольних частот	5,0	0,5	1,7	2,5
3	Залишкове згасання ⁵ , дБ	5,1	0,51	1,7	2,55
3	Відносний рівень передачі на передачі ⁵ ,	3,9	0,39	1,3	1,95
9	дБ				
4	Теж саме, передачі на прийом ⁵ , дБ	1,2	0,12	0,4	0,6

№ № п/п	Параметри	Відхилення параметрів системи управління			
		До вмикання	Після вмикання	При вмиканні	Під час вимкнення через випадков. фактор
1	2	3	4	5	6
4	Потужність навантаження ³ , Вт	6,6	0,66	2,2	3,3
4	Рівень потужності завод (для довжини	16,5	1,65	5,3	8,25
2	320...640 км) ³ , дБ				
4	Електроживлення ⁶ , В				
3	Рівень струмів КЧ 18,432; 9,216: 0.768	±7,2	±0,72	±2,4	3,6
4	МПЦ*, дБ	±3,0	±0,3	±1,0	±1,5
4	Напруга джерел електроживлення	±50	±5	±17	±25,5
4	(акумуляторні батареї, ДП), В				
5					

Примітки: 1 —мережний третинний тракт; 2—мережний вторинний тракт; 3—мережний первинний тракт; 4—мережний третинний и вторинний тракти; 5 — параметр каналів ТЧ; 6 — параметр апаратури і обладнання; 7 — параметр лінії передач.

Розрахункові данні кількості інформації за зміною параметрів

Номер параметру	Кількість інформації управління за 1 с, біт	Приріст ентропії керованої мережі, біт	Кількість інформації з урахуванням інваріантності системи, біт керування, біт	Загальна кількість інформації біт
1	2	3	4	5
1	1.74	3.32	2.32	3296
2	6.64	13.29	9.97	12657
3	1.73	3.31	2.31	3275
4	1.58	3.32	2.32	3022
5	1.77	3.32	2.32	3347
6	1.77	3.32	2.32	3347
7	1.74	3.32	2.32	3296
8	1.74	3.32	2.32	3296
9	1.73	3.32	2.32	3289
10	1.74	3.32	2.32	3296
11	1.74	3.32	2.32	3296
12	6.64	13.29	9.97	12657
13	1.77	3.32	2.32	3347
14	1.74	3.32	2.32	3296
15	1.74	3.32	2.32	3296
16	1.74	3.32	2.32	3296
17	1.74	3.32	2.32	3296
18	1.74	3.32	2.32	3296
19	1.74	3.32	2.32	3296
20	1.74	3.32	2.32	3296
21	1.74	3.32	2.32	3296
22	1.77	3.32	2.32	3347
23	1.72	3.32	2.32	3270
24	1.74	3.32	2.32	3296
25	1.74	3.32	2.32	3296
26	1.74	3.32	2.32	3296
27	1.74	3.32	2.32	3296
28	1.74	3.32	2.32	3296
29	1.74	3.32	2.32	3296
30	1.77	3.32	2.32	3347
31	1.74	3.32	2.32	3296
32	1.72	3.32	2.32	3270
33	1.74	3.32	2.32	3296
34	1.74	3.32	2.32	3296
35	1.74	3.32	2.32	3296
36	1.77	3.32	2.32	3347
37	1.81	3.32	2.32	3423
38	1.77	3.32	2.32	3347
39	1.77	3.32	2.32	3347
40	1.74	3.32	2.32	3296
41	1.74	3.32	2.32	3296
42	1.74	3.32	2.32	3296

Номер параметру	Кількість інформації управління за 1 с, біт	Приріст ентропії керованої мережі, біт	Кількість інформації з урахуванням інваріантності системи, біт керування, біт	Загальна кількість інформації біт
1	2	3	4	5
43	1.74	3.32	2.32	3296
44	1.74	3.32	2.32	3296
45	1.74	3.32	2.32	3296

Для наведеного прикладу число об'єктів управління кожного типу прийнято рівним 30. Отримані вихідні дані представляють проміжні і кінцеві результати розрахунків, а саме:

- кількість інформації управління, що видається системою управління за рахунок відхилення параметрів при вимкненні системи управління на 1 с (див. табл. 6.1, графу 2);

- приріст ентропії стану керованою мережею (табл. 6.1, графа 3);

- загальна кількість інформації, що передається за проміжок часу перехідного процесу (60с) в системі управління, з урахуванням заданного числа об'єктів управління та вибраних для них параметрів (див. табл. 6.1, графа 5);

- **число каналів, необхідних для передачі розрахованої кількості інформації за різною пропускнуною спроможністю каналів (табл. 6.3).**

Число каналів, необхідних для передачі розрахованої кількості інформації за різною пропускнуною спроможністю каналів

Таблиця 6.3.

Пропускна спроможність каналу, біт/с	Загальна кількість інформації управління з урахуванням усіх параметрів, що регулюються,	Число каналів, необхідних для передачі інформації управління,
1200	177063	1148
2400	177063	74
4800	177063	37
9600	177063	18

Розрахунок параметрів для системи управління реальної мережі був виконаний запропонованим способом. Результати розрахунку були підтверджені даними, отриманими практичними методами на первинних мережах.

Інформаційно-ентропійний метод не залежить від характеру і кількості об'єктів управління та їх параметрів і тому є універсальним, тобто придатним для будь-якого типу мережі.

Контрольні питання

1. В яких одиницях вимірюється інформація?
2. В якому повідомленні міститься інформація, яка дорівнює одній двійковій одиниці?
3. Дайте визначення ентропії.
4. Які Ви знаєте властивості ентропії джерела незалежних дискретних повідомлень?
5. Як визначається ентропія джерела залежних дискретних повідомлень?
6. Як визначається середня кількість інформації в дискретному каналі з завадами?
7. Як визначити ентропії неперервних повідомлень?
8. Як визначити середню кількість інформації в неперервному каналі з завадами?
9. Якою є розмірність диференційної ентропії?
10. Як визначається ефективність систем передавання інформації?

РОЗДІЛ 7. ОПТИМАЛЬНИЙ ПРИЙОМ ДИСКРЕТНИХ СИГНАЛІВ. ОСНОВИ ТЕОРІЇ ЗАВАДОСТІЙКОСТІ

7.1. Загальні відомості про завадостійкість

Перешкодостійкість системи зв'язку називається здатність системи розрізняти (відновлювати) сигнали із заданою достовірністю.

Зазвичай спосіб передачі (спосіб кодування та модуляції) заданий і потрібно визначити завадостійкість, яку забезпечують різні способи прийому. Який із можливих способів прийому є оптимальним? Зазначені питання є предметом розгляду теорії завадостійкості, основи якої розроблено академіком В.А. Котельниковим.

Завдання визначення завадостійкості всієї системи загалом дуже складне. Тому часто визначають завадостійкість окремих ланок системи: приймача при заданому способі передачі, системи кодування або системи модуляції при заданому способі прийому і т.д.

Гранично досяжна завадостійкість називається, за Котельниковим, потенційною завадостійкістю. Порівняння потенційної та реальної перешкодостійкості пристрою дозволяє дати оцінку якості реального пристрою та знайти ще невикористані резерви. Знаючи, наприклад, потенційну завадостійкість приймача, можна судити, наскільки близька до неї реальна завадостійкість існуючих способів прийому і наскільки доцільно їхнє подальше вдосконалення при заданому способі передачі.

Відомості про потенційну завадостійкість приймача при різних способах передачі дозволяє порівняти ці способи передачі між собою та вказати, які з них у цьому відношенні є найбільш досконалими.

За відсутності перешкод кожному прийнятому сигналу відповідає цілком певний сигнал. За наявності перешкод ця однозначна відповідність порушується. Перешкода, впливаючи на сигнал, що передається, вносить невизначеність щодо того, яке з можливих повідомлень було передано, і по прийнятому сигналу x тільки з деякою ймовірністю можна судити про те, що був переданий той або інший сигнал s . Ця невизначеність описується апостеріорним розподілом ймовірностей $P(s/x)$.

Якщо відомі статистичні властивості сигналу та перешкоди ω , то можна створити приймач, який на підставі аналізу сигналу x знаходитиме апостеріорний розподіл $P(s/x)$. Потім за видом цього розподілу приймається рішення про те, яке з можливих повідомлень було передано. Рішення приймається оператором чи самим приймачем за правилом, що визначається заданим критерієм.

Завдання полягає в тому, щоб відтворити повідомлення, що передається найкращим чином у сенсі обраного критерію. Такий приймач називається оптимальним, яке перешкодостійкість буде максимальної при заданому способі передачі.

Незважаючи на випадковий характер сигналів x , у більшості випадків є можливість виділити безліч найбільш ймовірних сигналів $\{x_i\}$, що відповідають передачі деякого сигналу s_i . Геометрична вистава

дозволяє безліч сигналів замінити областю багатовимірного простору.

Нехай область X прийнятих сигналів розбита на області, що не перекриваються $X_j (j = 1, 2, \dots, m)$, причому кожному сигналу s_i відповідає область X_i . Якщо прийнятий сигнал потрапив у цю область, приймач приймає рішення про те, що передавався сигнал s_i . З деякою ймовірністю сигнал x_i може потрапити в будь-яку іншу область $X_j (j \neq i)$, тоді приймається помилкове рішення: замість сигналу s_i відтворюється сигнал s_j . Ймовірність того, що переданий сигнал прийнято правильно, дорівнює $P(x_i/s_i)$, а ймовірність того, що він прийнятий помилково, дорівнює

$$1 - P(x_i/s_i) = \sum_{j \neq i} P(x_j/s_i). \quad (7.1)$$

Умовна ймовірність $P(x_j/s_i)$ залежить від способу формування сигналу, від перешкод, що є в каналі, та від обраної вирішальної схеми приймача.

Повна ймовірність помилкового прийому елемента сигналу, очевидно, дорівнюватиме:

$$P_0 = \sum_{i=1}^m P(s_i) [1 - P(x_i/s_i)]. \quad (7.2)$$

де $P(s_i)$ – априорні ймовірності переданих сигналів.

У разі двійкового каналу область сигналів, що приймаються, розбивається на дві області X_1 і X_2 . Якщо сигнал x потрапляє у область X_1 , то відтворюється сигнал s_1 , а якщо область X_2 , то $-s_2$.

7.2. Критерії оптимального прийому сигналів

Щоб визначити, яка з вирішальних схем є оптимальною, необхідно насамперед встановити, у сенсі розуміється оптимальність. Вибір критерію оптимальності не є універсальним, він залежить від поставленого завдання та умов роботи системи.

Нехай на вхід приймача надходить сума сигналу та перешкоди $x(t) = s_k(t) + \omega(t)$,

де $s_k(t)$ - Сигнал, якому відповідає кодовий символ a_k , $\omega(t)$ - адитивна

перешкода із відомим законом розподілу. Сигнал s_k у місці прийому є випадковим з апіорним розподілом $P(s_k)$. На підставі аналізу коливання

$x(t)$ приймач відтворює сигнал s_i . За наявності перешкод це відтворення може бути абсолютно точним. За прийнятою реалізацією сигналу приймач обчислює апостеріорний розподіл $P(s_i/x)$ містить всі відомості, які можна витягти з прийнятої реалізації сигналу $x(t)$. Тепер необхідно встановити критерій, за яким приймач видаватиме на основі апостеріорного розподілу $P(s_i/x)$ рішення щодо переданого сигналу s_k .

Під час передачі дискретних повідомлень широко використовується критерій Котельникова (критерій ідеального спостерігача). Відповідно до цього критерію приймається рішення, що переданий сигнал s_i , для якого апостеріорна ймовірність $P(s_i/x)$ має найбільше значення, тобто. реєструється сигнал s_i , якщо виконуються нерівності

$$P(s_i/x)P(s_j/x), j \neq i. \quad (7.3)$$

При використанні такого критерію, повна ймовірність помилкового рішення буде мінімальною. Дійсно, якщо за сигналом x приймається рішення про те, що був переданий сигнал s_i , то, очевидно, вірогідність правильного рішення дорівнюватиме $P(s_i/x)$, А ймовірність помилки $1 - P(s_i/x)$. Звідси випливає, що максимуму апостеріорної ймовірності $P(s_i/x)$ відповідає мінімум повної ймовірності помилки (7.2).

На підставі формули Байєса

$$P(s_i/x) = \frac{P(s_i)p(x/s_i)}{p(x)}. \quad (7.4)$$

Тоді нерівність (7.3) можна записати в іншому вигляді

$$P(s_i)p(x/s_i)P(s_j)p(x/s_j) \quad (7.5)$$

або

$$\frac{p(x/s_i) P(s_i)}{p(x/s_j) P(s_j)}. \quad (7.6)$$

Функцію $p(x/s)$ часто називають функцією правдоподібності. Чим більше значення цієї функції при даній реалізації сигналу x , тим більш правдоподібним, що передавався сигнал s . Відношення, що входить у нерівність (5.5),

$$\Lambda = \frac{p(x/s_i)}{p(x/s_j)}. \quad (7.7)$$

називається ставленням правдоподібності. Користуючись цим поняттям, правило рішення (7.6), яке відповідає критерію Котельникова, можна записати у вигляді

$$L \frac{p(s_j)}{p(s_i)}. \quad (7.8)$$

Якщо сигнали, що передаються, рівноймовірні $P(s_i) = P(s_j) = \frac{1}{m}$, то це правило рішення набуває більш простого вигляду

$$L1. \quad (7.9)$$

Таким чином, критерій ідеального спостерігача зводиться до порівняння відносин правдоподібності (7.8). Цей критерій є загальнішим і називається критерієм максимальної правдоподібності.

Розглянемо бінарну систему, в якій передача повідомлень здійснюється за допомогою двох сигналів $s_1(t)$ та $s_2(t)$, відповідних двом кодовим символам a_1 та a_2 . Рішення приймається за результатом обробки прийнятого коливання $x(t)$ пороговим методом: реєструється s_1 , якщо xx_0 , і s_2 , якщо xx_0 де x_0 - деякий пороговий рівень x . Тут можуть бути помилки двох видів: відтворюється s_1 , коли передавався s_2 та s_2 , коли передавався s_1 . Умовні ймовірності цих помилок (ймовірності переходів) дорівнюватимуть:

$$P12=P(s1/s2)=\int_{-\infty}^{x_0} p(x/s_2)dx, \quad (7.10)$$

$$P21=P(s2/s1)=\int_{x_0}^{\infty} p(x/s_1)dx. \quad (7.11)$$

Значення цих інтегралів може бути обчислені як відповідні площі, обмежені графіком щільностей умовного розподілу ймовірностей (рис.7.1). Ймовірності помилок першого та другого виду відповідно:

$$P1=P(s2)P(s1/s2)=P2P12, \quad (7.12)$$

$$P11=P(s1)P(s2/s1)=P1P21. \quad (7.13)$$

Повна ймовірність помилки при цьому

$$P0 = P1 + P11 = P2P12 + P1P21. \quad (7.14)$$

Нехай $P1 = P2$, тоді

$$P0=\frac{1}{2}(P12+P21).$$

Неважко переконатися, що у цьому випадку мінімум $P0$ має місце за $P12=P21$, тобто. при виборі порога відповідно до рис. 7.1. Для такого порога $P0 = P12 = P21$. На рис.

7.1 значення P_0 визначається заштрихованою площею. За будь-якого іншого значення порога величина P_0 буде більшою.

Незважаючи на природність та простоту, критерій Котельникова має недоліки. Перший полягає в тому, що для побудови вирішальної схеми, як це випливає із співвідношення (7.5), необхідно знати апіорні можливості передачі різних символів коду. Другим недоліком цього критерію є те, що всі помилки вважаються однаково небажаними (мають однакову вагу). У деяких випадках таке припущення не є правильним. Наприклад, при передачі чисел помилка в перших цифрах більш небезпечна, ніж помилка в останніх цифрах. Перепустка команди або помилкова тривога в різних системах оповіщення можуть мати різні наслідки.

Отже, в загальному випадку при виборі критерію оптимального прийому необхідно враховувати втрати, які несе одержувач повідомлення при різних видах помилок. Ці втрати можна висловити деякими ваговими коефіцієнтами, що приписуються кожному з помилкових рішень. Позначимо втрати помилкових рішень першого та другого видів відповідно L_{12} та L_{21} . Тоді можна визначити середні очікувані втрати чи середній ризик

$$r = L_{12}P_1 + L_{21}P_{11} = L_{12}P_2P_{12} + L_{21}P_1P_{21}. \quad (7.15)$$

Оптимальною вирішальною схемою буде така, що забезпечує мінімум середнього ризику. Критерій мінімального ризику належить до класу про байесових критеріїв.

У радіолокації широко використовується критерій Неймана-Пірсона. При виборі цього критерію враховується, по-перше, що помилкова тривога і пропуск мети є рівноцінним за своїми наслідками, і, по-друге, що невідома апіорна ймовірність переданого сигналу. Якщо перепустка мети є більш небажаною, то можна задати деяку величину β допустимої ймовірності помилкової тривоги і зажадати, щоб вирішальна схема максимізувала ймовірність правильного виявлення Робн (або, що те саме, мінімізувати ймовірність пропуску $P_{пр}$).

Згідно з критерієм Неймана-Пірсона приймач є оптимальним у тому випадку, якщо при заданій ймовірності помилкової тривоги

$$P_{л} = \int_{x_0}^{\infty} p(x/0) dx = \beta \quad (7.16)$$

він забезпечує найбільшу ймовірність правильного виявлення

$$P_{обн} = 1 - P_{пр} = 1 - \int_{-\infty}^{x_0} p(x/s) dx. \quad (7.17)$$

Можна показати, що критерій Неймана-Пірсона призводить до наступного правила рішення: ціль вважається виявленою, якщо

$$A = \frac{p(x/s)}{p(x/0)} \lambda, \quad (7.18)$$

де λ - Деяке число, що визначається допустимою ймовірністю помилкової тривоги β .

7.3. Оптимальний прийом дискретних сигналів

Джерело дискретних повідомлень характеризується сукупністю можливих елементів повідомлення $u_1, u_2, \dots, u_m$ та ймовірностями появи цих елементів на виході джерела $P(u_1), P(u_2), \dots, P(u_m)$. У передавальному пристрої повідомлення перетворюється на сигнал таким чином, що кожному елементу повідомлення відповідає певний сигнал. Позначимо ці сигнали через $s_1, s_2, \dots, s_m$, а їх ймовірність появи на виході передавачів (апріорні ймовірності) відповідно через $P(s_1), P(s_2), \dots, P(s_m)$. Очевидно, апріорні ймовірності сигналів $P(s_i)$ дорівнюють апріорним ймовірностям $P(u_i)$ відповідних повідомлень $P(s_i) = P(u_i)$. У процесі передачі сигнал накладається перешкода. Нехай ця перешкода має рівномірний спектр потужності з інтенсивністю $\frac{1}{2} N_0$. Тоді сигнал на вході можна подати як суму переданого сигналу $s_i(t)$ та перешкоди $\omega(t)$:

$$x(t) = s_i(t) + \omega(t) \quad (i = \overline{1, m}). \quad (7.19)$$

Оскільки сигнали $s_i(t)$, $x(t)$ та перешкода $\omega(t)$ задані на кінцевому інтервалі (0), то їх можна подати у вигляді розкладів по ортогональних функцій:

$$s_i(t) = \sum_l s_{il} \phi_l(t), \quad (7.20)$$

$$\omega(t) = \sum_l \omega_l \phi_l(t), \quad (7.21)$$

$$x(t) = \sum_l x_l \phi_l(t), \quad (7.22)$$

де

$$x_l = s_{il} + \omega_l, \quad (7.23)$$

$$s_{il} = \int_0^T s_i(t) \phi_l(t) dt, \quad (7.24)$$

$$\omega_l = \int_0^T \omega(t) \phi_l(t) dt. \quad (7.25)$$

Так як ми припускаємо, що перешкода має нормальний розподіл, то й коефіцієнти Фур'є у виразі (7.25) матимуть нормальний розподіл із дисперсією $\sigma_l^2 = \frac{1}{2} N_0$ і середнім значенням, що дорівнює нулю:

$$p(\omega_l) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{\omega_l^2}{N_0}\right). \quad (7.26)$$

Коефіцієнти x_l також мають нормальний розподіл із тією ж дисперсією $\frac{1}{2}N_0$ та середнім значенням s_{il}

$$p(x_l/s_{il}) = \frac{1}{\sqrt{\pi N_0}} \exp \left[-\frac{(x_l - s_{il})^2}{N_0} \right]. \quad (7.27)$$

Через незалежність коефіцієнтів ω_l багатовимірний розподіл коефіцієнтів x_l , тобто. умовний розподіл $p(x/s_i)$, дорівнюватиме добутку одномірних розподілів (5.25):

$$p(x/s_i) = \prod_{l=l_1}^{l_2} p(x_l/s_{il}) = (\pi N_0)^{-n/2} \exp \left[-\frac{1}{N_0} \sum_{l=l_1}^{l_2} (x_l - s_{il})^2 \right]. \quad (7.28)$$

Підставляючи цей вираз у (7.6), отримаємо наступну нерівність, що визначає умову оптимального прийому за Котельниковим:

$$\frac{\prod p(x/s_i)}{\prod p(x/s_j)} = \frac{\exp \left[-\frac{1}{N_0} \sum_{l=l_1}^{l_2} (x_l - s_{il})^2 \right] P(s_j)}{\exp \left[-\frac{1}{N_0} \sum_{l=l_1}^{l_2} (x_l - s_{jl})^2 \right] P(s_i)}. \quad (7.29)$$

Логарифмуючи обидві частини нерівності, приходимо до еквівалентного виразу

$$\sum_{l=l_1}^{l_2} (x_l - s_{il})^2 - \sum_{l=l_1}^{l_2} (x_l - s_{jl})^2 \quad (7.30)$$

Відповідно до виразів (5.20) та (5.18) маємо

$$x(t) - s_i(t) = \sum_{l=l_1}^{l_2} (x_l - s_{il}) \phi_l(t). \quad (7.31)$$

Після зведення у квадрат та усереднення за часом виразу (7.31) з урахуванням властивостей ортогональних функцій $\phi_l(t)$ отримуємо

$$\int_0^{T_0} [x(t) - s_i(t)]^2 dt = \sum_{l=l_1}^{l_2} (x_l - s_{il})^2. \quad (7.32)$$

Тоді умову оптимального прийому (7.30) можна записати в іншому вигляді:

$$\int_0^T [x(t) - s_i(t)]^2 dt - \int_0^T [x(t) - s_j(t)]^2 dt \quad (7.33)$$

Нерівності (7.29) або їм еквівалентні нерівності (7.30) та (7.33) визначають умови правильного прийому сигналів $s_i(t)$.

У випадку, коли апіорні ймовірності сигналів однакові $P(s_1) = P(s_2) = \dots = P(s_m) = \frac{1}{m}$, критерій Котельникова набуває більш простого вигляду:

$$\int_0^T [x(t) - s_i(t)]^2 dt \quad j \neq i. \quad (7.34)$$

Звідси випливає, що при рівноймовірних сигналах оптимальний приймач відтворює повідомлення, відповідне переданому сигналу, який має найменше середньоквадратичне відхилення від прийнятого сигналу.

Нерівність (7.34) можна записати в іншому вигляді, розкривши дужки:

$$\int_0^T x^2(t) dt + \int_0^T s_i^2(t) dt - 2 \int_0^T x(t) s_i(t) dt.$$

Для сигналів, енергії яких однакові, це нерівність всім $j \neq i$ набуває більш простої форми:

$$\int_0^T x(t) s_i(t) dt \quad j \neq i \quad (7.35)$$

І тут умова оптимального прийому можна сформулювати в такий спосіб. Якщо всі можливі сигнали рівноймовірні і мають однакову енергію, оптимальний приймач відтворює повідомлення, відповідне переданому сигналу, взаємна кореляція якого з прийнятим сигналом максимальна.

Для двійкової системи отриманим результатам можна дати дуже наочне геометричне трактування. Нехай передаються s_2 . Першому сигналу відповідає вектор s_1 у n -мірному просторі, а другому – вектор s_2 . Прийнятому сигналу відповідає вектор x , що дорівнює сумі векторів сигналу s і перешкоди ω . Простір можливих значень можна розбити на дві області так, щоб при попаданні кінця вектора x в першу область відтворювався сигнал s_1

(область сигналу s_1), а при попаданні в іншу область – відтворювався сигнал s_2 (область сигналу s_2). Якщо x , відповідний даному сигналу, потрапляє у область іншого сигналу, відбувається помилка (замість s_1 відтворюється s_2 або навпаки, s_2 замість s_1). Імовірність помилки, очевидно, залежить від конфігурації областей сигналу. У оптимальному приймачі Котельникова простір сигналів розбивається області сигналу s_1 та сигналу s_2 так, щоб повна ймовірність помилки (5.13) була мінімальною.

У разі рівноймовірних сигналів та перешкод з рівномірним розподілом оптимальним розбиттям простору буде таке, при якому будь-яка точка відноситься до області сигналу s , кінець вектора якого ближче всього до точки x . У двовірній моделі (рис. 7.2 а). Для двійкової системи межа областей сигналів s_1 і s_2 є геометричне місце точок, рівновіддалених від s_1 і s_2 , тобто. гіперплощина, перпендикулярна до вектора різниці $\Delta s = s_1 - s_2$ і ділить його навпіл.

Якщо, наприклад, передавався сигнал s_1 , то помилка станеться у тому випадку, коли виконується нерівність

$$r_2 \quad (7.36)$$

або

$$\mu \frac{1}{2} d, \quad (7.37)$$

де $r_1 = \|x - s_1\|, r_2 = \|x - s_2\|$. Проекція ω на вектор, колінеарний Δs , тобто $\mu = \frac{\omega \Delta s}{\|\Delta s\|}$. Замість нерівності (5.43) можна записати $\|x - s_2\|^2$ або в евклідовій метриці:

$$\int_0^T [x(t) - s_2(t)]^2 dt. \quad (7.38)$$

Умова (7.38) повністю збігається з умовою (7.34).

Структурна схема приймача, реалізує операції (7.38), наведено на рис.7.26. Тут Γ_1 і Γ_2 - генератори опорних сигналів, що формують точні копії переданих сигналів s_1 і s_2 , В - пристрій, що віднімає, КВ - квадратує пристрій, І - інтегратор, РУ - схема порівняння і вибору (вирішальний пристрій).

Нерівність (7.38) можна записати в іншому вигляді, розкривши дужки під інтегралами:

$$-2 \int_0^T x(t) s_2(t) dt + \int_0^T s_2^2(t) dt$$

або

$$\int_0^T x(t) [s_1(t) - s_2(t)] dt \frac{1}{2} (E_2 - E_1). \quad (7.39)$$

Ця нерівність абсолютно еквівалентна нерівності (7.38), але вона веде до іншої схемної реалізації оптимального приймача.

На рис.7.3а наведено структурну схему приймача, що реалізує умови роботи (7.39). У цій схемі після операції перемноження (П) та інтегрування (І) проводиться порівняння отриманого результату з постійним порогом, рівним різниці енергій сигналів $\frac{1}{2} (E_2 - E_1)$. Ця схема простіше, ніж схема рис.7.2б. Однак вона має той недолік, що при зміні рівня сигналів поріг потрібно автоматично регулювати. Цей недолік усувається, якщо сигнали мають рівну енергію $E_2 = E_1$, тоді поріг дорівнює нулю та вирішальна схема визначає лише знак сигналу на виході.

При $E_2 = E_1$ спрощується і схема приймача рис.7.2б. Розкриваючи дужки (7.38), отримуємо умови оптимального прийому в наступному вигляді:

$$\int_0^T x(t) s_1(t) dt, \quad (7.40)$$

що збігається з умовою (5.33). Таким чином, при $E_2 = E_1$ приймач Котельникова перетворюється на кореляційний (когерентний) (рис.7.3б).

Оптимальний прийом можна також реалізувати у схемі з узгодженими лінійними фільтрами (рис.7.3в), імпульсні реакції яких мають бути $g_1(t) = cs_i(T - t)$, де c - постійний коефіцієнт.

Розглянуті схеми оптимальних приймачів належать до типу когерентних, у яких враховується як амплітуда, а й фаза високочастотного сигналу.

Зауважимо, що у схемах оптимальних приймачів відсутні фільтри на вході, які у реальних приймачах завжди є. Це означає, що оптимальний приймач флуктуаційних перешкод не вимагає фільтрації на вході. Його завадостійкість, як ми побачимо далі, не залежить від ширини смуги пропускання приймача.

7.4. Ймовірність помилки при когерентному прийому двоєкових сигналів

Визначимо ймовірність помилки у системі передачі двоєкових сигналів прийому оптимальний приймач. Ця ймовірність, очевидно, буде мінімально можливою і характеризуватиме потенційну завадостійкість при даному способі передачі. При прийомі на реальний приймач завадостійкість може дорівнювати потенційної, але не може бути більше її.

Нехай сигнал приймає значення $s_1(t)$ з ймовірністю P_1 та значення $s_2(t)$ з ймовірністю P_2 . Якщо передавався сигнал s_1 , то згідно з умовою (7.33) помилка відбудеться в тому випадку, коли

$$\int_0^T [x(t) - s_1(t)]^2 dt - \int_0^T [x(t) - s_2(t)]^2 N_0 \ln \frac{P_1}{P_2}. \quad (7.41)$$

Так як $x(t) = s_1(t) + \omega(t)$, то нерівність (7.41) може бути приведена до вигляду

$$\int_0^T \omega^2(t) dt - \int_0^T [s_1(t) - s_2(t) + \omega(t)]^2 dt N_0 \ln \frac{P_1}{P_2}$$

або

$$\int_0^T \omega(t)[s_1(t) - s_2(t)] dt. \quad (7.42)$$

Відповідно до виразів (7.20) і (7.21) ліву частину нерівності (7.42) можна записати в наступному вигляді:

$$\xi = \int_0^T \omega(t)[s_1(t) - s_2(t)] dt = \sum_l \omega_l (s_{1l} - s_{2l}). \quad (7.43)$$

Оскільки кожен коефіцієнт ω_l має нормальний розподіл із середнім значенням, рівним нулю, то сума (7.43) буде також являти собою нормальну випадкову величину ξ з нульовим середнім значенням та дисперсією

$$D\xi = \overline{\xi^2} = \sum_l \overline{\omega_l^2} (s_{1l} - s_{2l})^2 = \frac{1}{2} N_0 \int_0^T [s_1(t) - s_2(t)]^2 dt = \sigma_\xi^2. \quad (7.44)$$

Щільність ймовірності випадкової величини ξ

$$P(\xi) = \frac{1}{\sqrt{2\pi\sigma_\xi^2}} e^{-\frac{\xi^2}{2\sigma_\xi^2}}.$$

Згідно з (7.42) помилка відбудеться при передачі сигналу s_1 , якщо

$$\xi.$$

Величина цієї помилки дорівнюватиме:

$$P_{21} = P(\xi. \quad (7.45)$$

Вводячи нову змінну $u = \xi/\sigma_\xi$, на підставі співвідношення, маємо

$$P_{21} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{A}{\sigma_\xi}} e^{-\frac{u^2}{2}} du = \frac{1}{2} \left[1 + \Phi\left(\frac{A}{\sigma_\xi}\right) \right],$$

де

$$\frac{A}{\sigma_\xi} = \frac{\frac{N_0}{2} \ln \frac{P_2}{P_1} - \frac{1}{2} \int_0^T [s_1(t) - s_2(t)]^2 dt}{\sqrt{\frac{N_0}{2} \int_0^T [s_1(t) - s_2(t)]^2 dt}}.$$

Після нескладних перетворень остаточно отримуємо

$$P_{21} = \frac{1}{2} [1 - \Phi(a_{21})], \quad (7.46)$$

де

$$a_{21} = a + \frac{1}{2a} \ln \frac{P_1}{P_2}, \quad (7.47)$$

$$a^2 = \frac{1}{2N_0} \int_0^T [s_1(t) - s_2(t)]^2 dt. \quad (7.48)$$

Цілоком аналогічно визначається ймовірність помилки передачі сигналу $s_2(t)$:

$$P_{12} = \frac{1}{2} [1 - \Phi(a_{12})], \quad (7.49)$$

де

$$a_{12} = a + \frac{1}{2a} \ln \frac{P_2}{P_1}. \quad (7.50)$$

Повна ймовірність помилки при оптимальному прийомі бінарних сигналів $s_1(t)$ і $s_2(t)$ дорівнюватиме:

$$P_0 = P_1 P_{21} + P_2 P_{12} \quad (7.51)$$

або згідно (7.46) та (7.49)

$$P_0 = \frac{1}{2} P_1 [1 - \Phi(a_{21})] + \frac{1}{2} P_2 [1 - \Phi(a_{12})]. \quad (7.52)$$

З отриманих формул випливає, що ймовірність помилки, що визначає потенційну стійкість до перешкод, залежить від двох величин: $a^2 P_2 / P_1$. Перша величина визначається ставленням питомої енергії різниці сигналів до інтенсивності перешкоди N_0 . Чим більше це ставлення, тим більша потенційна завадостійкість. Відношення апіорних ймовірностей P_2 / P_1 визначається статистичними властивостями повідомлень, що передаються.

Якщо сигнали, що передаються, рівноймовірні $P_1 = P_2 = 0,5$, то ф-ла (7.52) спрощується і набуває вигляду

$$P_0 = \frac{1}{2} [1 - \Phi(a)]. \quad (7.53)$$

Формулу (7.53) легко одержати з геометричних уявлень. Як видно з рис. (5.3), при передачі сигналу s_1 помилка відбудеться в тому випадку, якщо виконуватиметься нерівність $r_1 r_2$ або $\mu \frac{1}{2} d$. Отже, ймовірність помилки можна як ймовірність виконання однієї з цих нерівностей, тобто.

$$P_0 = P(r_1 | r_2) = P\left(\mu \left| \frac{1}{2} d \right.\right).$$

Помноживши обидві частини нерівності на d , одержуємо

$$P_0 = P\left(\xi \left| \frac{1}{2} d^2 \right.\right),$$

де $\xi = \mu d \omega \Delta s$ - Випадкова величина (7.43), що має нормальний розподіл з дисперсією $D_\xi = \frac{1}{2} N_0 d^2$ (7.44). Тоді маємо

$$P_0 = P\left(\xi \frac{1}{2} d^2\right) = \frac{1}{2} [1 - \Phi(a)],$$

де

$$a^2 = \frac{1}{4} \left(\frac{d^2}{\sqrt{D_\xi}} \right)^2 = \frac{1}{2} \frac{d^2}{N_0} = \frac{1}{2N_0} \int_0^T [s_1(t) - s_2(t)]^2 dt,$$

що збігається з (7.53) та (7.48).

При малій інтенсивності перешкод, коли $a \ll 1$, у ф-лах (7.47) та (7.50) другим членом можна знехтувати. І тут ф-ла (7.52) також наводиться до ф-ле (7.53). Ймовірність помилки при цьому практично не залежить від P_1 і P_2 . При великому рівні перешкод, коли мало, залежність ймовірності помилки від ставлення апріорних ймовірностей P_1/P_2 стає помітною. Зі збільшенням цього відношення ймовірність помилки збільшується.

Таким чином, при рівноймовірних сигналах ймовірність помилки повністю визначається величиною a . Значення цієї величини залежить від спектральної густини перешкод N_0 і переданих сигналів $s_1(t)$ і $s_2(t)$.

Для систем з активною паузою, у яких сигнали мають однакову енергію $\overline{Ts_1^2(t)} = \overline{Ts_2^2(t)} = E$, вираз для a^2 можна уявити в наступному вигляді:

$$a^2 = \frac{E(1-p_{12})}{N_0} = q_0(1-p_{12}), \quad (7.54)$$

де $p_{12} = \frac{1}{E} \int_0^T s_1(t)s_2(t)dt$ - Коефіцієнт взаємної кореляції між сигналами, $q_0 = E/N_0$ - відношення енергії сигналу до питомої потужності перешкоди.

Імовірність помилки для таких систем визначається формулою

$$P_0 = \frac{1}{2} [1 - \Phi(\sqrt{q_0(1-p_{12})})]. \quad (7.55)$$

Звідси випливає, що за $p_{12} = -1$, тобто, $s_1(t) = -s_2(t)$, Система забезпечує найбільшу потенційну завадостійкість. Ця система із протилежними сигналами. Для неї $a_{\max}^2 = 2q_0$. Практичною реалізацією системи із протилежними сигналами є система з фазовою маніпуляцією.

Порівняння різних систем передачі дискретних повідомлень зручно проводити за параметром a^2 , що представляє собою наведене відношення сигналу до перешкоди на виході оптимального приймача при заданому способі передачі $a^2 = q_{\text{вых}} = q_0(1-p_{12}) \approx \text{TF}(1-p_{12})q_{\text{вх}}$, або за величиною виграшу

$$\gamma = \frac{q_{\text{вых}}}{q_{\text{вх}}} = \text{TF}(1-p_{12}), \quad (7.56)$$

де $q_0 = \text{TF}q_{\text{вх}}$. Множник TF визначає виграш за рахунок оптимальної обробки сигналу на прийомі, а $(1-p_{12})$ - за рахунок способу передачі.

Загалом радіотелеграфний сигнал можна записати

$$s_i(t) = A_i(t)\cos(\omega_i t + \phi_i), 0, i = 1, 2, \quad (7.57)$$

де параметри коливання A_i , ω_i , ϕ_i набувають певних значень залежно від виду маніпуляції. Відповідно до (7.48) для сигналів (7.57) маємо:

$$a^2 = \frac{\tau_0}{4N_0} \left\{ \overline{A_1^2(t)} + \overline{A_2^2(t)} - \frac{2\overline{A_1(t)A_2(t)}}{\tau_0(\omega_1 - \omega_2)} \times [\sin[(\omega_1 - \omega_2)\tau_0 + \phi_1 - \phi_2] - \sin(\phi_1 - \phi_2)] \right\}. \quad (7.58)$$

Для амплітудної маніпуляції $A_1(t) = A_0, A_2(t) = 0, \omega_1 = \omega_2 = \omega_0$

$$a_{AM}^2 = \frac{A_0^2 \tau_0}{4N_0} = \frac{1}{2} q_0.$$

Для частотної маніпуляції $A_1(t) = A_2(t) = A_0, \omega_1 \neq \omega_2$. При оптимальному виборі рознесення частот $(\omega_1 - \omega_2)2\pi = k/\tau_0$, де k - ціле число і $\phi_1 = \phi_2$. Тоді на підставі (5.56) отримуємо

$$a_{ЧМ}^2 = q_0.$$

Для фазової маніпуляції $A_1(t) = A_2(t) = A_0, \omega_1 = \omega_2 = \omega_0, \phi_1 - \phi_2 = \pi$

$$a_{ФМ}^2 = 2q_0.$$

Порівняння отриманих формул показує, що з усіх систем передачі бінарних сигналів найбільшу потенційну стійкість до перешкод забезпечує система з фазовою маніпуляцією. Порівняно з ЧС вона дозволяє здобути дворазовий, а порівняно з АМ - чотириразовий виграш за потужністю.

У системах зв'язку сигнал зазвичай складається із послідовності простих сигналів. Так, у телеграфії кожній букві відповідає кодова комбінація, що складається з п'яти елементарних посилок. Можливі складніші комбінації. Якщо елементарні сигнали, що становлять кодову комбінацію, незалежні, то ймовірність помилкового прийому кодової комбінації визначається такою формулою:

$$P_{ок} = 1 - (1 - P_0)^n, \quad (7.59)$$

де P_0 – ймовірність помилки елементарного сигналу, p – число елементарних сигналів кодової комбінації (значність коду).

Слід зауважити, що ймовірність помилки у розглянутих вище випадках повністю визначається ставленням енергії сигналу до спектральної густини перешкоди і не залежить від форми сигналу. У випадку, коли спектр перешкоди відрізняється від рівномірного, ймовірність помилки можна зменшити, змінюючи спектр сигналу, тобто. його форму.

7.5. Ймовірність помилки при когерентному прийому багатопозиційних сигналів

Робота приймача у багатопозиційних системах зводиться до розрізнення m сигналів, відповідних m позицій коду. Схему приймача можна уявити, що складається з m каналів (гілок), кожен з яких розрахований на прийом одного певного сигналу. Одним із прикладів багатопозиційної системи є система з частотною маніпуляцією, в якій сигнали є гармонійними коливаннями різних частот. Приймач у цій системі містить m фільтрів, налаштованих на частоти переданих сигналів. За допомогою цих фільтрів здійснюється поділ (розрізнення) сигналів.

Нехай $s_1(t), s_2(t), \dots, s_m(t)$ – сигнали, що використовуються для передачі, та $\omega(t)$ – Адитивна перешкода, що впливає на приймач. Прийняті сигнали при цьому будуть $x_1, x_2, \dots, x_m$. Якщо передавався сигнал $s_1(t)$, то в першому фільтрі будуть сигнал та перешкода $x_1 = s_1 + \omega$, а інших фільтрах – лише перешкода. Приймач порівнює прийняті сигнали і відтворює найбільший їх, тобто. виносить рішення про те, що переданий k -й сигнал, якщо $x_k = \max(x_1, x_2, \dots, x_m)$. Оскільки ми припустили, що передавався сигнал s_1 , то вірогідність правильного рішення дорівнюватиме:

$$P_{\text{прав}} = P(x_1 x_2, x_3, \dots, x_m),$$

а ймовірність помилки

$$P_0 = 1 - P_{\text{прав}}.$$

Оптимальний когерентний приймач m -позиційної системи є багатоканальний корелятор або систему з m узгоджених фільтрів. Структурна схема такого приймача

аналогічна схемам рис.7.3б чи рис.7.3в для двійкових сигналів (різниця лише у числі каналів). У цьому випадку приймач відповідно до умов (7.35) обчислює функцію взаємної кореляції прийнятого сигналу $x(t)$ з усіма m опорними сигналами

$$y_k = \int_0^T s_k(t) x(t) dt = \int_0^T s_k(t) [s_1(t) + \omega(t)] dt, k = \overline{1, m}, \quad (7.60)$$

і видає рішення про те, що був переданий сигнал, для якого кореляція має найбільше значення. Можливість правильного рішення (якщо передається сигнал s_1) Дорівнюватиме:

$$P_{\text{прав}} = P(y_1 y_2, y_3, \dots, y_m), \quad (7.61)$$

На рис. 7.4,а наведено структурну схему когерентного приймача багатопозиційних сигналів, побудовану на базі кореляційної техніки. У кожному каналі цієї схеми проводяться синхронне детектування прийнятих сигналів, інтегрування та відлік наприкінці кожного елемента сигналу. Отримані відліки надходять на схему порівняння (вирішальний пристрій РУ). В результаті порівняння видається рішення про те, який з сигналів було передано.

Визначимо ймовірність помилки при оптимальному когерентному прийомі ортогональних m -позиційних сигналів. Вважатимемо, що всі можливі сигнали рівноймовірні і мають однакову енергію E . Для цього випадку умови правильного прийому сигналу s_1 згідно (7.35) запишуться

$$\int_0^T x(t) s_1(t) dt \geq \int_0^T x(t) s_j(t) dt, j = \overline{2, m}.$$

Так як $x(t) = s_1(t) + \omega(t)$, $\int_0^T s_1(t) s_j(t) dt = 0$ і $\int_0^T s_1^2(t) dt = E$, то ця нерівність набуває вигляду

$$\int_0^T \omega(t) [s_1(t) - s_j(t)] dt \geq E \quad (7.62)$$

або

$$\int_0^T s_j(t) \omega(t) dt. \quad (7.63)$$

Розглянемо функцію $\xi_j = \int_0^T s_j(t) \omega(t) dt$, Що представляє собою нормальну випадкову величину (перешкоду на виході j-го каналу) з дисперсією

$$D\xi_j = \sigma_\xi^2 = \overline{\xi_j^2} = \frac{1}{2} N_0 E.$$

Щільність ймовірності величини ξ_j

$$p(\xi_j) = \frac{1}{\sqrt{\pi N_0 E}} \exp\left(-\frac{\xi_j^2}{N_0 E}\right).$$

Ймовірність того, що перешкода в j-ном каналі не перевищить сумарного значення сигналу та перешкоди у першому каналі, тобто, ймовірність того, що ξ_j , дорівнюватиме:

$$\int_{-\infty}^{\xi_1 + E} \frac{\exp(-\xi_j^2 / N_0 E)}{\sqrt{\pi N_0 E}} d\xi_j = \frac{1}{2} [1 + \Phi(t + \sqrt{2q_0})],$$

$$\text{де } \Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, t = \frac{\xi_1}{\sqrt{\frac{N_0 E}{2}}} = \frac{\xi_1}{\sigma_\xi}, q_0 = \frac{E}{N_0}.$$

Ймовірність того, що перешкода ξ_j у всіх m-1 каналах без сигналу не перевищить сумарного значення сигналу та перешкоди в першому каналі,

$$\left(\frac{1}{2}\right)^{m-1} [1 + \Phi(t + \sqrt{2q_0})]^{m-1}.$$

Інтегрування цього виразу за всіма можливими значеннями перешкоди ξ_1 дає можливість правильного прийому

$$P_{\text{прав}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \left\{ \frac{1}{2} [1 + \Phi(t + \sqrt{2q_0})] \right\}^{m-1} dt. \quad (7.64)$$

Імовірність помилки при цьому

$$P_{\text{окг}} = 1 - P_{\text{прав}} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \left[\frac{1}{2} + \frac{1}{2} \Phi(t + \sqrt{2q_0}) \right]^{m-1} dt. \quad (7.65)$$

У окремому випадку при $m=2$ вираз (7.65) перетворюється на ф-лу (7.53) для двійкових систем

$$P_0 = \frac{1}{2} [1 - \Phi(\sqrt{q_0})].$$

При $m=2$ інтеграл у правій частині (7.65) можна обчислити наближеними методами. При відносно великих значеннях відношення сигналу до перешкод ($q_0 \gg 1$) має місце асимптотичний вираз

$$P_{\text{окг}} \approx \frac{m-1}{2} [1 - \Phi(\sqrt{q_0})]. \quad (7.66)$$

За неоптимального когерентного прийому дискретних сигналів у схемі рис. 7.4,а інтегратор відсутня. У цьому випадку після синхронного детектора ставиться фільтр нижніх частот і береться відлік (стробування) на виході фільтра всередині послілки. Можна показати, що ймовірність помилки при неоптимальному когерентному прийомі визначається отриманими вище виразами (7.65) та (7.66), якщо в останніх замість q_0 підставити

$$q = q_0 / \text{TF} = P_c / P_n.$$

7.6. Некогерентний приймання дискретних сигналів

При некогерентному прийомі інформація про фазу сигналів, що приймаються, не використовується. Такий спосіб прийому застосовується у каналах зі змінними параметрами, коли фаза сигналу випадково змінюється та її визначення викликає значні труднощі, а також у каналах із постійними параметрами з метою спрощення схеми приймача.

Оптимальний некогерентний приймач обчислює модуль (огональну) функції взаємної кореляції $z_k = \left| \int_0^T \dot{x}(t) \dot{s}_k(t) dt \right|$ і вирішує, що був переданий сигнал, для якого z в деякий момент часу $t = t_0$ має найбільше значення. Нехай передавався сигнал $s_1(t)$ тоді умову правильного прийому цього сигналу можна записати в наступному вигляді: z_k або

$$\left| \int_0^T \dot{x}(t) \dot{s}_k(t) dt \right|. \quad (7.67)$$

Схема приймача, що реалізує умову (5.65), наведено на рис.5.9. Ця схема містить m узгоджених фільтрів (Φ), відповідних m окремим сигналам. На виході кожного фільтра виходить напруга, пропорційна функції взаємної кореляції $y_k = \int_0^T \dot{x}(t) \dot{s}_k(t) dt$. Амплітудний детектор (Δ) виділяє обгинальну (модуль) цієї функції. Потім провадиться відлік і приймається рішення.

Припустимо, що сигнали рівноймовірні, мають однакову енергію і є ортогональними у посиленому значенні. За цих умов маємо

$$z_k^2 = \left[2 \int_0^T \omega(t) s_k(t) dt \right]^2 + \left[2 \int_0^T \omega(t) \dot{s}_k(t) dt \right]^2 = \xi_k^2 + \eta_k^2,$$

$$z_1^2 = (2E + \xi_1)^2 + \eta_1^2, \quad (7.68)$$

де

$$\xi_k = 2 \int_0^T \omega(t) s_k(t) dt, \eta_k = 2 \int_0^T \omega(t) \dot{s}_k(t) dt.$$

Випадкові величини ξ і η мають нормальний розподіл з нульовим середнім значенням та дисперсією, що дорівнює $\sigma^2 = \sigma_\xi^2 = \sigma_\eta^2 = 2N_0E$. У цьому вся легко переконатися як і, як було зроблено під час виведення ф-лы (7.44).

Випадкова величина z_k^2 є сумою квадратів двох незалежних випадкових величин ξ_k і η_k з нормальним розподілом, нульовим середнім значенням та однаковими дисперсіями, рівними $2N_0E$. Така величина, як відомо, має розподіл Релея. У нашому випадку

$$p(z_k) = \frac{z_k}{2N_0E} e^{-\frac{z_k^2}{4N_0E}}. \quad (7.69)$$

Випадкову величину z_1^2 можна розглядати як квадрат довжини векторної суми постійного вектора довжиною $L = 2E$ і випадкового вектора з нормально розподіленими незалежними складовими, що мають дисперсію $\sigma_1^2 = 2N_0E$. Тому величина z_1 підпорядковується узагальненому розподілу Релея із щільністю ймовірностей

$$p(z_1) = \frac{z_1}{2N_0E} \exp\left(-\frac{z_1^2 + L^2}{4N_0E}\right) I_0\left(\frac{z_1 L}{2N_0E}\right). \quad (7.70)$$

Випадкові величини z_k є нічим іншим, як огибающие напруги у каналах без сигналу, тобто. огинають перешкод. Оскільки перешкоди ми вважаємо гаусовими, то цим пояснюється, що z_k матиме релеївський розподіл. Випадкова величина z_1 є загальна сумарного колювання сигналу і перешкоди в каналі з сигналом, тому вона і підпорядковується закону узагальненого розподілу Релея.

Тепер можна визначити можливість помилки при некогерентному прийомі. У загальному випадку ця ймовірність дорівнюватиме:

$$P_{\text{онкг}} = 1 - P(z_2, z_3, \dots, z_m). \quad (7.71)$$

При бінарній передачі ($m = 2$)

$$P_{\text{онкг}} = 1 - P(z_1 z_2) = P(z_2 z_1).$$

Для обчислення ймовірності помилки спочатку обчислюється за певного фіксованого значення z_1 ймовірність того, що $z_2 z_1$. Ця можливість виражається інтегралом

$$I(z_1) = \int_{z_1}^{\infty} p(z_2) dz_2,$$

який має різні значення при різних z_1 . Для того, щоб знайти повну ймовірність $z_2 z_1$, необхідно $I(z_1)$ усереднити за всіма можливими значеннями z_1 відповідно до розподілу $p(z_1)$. Таким чином,

$$P_{\text{онкг}} = P(z_2 | z_1) = \int_0^{\infty} I(z_1) p(z_1) dz_1 = \int_0^{\infty} p(z_1) dz_1 \int_{z_1}^{\infty} p(z_2) dz_2. \quad (7.72)$$

Після підстановки у (7.72) виразів $p(z_2)$ і $p(z_1)$ відповідно (7.69) та (7.70) та інтегрування отримуємо наступний вираз для ймовірності помилки при оптимальному некогерентному прийомі двійкових сигналів:

$$P_{02} = \frac{1}{2} e^{-\frac{q_0}{2}}, \quad (7.73)$$

де $q_0 = E/N_0$.

Для m -позиційних систем справедливо наближене співвідношення

$$P_{\text{онкг}} \approx \frac{m-1}{2} e^{-\frac{q_0}{2}}. \quad (7.74)$$

З порівняння ф-л (7.66) і (7.74) випливає, що ймовірність помилки у багатопозиційних системах $P_{\text{от}}$ приблизно визначається через ймовірність помилки у відповідній двійковій системі P_{02} . Це співвідношення має такий вигляд:

$$P_{\text{от}} \approx (m - 1)P_{02}. \quad (7.75)$$

На рис.7.5 наведено графіки залежності ймовірності помилки в двійковій системі з активною паузою від відношення сигналу до перешкод при когерентному та некогерентному прийомі. Порівняння кривих показує, що оптимальний когерентний прийом несуттєво відрізняється за стійкістю до перешкод від оптимального некогерентного прийому.

7.7. Передача дискретних сигналів по каналах з змінними параметрами

Досі розглядали системи передачі, у яких параметри каналу незмінні. Ймовірність помилки у таких системах обумовлена лише наявністю адитивних перешкод. Оптимальний приймач використовує всі параметри сигналу та забезпечує максимально можливу завадостійкість. У реальних умовах параметри каналу і параметри сигналу можуть випадково змінюватися. Ефективне використання параметрів сигналу на прийомі в цих умовах утруднюється і стійкість до перешкод неминуче погіршується.

У реальних каналах сигнали на вхід приймача можуть приходити різними шляхами з різними згасаннями μ_k та різними запізнюваннями τ_k . Прийнятий сигнал $x(t)$ у цьому випадку можна подати у вигляді суми

$$x(t) = \sum_{k=1}^n \mu_k s(t - \tau_k) + \omega(t), \quad (7.76)$$

де μ_k і τ_k - випадкові процеси, взагалі кажучи, що залежать від часу, $\omega(t)$ - адитивна перешкода.

Канали, що описуються виразом (7.76), називаються багатопроменивими.

У багатьох випадках має місце лише один шлях поширення

$$x(t) = \mu s(t - \tau) + \omega(t). \quad (7.77)$$

Такі канали називаються однопроменивими. Якщо величини μ і τ фіксовані в часі, маємо канал з постійними параметрами. До таких каналів можна віднести канали провідного зв'язку та канали укв при передачі в межах прямої видимості.

Майже всі види радіоканалів відносяться до каналів з параметрами, що випадково змінюються. До цього класу каналів належать короткохвильові лінії зв'язку, в яких завдяки зміні стану іоносфери відбуваються безперервні коливання сигналу амплітуди в точці прийому. Випадкові зміни умов поширення радіохвиль мають місце на ультракоротких хвилях, а також середніх і навіть довгих хвилях.

Будь-які зміни коефіцієнта передачі та часу поширення сигналів викликають флуктуацію сигналів на виході каналів і можуть розглядатися як дія перешкод. Зміни коефіцієнта виявляються у вигляді флуктуацій амплітуди сигналу та називаються мультиплікативною перешкодою. Випадкові затримки променів викликають фазові та часові флуктуації сигналів.

Вплив мультиплікативної перешкоди на сигнал, що передається, можна розглядати як модуляцію цього сигналу випадковим процесом μ . Таку перешкоду можна звести до адитивної:

$$x = \mu s = \mu_0 s + \omega_3, \quad (7.78)$$

де ω_3 - еквівалентна адитивна перешкода, рівна $\omega_3 = s(\mu - \mu_0) = s\xi$.

Еквівалентна перешкода ω_3 є твором випадкового процесу $\xi = \mu - \mu_0$ детерміновану функцію часу s . Це означає що ω_3 - нестационарний процес і всі його розподіли та його моменти залежить від часу. Фактично це означає, що з аналізу необхідно після усереднення по безлічі вдаватися до усереднення за часом. Зауважимо також, що це випадковий процес з ненульовим середнім $\bar{\mu} = \mu_0 \neq 0$ і що μ_0 . Процес $\xi = \mu - \mu_0$ має вже нульовий середній і відповідно зміщений розподіл. Величини μ і ξ безрозмірні, тоді як ω_3 має розмірність сигналу.

Еквівалентна потужність мультиплікативної перешкоди, очевидно, дорівнюватиме її дисперсії:

$$P_{\text{пз}} = \overline{D\omega_3} = \overline{Ds\xi} = \overline{s^2 D\xi} = \overline{s^2} D\mu.$$

Відношення сигналу до еквівалентної перешкоди при цьому дорівнює:

$$q_3 = \frac{\overline{\mu_0 s^2}}{\overline{D\omega_3}} = \frac{\mu_0^2}{D\xi} = \frac{\mu_0^2}{D\mu}, \quad (7.79)$$

тобто, воно визначається лише середнім значенням та дисперсією процесу $\mu(t)$, що характеризує мультиплікативну перешкоду

Поширеним явищем, що призводить до випадкових коливань параметрів сигналу в точці прийому, є багатопроменеве поширення радіохвиль. Багатопроменевість - основна причина завмирання сигналу. Внаслідок різниці ходу променів, що приходять від передавача до приймача, сигнал у приймальній антені являє собою суму окремих коливань з різними фазами та амплітудами. Інтерференція цих коливань в умовах, коли різниці ходу променів не залишаються постійними, і є основною причиною флуктуацій амплітуд, так і фаз складових сигналу. Залежно від ширини спектра сигналу F та властивостей каналу розрізняють загальні (або гладкі) та селективні завмирання, які, у свою чергу, можуть бути швидкими та повільними.

Загальні завмирання мають місце, коли час запізнення променів t . При цьому коефіцієнт передачі каналу μ (або амплітуда сигналу) та фаза сигналу ϕ всім частотних складових змінюються однаково. При великій кількості променів вважатимуться, що випадкова величина ϕ має рівномірну щільність ймовірності на інтервалі від 0 до 2π , а μ - розподіл Релея

$$p(\mu) = \frac{2\mu}{\mu_0^2} e^{-\frac{\mu^2}{\mu_0^2}}, \quad (7.80)$$

де $\mu_0 = \sqrt{\mu^2}$ - Середньоквадратичне значення коефіцієнта передачі μ .

У ряді випадків сигнал у точці прийому складається з двох складових: регулярної з параметрами, що повільно змінюються, і розсіяної швидко флуктуюючої складової. У таких випадках коефіцієнт передачі μ описується узагальненим розподілом Релея.

$$p(\mu) = \frac{2\mu}{\mu_\phi^2} \exp\left(-\frac{\mu^2 + \mu_p^2}{\mu_\phi^2}\right) I_0\left(\frac{2\mu\mu_p}{\mu_\phi^2}\right), \quad (7.81)$$

де μ_p - регулярна складова коефіцієнта передачі, $\overline{\mu_\phi^2} = \mu_0^2 - \mu_p^2$ - середнє значення квадрата флуктуючої складової, $\mu_0 = \sqrt{\overline{\mu^2}}$ - середнє квадратичне значення μ . Фаза сигналу ϕ у цьому випадку розподілена нерівномірно.

У разі сильних завмирань, коли $\overline{\mu_\phi^2} \mu_p^2$, розподіл (7.81) наближається до релєєвського (7.80). Такі завмирання іноді називають релєєвськими. Завмирання, що підкоряються узагальненому закону Релєя, називають квазірелєєвськими.

Для слабких завмирань, коли $\overline{\mu_\phi^2}$, на підставі (7.81) отримуємо розподіл, близький до нормального:

$$p(\mu) \approx \frac{1}{\sqrt{2\pi\overline{\mu_\phi^2}}} \sqrt{\frac{\mu}{\mu_p}} e^{-\frac{(\mu-\mu_p)^2}{2\overline{\mu_\phi^2}}}.$$

Слабкі завмирання тому називають гаусовими завмираннями.

Згідно з експериментальними даними в діапазонах середніх і коротких хвиль релєєвські та квазірелєєвські завмирання зустрічаються приблизно однаково часто. В діапазоні при дальньому іоносферному або тропосферному поширенні переважають релєєвські завмирання, при ближньому поширенні - квазірелєєвські.

Селективні завмирання спостерігаються тоді, коли час запізнення променів можна порівняти з величиною $1/F$. У цьому випадку амплітуди та фази частотних складових сигналу змінюються незалежно один від одного. При швидких завмираннях амплітуди і фази суміжних елементів сигналу не кореліровані між собою, а при повільних завмирання вони змінюються однаково.

У разі повільних замирань, які ми і розглянемо, коефіцієнт передачі каналу μ та фази сигналу ϕ практично не змінюються протягом тривалості кількох елементів сигналу. Аналіз раніше прийнятих елементів сигналу дозволяє з достатньою мірою точності передбачити очікувані параметри наступного елемента. У цих умовах прийом може бути здійснений так само, якби завмирання були відсутні, і оптимальними будуть

схеми, розглянуті в попередніх параграфах, з тією різницею, що в схемах повинні проводитися безперервні регулювання μ (регулювання посилення) та ϕ (регулювання фази) відповідно до очікуваних значень μ і ϕ . У зв'язку з труднощами підстроювання фази ширше застосування знаходять некогерентні методи прийому, у яких відомості про фазу сигналу не використовуються. У каналах, де випадково змінюється лише фаза сигналу, некогерентний прийом є оптимальним.

Формули, що визначають можливість помилки елемента сигналу при повільних завмираннях для даного значення q_0 , залишаються тими самими, що у каналі без завмирань. Але в процесі завмирань величина q_0 змінюється пропорційно μ^2 . Тому для визначення повної ймовірності помилки необхідно усереднити ймовірність $P_0(q_0)$ відповідно до розподілу $p(\mu)$, тобто.

$$\overline{P}_0 = \int_0^\infty p(\mu) P_0(q_0) d\mu = \int_0^\infty p(\mu) P_0\left(\frac{\mu\sqrt{q_0}}{\mu_0}\right) d\mu, \quad (7.82)$$

де враховано, що $q_0 = (\mu^2/\mu_0^2)\overline{q}_0$.

Ймовірність помилки при когерентному прийомі бінарних сигналів в умовах повільних релєвських замирань визначається наступним виразом

$$\overline{P}_0 = \frac{1}{2} \left(1 - \sqrt{\frac{\overline{q}_0}{\frac{2}{1-p_{12}} + \overline{q}_0}} \right). \quad (7.83)$$

При $q_0 \gg 1$ маємо наближену формулу

$$\overline{P}_0 \approx \frac{1}{2q_0(1-p_{12})}. \quad (7.84)$$

При некогерентному прийомі для систем з активною паузою ймовірність помилки при повільних релєвських завмираннях визначається шляхом усереднення (5.71) μ відповідно (5.77)

$$\overline{P_0} = \int_0^\infty \frac{2\mu}{\mu_0^2} e^{-\frac{\mu^2}{\mu_0^2}} \frac{1}{2} e^{-\frac{\mu^2}{\mu_0^2} q_0} d\mu = \frac{1}{2+q_0}. \quad (7.85)$$

Графік залежності (5.81) наведено на рис.7.5. З порівняння кривих видно, що наявність замирань сигналу значно знижує стійкість до перешкод системи зв'язку.

У разі квазірелеївських замирань щільність ймовірності $p(\mu)$ визначається узагальненим розподілом Релея (7.81). Для цього випадку повна ймовірність помилки при некогерентному прийомі бінарних сигналів визначається виразом

$$\overline{P_0} = \frac{c^2+1}{q_0+2c^2+2} \exp\left(-\frac{c^2 \overline{q_0}}{q_0+2c^2+2}\right), \quad (7.86)$$

де $c^2 = \mu_p^2 / \mu_\Phi^2$.

При $c = 0$ цей вислів перетворюється на ф-лу (7.85) для релеївських замирань.

Канал із квазірелеївськими замираннями є проміжним випадком між каналом без замирань і каналом із релеївськими замираннями. Графіки ймовірності помилки при цьому випадку на рис. 7.5 розміщуватимуться між кривими 2 і 3. При малих значеннях з графіки будуть ближчими до кривої 3, а при великих – до кривої 2.

Природним способом усунення мультиплікативної перешкоди, зокрема і замирань, є застосування автоматичного регулювання посилення (АРУ). За відсутності адитивної завади ідеальна система АРУ дозволяє повністю усунути мультиплікативну перешкоду. Дія такої системи має зводитися до множення сигналу x на $1/\mu$, тобто, $z = (1/\mu)x = (1/\mu)\mu s = s$. За наявності адитивної перешкоди $z = (1/\mu)x = (1/\mu)(\mu s + \omega) = s + \omega/\mu$. Тут хоч і виходить сигнал постійної інтенсивності, але з адитивною перешкодою, що флюктує по інтенсивності.

Ефективним заходом боротьби із замираннями є рознесений прийом сигналів. Суть його полягає в тому, що на прийомі передане повідомлення відтворюється не по одному прийнятому сигналу, а по двох або кількох сигналах, що несуть одне й те повідомлення. Цими сигналами можуть бути сигнали кількох передавачів, що працюють на різних частотах (рознесення по частоті), або сигнали

одного і того ж передавача, прийняті на різні антени, рознесені простором або поляризації. Можливе також рознесення за часом (наприклад, повторення передачі).

Обробка кількох сигналів прийомі у випадку зводиться до підсумовування з вагою. В окремих випадках це може бути просте складання прийнятих сигналів або вибір найбільшого з них.

Контрольні питання

1. Дайте визначення перешкодостійкості системи зв'язку.
2. Яка завадостійкість називається потенційною?
3. Який приймач називається оптимальним?
4. Які критерії оптимального прийому Ви знаєте?
5. Дайте визначення критерію ідеального спостерігача В. А. Котельникова.
6. Що таке функція правдоподібності та відношення правдоподібності?
7. Дайте визначення критерію мінімального ризику.
8. Дайте визначення критерію Неймана Пірсона.
9. Запишіть умову оптимального приймача, що працює на порівняння середньоквадратичних помилок та намалюйте його схему.
10. Намалюйте схему оптимального приймача кореляційного.
11. Зобразіть схему оптимального приймача на узгоджених фільтрах.
12. За яких умов схема оптимального приймача В.А.Котельникова перетворюється на схему кореляційного приймача?
13. Що називається мультиплікативною перешкодою? Як таку перешкоду можна звести до еквівалентної адитивної перешкоди?
14. Які види замирань Вам відомі? Поясніть причини появи замирань.
15. У чому різниця між релеївськими та квазірелеєвськими замираннями?
16. Поясніть метод обчислення ймовірності помилки при повільних загальних замираннях.

РОЗДІЛ 8. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІНФОРМАЦІЇ

8.1. Кількісна міра інформації

В теорії інформації вивчаються кількісні закономірності передавання, збереження й оброблення інформації. Зокрема, основна увага приділяється визначенню середньої швидкості передавання інформації та розв'язанню задачі максимізації цієї швидкості застосуванням відповідного кодування. Граничні співвідношення теорії інформації дають можливість оцінити ефективність різних систем зв'язку та встановити умови узгодження в інформаційному відношенні джерела з каналом і каналу зі споживачем.

Для дослідження цих питань з обох позицій необхідно насамперед встановити універсальну кількісну міру інформації, що не залежить від конкретної фізичної природи повідомлень, які передаються.

Статистичні методи визначення кількісної міри інформації. При визначенні одиниці кількості інформації слід мати на увазі, що кількість інформації в повідомленні про деяку подію істотно залежить від імовірності цієї події:

$$I(a) = \log_k \frac{1}{P(a)} = -\log_k P(a). \quad (8.1)$$

Оскільки $0 < P(a) < 1$, то $I(a)$ - величина додатна і скінченна. При $P(a) = 1$ кількість інформації дорівнює нулеві, тобто повідомлення про відому подію ніякої інформації не несе.

Логарифмічна міра у даному випадку має природну властивість адитивності, відповідно до якої кількість інформації, що міститься в декількох незалежних повідомленнях, дорівнює сумі кількості інформації, що міститься в усіх повідомленнях.

Дійсно, якщо загальна ймовірність n незалежних повідомлень $P(a_1 \text{ і } a_2 \text{ і } \dots \text{ і } a_n) = P(a_1)P(a_2) \dots P(a_n)$, то кількість інформації в цих повідомленнях дорівнює:

$$I(a_1 i a_2 i \dots i a_n) = -\log_k P(a_1 i a_2 i \dots i a_n) = -\sum_{i=1}^n \log_k P(a_i) = \quad (8.2)$$

При $k = 2$ кількість інформації виражається в двійкових одиницях (дв. од.):

$$I(a_i) = \log P(a_i).$$

Визначити інформацію можна по-іншому. Розглянемо як повідомлення не окрему букву, а ціле число. Якщо всі букви рівноймовірні і незалежні, то всі слова також рівноймовірні і:

$$P_{сл} = 1/N,$$

де $N = m^n$ - кількість слів.

Тоді можна записати:

$$I_{сл} = -\log P_{сл} = -\log(1/m^n) = n \log m. \quad (8.3)$$

Для двійкового каналу $m = 2$. Тоді $n \log 2 = n$ дв. од.

У загальному випадку при передаванні повідомлень невизначеність зменшується. Так, у каналі з шумами можливі помилки. За прийнятим повідомленням тільки з деякою ймовірністю $P(u/v) < 1$ можна зрозуміти, що було передано повідомлення u при прийнятому повідомленні v . Тому після одержання повідомлення залишається деяка невизначеність, що характеризується величиною апостеріорної ймовірності $P(u/v)$, а кількість інформації, яка міститься в сигналі v , визначається ступенем зменшення невизначеності при його прийманні.

Якщо $P(u)$ - апіорна ймовірність, то кількість інформації в прийнятому повідомленні v щодо переданого повідомлення u дорівнює:

$$I(u, v) = \log \quad (8.4)$$

Цей вираз можна розглядати також як різницю між кількістю інформації, що надійшла від джерела повідомлень, і тією кількістю інформації, що втрачена в каналі внаслідок дії шумів.

Наведена вище кількісна міра інформації стосується тільки статистичної властивості повідомлень, що передаються. Для придбання, збереження, перероблення і передавання повідомлень (інформації, знань) використовуються *знаки* (символи) - матеріальні предмети, явища, події, що виступають як представники деякого іншого предмету, властивості або відносин. Розрізняють два види знаків: *мовні*, що входять у деяку знакову систему (природні або штучні мови), і *немовні*. Прикладами останніх можуть бути відбитки пальців, симптоми захворювань, зображення маски як символу театрального мистецтва і т.д. Наука про властивості знаків називається *семіотикою*. Семіотика вивчає знаки в трьох аспектах. *Синтактика* вивчає знаки і співвідношення між ними (правила побудови складених знаків у рамках знакової системи), наприклад, правила побудови слів з букв, речень зі слів безвідносно від інтерпретації. *Семантика* розглядає співвідношення між знаками та їх інтерпретаціями незалежно від того, хто є інтерпретатором (адресатом). Так, знак у вигляді червоного хреста, розташований на будинку, означає пункт медичної допомоги. *Прагматика* досліджує зв'язок знаків з тими, хто використовує їх, тобто проблеми тлумачення знаку "адресатами", його корисності та цінності для інтерпретатора. Очевидно, корисність і цінність знаку істотно різні для здорової й хворої людини. Всі три аспекти вивчення знаків взаємопов'язані, і кожен наступний містить у собі попередній. Оскільки знак є носієм інформації, то відповідно до наведеної структури семіотики виділяють синтактичний, семантичний і прагматичний аспекти теорії інформації.

Структурні міри інформації. Дискретні повідомлення складаються зі зліченої множини символів, створюваних джерелом інформації. Набір символів, якими можуть бути цифри, букви, імпульси й інші знаки, називають *алфавітом джерела*. Кількість символів у алфавіті визначає обсяг алфавіту. Типовими дискретними повідомленнями є текст, записаний за допомогою якого-небудь алфавіту, та послідовність чисел. Символи можуть мати різні фізичні властивості, які дають змогу однозначно відрізнити їх один від одного. Ці властивості називають *якісними ознаками*. При використанні структурних мір інформації враховується кількість символів, зв'язків між ними або комбінацій з них. *Кількість інформації* в комбінаторній мірі обчислюється як кількість комбінацій символів, що входять до алфавіту. Отже, оцінюється *комбінаторна властивість* потенціальної структурної різноманітності джерела інформації.

Комбінування можливе при двох і більше неоднакових символах, наявності змінних або зв'язків, коли якісною ознакою символу є його положення (позиція) у послідовності. В останньому випадку місце розташування символу впливає на ціле, що має місце, наприклад, у позиційній системі числення. Нехай є двійкові числа 11110 і 01111 або 00001 і 10000. У першому випадку змінює положення нуля, у другому - одиниця. Відповідно число 30 перетворюється в 15, а одиниця - в 16.

У комбінаториці розглядаються різні види сукупності (сполучення), що утворилася з елементів деякої множини M , яка містить n різних елементів.

Види сполучень - розміщення, перестановки, поєднання.

Кількість різних сполучень з n елементів по m різних елементах множини M :

$$N = C_m^n = m! / [n!(m-n)!]$$

Якщо, наприклад, символами є a, b, c, d , то число комбінацій за двома символами дорівнює 6. Це будуть комбінації: ab, ac, ad, bc, bd, cd .

Сполучення з повтореннями також відрізняються складом елементів, але елементи в них можуть повторюватися до n раз. При цьому:

$$N = C_{m(\text{повт})}^n = (m+n-1)! / [n!(m-1)!]$$

При вищевказаних чотирьох символах і $n = 2$ матимемо 10 комбінацій. До наведених раніше шести додадуться комбінації aa, bb, cc, dd .

Перестановки m елементів розрізняються їх порядком у послідовності. Число можливих перестановок m символів $N = m$. Як випливає з наведеного виразу, при визначенні числа перестановок завжди передбачається, що $n = m$. В усіх попередніх і наступних прикладах довжина n комбінацій дорівнюватиме двом, тому для одержання порівнянних результатів покладемо $n = m = 2$. Тоді матимемо тільки дві комбінації: ab і ba .

Розміщення з m символів по n елементів розрізняються складом елементів і їх порядком. Можливе число розміщень з m символів по n :

$$N = C_{m(\text{повт})}^n = m! / (m-n)!$$

При $m = 4$ і $n = 2$ маємо 12 комбінацій. Шість комбінацій ідентичні тим, що були отримані при сполученні символів, у шести інших символи в кожній комбінації міняються місцями (досить їх записати в зворотному порядку).

Можливе число розміщень з повтореннями по n символів з m :

$$N = C_{m(\text{повт})}^n = m^n. \quad (8.5)$$

При $m = 4$ і $n = 2$ маємо 16 комбінацій. До наведених вище 12 комбінацій додаються ще чотири наявності: aa, bb, cc, dd . Вираз (8.5) визначає максимальне число комбінацій, які можна одержати, розміщуючи m різних символів по n позиціях.

При застосуванні комбінаторної міри кількість інформації збігається з числом можливих з'єднань. Отже, визначення кількості інформації в комбінаторній мірі полягає в наявній кількості можливих або справді здійснених комбінацій, тобто в оцінці *структурної різноманітності*. Внаслідок показникового закону залежності числа можливих з'єднань від n - довжини послідовності символів у з'єднанні (8.5) - число можливих комбінацій не є зручною мірою кількості інформації У статті, яку було опубліковано в 1928 р., Р. Хартлі запропонував за міру кількості інформації двійковий логарифм числа можливих комбінацій символів:

$$I = \log_2 N = \log_2 m^n = n \log m. \quad (8.6)$$

З (8.6) випливає, що кількість інформації пропорційна довжині n послідовності символів у комбінації. Логічно за одиницю взяти кількість інформації, що припадає на одну позицію.

Первинним і неподільним елементом інформації є *елементарна двійкова подія* A - вибір з твердження або заперечення, наявності або відсутності якого-небудь явища, істини і неправди і т.д. Двійковість події дає можливість зобразити її умовно в геометричній символіці точкою і пробілом, в арифметичній - одиницею і нулем, у сигнальній - імпульсом і паузою (відсутністю імпульсу). Тому дуже часто інформацію подають послідовністю, складеною тільки з двох символів (0 і 1), тобто $m = 2$. Якщо тепер у (8.6) покласти $n = 1, m = 2$, то кількість інформації $I = 1 \cdot \log_2 2 = 1$ біт.

Одиниця кількості інформації, отримана при зазначених вище умовах, називається *двійковою* або одиницею *bit*. Введене визначення кількості інформації еквівалентне числу двійкових знаків 0 і 1 при поданні повідомлень числами двійкової системи числення. Одному біту відповідає один двійковий розряд.

Особливості семантичної оцінки інформації. Якщо статистична теорія інформації будується на базі поняття математичної ймовірності, то семантична теорія інформації Р. Карнапа і Бар-Хілела ґрунтується на семантичному визначенні ймовірності, що називається *логічною*, або *індуктивною ймовірністю*. Під терміном *логічна, або індуктивна ймовірність* Р. Карнап розуміє вид ймовірності, який мається на увазі щоразу, коли роблять індуктивний висновок, що не випливає з логічною необхідністю (з ймовірністю "одиниця") зі свідчень, істинність яких гарантується. При цьому вважатимемо, що стосовно даних свідчень висновок має деякий ступінь ймовірності. Висновок між істиною і неправдою називається у ймовірній логіці гіпотезою, тобто припущенням, здогадом.

Коли Ваш співрозмовник-філателіст стверджує, що цього року, ймовірно, буде випущена поштова марка, присвячена Т.Г. Шевченку, то він не виходить за межі статистичної ймовірності і виражає свою впевненість у цьому. Така логічна ймовірність поширюється як на користь висловленої гіпотези, так і на спростування її. Нехай такими взаємними виключеннями є висловлення: A - Т.Г. Шевченко - видатний український поет, або ні - $\bar{A}$; B - цього року буде відзначатися 150-річчя з дня народження поета, або ні - $\bar{B}$; C - раніше випускалися марки, присвячені поетам, або ні - $\bar{C}$. Логічна ймовірність - це поняття, застосовне до пар тверджень - гіпотези (буде випущена поштова марка, присвячена Т.Г. Шевченку) і свідчення, яке виражає щось, що має відношення до гіпотези.

Нехай обговорюється можливість випуску зазначеної поштової марки. Потрібно оцінити апіорну логічну ймовірність того, що така марка буде випущена. Спочатку припустимо, що коли висловлюється гіпотеза H - що буде випущена поштова марка, присвячена Т.Г. Шевченку, немає ніяких свідчень щодо висловленої гіпотези. Тоді, відповідно до *принципу байдужості* (однакової ймовірності), передбачається, що з ймовірністю 0,5 марка буде або не буде випущена. Перед тим, як висловити наступну гіпотезу, мабуть, варто оцінити ймовірність того, що марка буде випущена залежно від співвідношення свідчень на користь висловлюваної гіпотези і проти неї. Дійсно, якщо першим свідченням є A - висловлення на користь гіпотези H , то логічно вважати гіпотезу H більш ймовірною. Аналогічно можна висловитись, якщо першим свідченням було $\bar{A}$ - висловлення на користь гіпотези $\bar{H}$. В обох варіантах першого свідчення логічна ймовірність гіпотези, що суперечить висловленому свідченню, оцінюватиметься меншою величиною (табл. 8.1).

Як же встановити числове значення ймовірності одних висловлень щодо інших? Однозначної відповіді на це питання поки немає. Ясно одне - ступінь ймовірності гіпотези залежить від стану накопичених свідчень - знань. Розглянемо спосіб присвоювання величини логічної ймовірності, запропонований Карнапом. Щоразу, коли змінюється кількість свідчень, переоцінюються присвоєні послідовності апріорних ймовірностей за таким правилом. Усі можливі послідовності з деякою кількістю розглянутих висловлень поєднуються в групи у межах кожної зібрані комбінації з однаковим відношенням висловлювань одного виду до висловлювань іншого виду, кожній групі присвоюється та сама ймовірність. Потім ймовірності груп поділяються порівну між окремими послідовностями, тобто визначаються апріорні логічні ймовірності кожної можливої послідовності.

З таким методом присвоювання апріорних ймовірностей можна враховувати накопичений досвід. Як впливає з табл. 8.1 зі збільшенням у послідовності відносного числа висловлювань логічні ймовірності гіпотези H зростають. Припустимо, що три свідчення дали таку послідовність: Т.Г. Шевченко - видатний український поет, у цьому році буде відзначатися 150-річчя з дня народження поета, раніше не випускалися марки, присвячені поетам (варіанти 3 і 10 табл. 8.1). Подібна послідовність показує перевагу випуску поштової марки, присвяченої Т.Г. Шевченку, над тим, що такої марки не буде. Дійсно, ймовірність того, що марка, присвячена Т.Г. Шевченку, буде випущена, дорівнює $3/60$ проти $2/60$ для гіпотези $\bar{H}$, тобто становить $3/2$. У цьому варіанті розподілу апріорних ймовірностей не враховується "вага" свідчення, тобто припускається, наприклад, що висловлювання: Т.Г. Шевченко - видатний український поет, цього року буде відзначатися 150-річчя з дня його народження - рівноцінні стосовно гіпотези.

Таблиця 8.1. Априорна ймовірність залежно від кількості свідчень

Номер варіанту	Ймовірна послідовність	Априорна присвоєна ймовірність	
		групі	послідовності
Кількість свідчень - 0			
1	H	1/2	1/2
2	$\overline{H}$	1/2	1/2
Кількість свідчень - 1			
1	AH	1/3	2/6
2	$A\overline{H}$	1/3	1/6
3	$\overline{A}H$	1/3	1/6
4	$\overline{A}\overline{H}$	1/3	2/6
Кількість свідчень – 2			
1	ABH	1/3	6/18
2	$AB\overline{H}$	1/3	1/18

Номер варіанту	Ймовірна послідовність	Апріорна присвоєна ймовірність	
		групі	послідовності
3	$\overline{ABH}$	1/3	1/18
4	$\overline{ABH}$	1/3	1/18
5	$\overline{ABH}$	1/3	1/18
6	$\overline{ABH}$	1/3	1/18
7	$\overline{ABH}$	1/3	1/18
8	$\overline{ABH}$	1/3	6/18
Кількість свідчень - 3			
1	$ABCH$	1/5	12/60
2	$ABCH$	1/5	3/60
3	$ABCH$	1/5	3/60
4	$ABCH$	1/5	3/60
5	$ABCH$	1/5	3/60
6	$ABCH$	1/5	2/60

7	$\overline{ABCH}$	1/5	2/60
8	$\overline{ABCH}$	1/5	2/60
9	$\overline{ABCH}$	1/5	2/60
10	$\overline{ABCH}$	1/5	2/60
11	$\overline{ABCH}$	1/5	2/60
12	$\overline{ABCH}$	1/5	3/60
13	$\overline{ABCH}$	1/5	3/60
14	$\overline{ABCH}$	1/5	3/60
15	$\overline{ABCH}$	1/5	3/60
16	$\overline{ABCH}$	1/5	12/60

Розглянута теорія встановлює міру кількості семантичної інформації, що міститься в простих твердженнях або пропозиціях. Кількість семантичної інформації $cont(i)$, що міститься в якому-небудь елементарному твердженні i (гіпотезі i), пропонується розглядати як функцію логічної ймовірності $L(i)$. Позначення $cont$ є скороченням англійського слова *content* - зміст, суть. Семантична інформація, як і в статистичній теорії інформації, розглядається тут як знята, усунута невизначеність при переході від відносного знання (гіпотези) до достовірного знання. Якщо i й j - два висловлення, і $L(i) > L(j)$, то слід вважати, що гіпотеза i містить меншу кількість інформації, ніж гіпотеза j , тобто $cont(i) < cont(j)$.

Чим більша логічна ймовірність гіпотези "буде випущена поштова марка, присвячена Т.Г. Шевченку", тим менша кількість семантичної інформації міститься в ній щодо достовірного знання "поштова марка, присвячена Т.Г. Шевченку, випущена". Дійсно, якщо Ваше припущення про ймовірний випуск марки підтверджується, то усунута при цьому невизначеність невелика, і кількість інформації мала. Якщо марка не вийде, логічна ймовірність такого результату передбачалася малою, результат для філателістів сенсаційний, і вони мають велику кількість інформації.

Бажано, щоб у випадку, коли i та j логічно незалежні, мало місце співвідношення $cont(i \cap j) = cont(i) + cont(j)$.

Прийняте тлумачення логічної незалежності можна показати на прикладі висловлення: книга художня і книга (та сама) не українська. Ці висловлення незалежні, але не виключають одне одного.

Вищевказані умови аналогічні прийнятим при розгляді статистичної теорії інформації, тому кількість семантичної інформації $cont(i) = \log[1/L(i)]$. Цей вираз формально подібний розглянутому в статистичній теорії, але він базується на логічній імовірності гіпотези, що дає змогу наблизитися до оцінки змісту інформації.

В іншому варіанті оцінки кількості семантичної інформації використовується поняття *тезауруса* (від грецького "скарбниця"), під яким розуміється запас знань або словник, використовуваний приймачем інформації. Відомості, отримані під час приймання повідомлення, змінюють вихідний тезаурус. Величина цієї зміни береться за характеристику кількості семантичної інформації, яка міститься в даному повідомленні щодо даного приймача. Недостатньо розвинений тезаурус одержує нульову або дуже малу інформацію з даного повідомлення (не може його зрозуміти). Але й надто повний тезаурус також не може одержати багато інформації, оскільки вона не буде для нього новою. Кількість інформації, яку одержує приймач з деякого повідомлення, залежить від кількості наявної в тезаурусі (приймачі) інформації I_T (рис. 8.1, г).

Цінність інформації. При розв'язанні задачі визначення цінності інформації зазвичай розглядають таку ситуацію. Хтось прагне до поставленої мети, маючи можливість наблизитися до неї, затративши певні зусилля. Завдяки одержуваним ззовні відомостям він може точніше наблизитися до мети або заощадити витрати на її досягнення. Цінність отриманої інформації вимірюється ступенем наближення до мети або величиною економії на витратах.

Якщо інформація використовується для управління досягненням мети, то її цінність можна виразити через зміну ймовірності її досягнення. Нехай до одержання інформації ймовірність досягнення мети дорівнювала P_0 , а після її одержання - P_1 , тоді за А.А. Харкевичем цінність інформації визначається як:

$$I = \log_2 P_1 - \log_2 P_0 = \log_2 (P_1 / P_0). \quad (8.7)$$

Тут ймовірність обчислюють як відношення числа сприятливих результатів до їх загального числа.

Розглянемо випадок, наведений на рис. 8.1, а,б. Пункти **A**, **B** і **C** з'єднані дорогами. Мандрівник знаходиться у точці **A**, точка **C** - мета подорожі. Перед ним два шляхи. Який з них веде до **C**, він не знає. Не маючи інформації для вибору, він з ймовірністю 0,5 може бути в **B** або **C**, тобто ймовірність досягнення мети до одержання інформації $P_0 = 0.5$. Нехай за отриманою інформацією $I = \log[1/P_0] = 1$ біт (наприклад, хтось сказав, що треба йти по лівій дорозі) він переходить у пункт **B**. Яка цінність отриманої інформації? Вона залежить від ймовірності досягнення мети з пункту **B**. Визначимо цінності отриманої інформації для трьох випадків:

1. Число хороших випадків дорівнює одному з двох можливих: $P_1 = 0.5$ (рис.8.1,а). Отже, $\ddot{O} = \log_2 [P_1 / P_0] = \log_2 1 = 0$. Тут ймовірність досягнення мети після одержання інформації не змінилася і, як наслідок, її цінність дорівнює нулеві.

2. Є один успішний результат із трьох: $P_1 = 1/3$ (рис.8.1,б) і $\ddot{O} = \log_2 [P_1 / P_0] = \log_2 (2/3) = -0.58$. Враховуючи, що мандрівник був посланий у напрямку, протилежному меті, негативна цінність дає змогу назвати таку пораду *дезінформацією*. Деінформація, збільшуючи вихідну невизначеність (від вибору одного з двох шляхів до одного з трьох), зменшує ймовірність досягнення мети і має негативну цінність.

3. Є два успішних результати з трьох: $P_1 = 2/3$ (рис.8.1,в) і $\ddot{O} = \log_2 [P_1 / P_0] = \log_2 (4/3) = 0.42$.

Найбільшу цінність має порада "йдіть по правій дорозі" ($I = 1$ біт), що направляє мандрівника з A до C , тоді $P_1 = 1$ і $\ddot{O} = \log_2 [P_1/P_0] = \log_2 2 = 1$. Наведений приклад показує, що цінність тієї самої кількості інформації може змінюватися в широких межах.

Інший підхід до визначення цінності інформації пов'язаний з працями Р.Л. Стратоновича. Розглянемо систему (рис.8.1,д), в якій деякий спостерігач робить пошук, здійснює вибір між гіпотезами або оцінює невідому величину. Підсумок своєї діяльності він подає в блок, де оцінюється результат і призначається штраф. Розмір штрафу чи втрат обчислюється відповідно до функції штрафу за величиною помилки, зробленої спостерігачем. *Функцією штрафу* називають залежність між розміром штрафу і величиною помилки, за яку він призначається. Якщо спостерігач нічого не знає про об'єкт і діє, наприклад, або наважання, або методом спроб і помилок, йому можна допомогти, вказавши, як звузити область пошуку. Завдяки обраному методу він буде діяти більш ефективно. Це проявиться в тому, що середнє значення його штрафів зменшиться.

Вказівки дає джерело інформації, їх повідомляють спостерігачу. Вважається, що кількість інформації, яка міститься в кожному отриманому спостерігачем повідомленні, відома. До одержання інформації спостерігач діє в умовах невизначеності, що оцінюється ентропією H і має суму втрат $R(H)$. Отримана і використана інформація спричинює нову, меншу, невизначеність H_1 і нові втрати $R(H_1)$. Різниця втрат $R(H) - R(H_1) = \Delta R$ кількісно показує користь, принесену інформацією чи *кількісну міру цінності інформації*. Отже, цінність інформації можна визначити за розміром штрафу.

Розмір втрат (штрафу) залежить не тільки від корисності інформації. Спостерігач може поводитися "нерозумно" і не користуватися отриманою цінною інформацією. Тому передбачається, що спостерігач отриману інформацію використовує для себе за допомогою найкращих доступних йому методів. Потрібна така функція штрафів, щоб максимально зменшити втрати, які відбуваються внаслідок надходження інформації.

8.2. Ентропія джерела дискретних повідомлень

У багатьох випадках, коли потрібно узгодити канал із джерелом повідомлень, знати кількість інформації, що міститься в окремих повідомленнях, недостатньо. Виникає потреба у характеристиках, які б давали можливість оцінювати інформаційні

властивості джерела повідомлень у цілому. Однією з таких важливих характеристик є середня кількість інформації одного повідомлення.

У найпростішому випадку, коли всі повідомлення рівноймовірні: $P(a_i) = 1/m$, середня кількість інформації дорівнює $\log m$. Отже, при рівноймовірних незалежних повідомленнях інформаційні властивості джерела залежать тільки від числа повідомлень в ансамблі m . Однак у реальних умовах повідомлення, як правило, мають різну ймовірність. Так, наприклад, одні букви алфавіту зустрічаються частіше, інші - рідше.

Тому крім знання числа повідомлень m в ансамблі треба мати відомості про ймовірності кожного повідомлення $P = (a_i), i = \overline{1, m}$.

Оскільки ймовірності повідомлень неоднакові вони несуть різну кількість інформації. Менш імовірні повідомлення несуть більшу кількість інформації і навпаки.

Середня кількість інформації одного повідомлення (біт/повідомлення) визначається як математичне сподівання:

$$H(a) = \overline{I(a_i)} = \sum_{i=1}^m P(a_i) I(a_i) = - \sum_{i=1}^m P(a_i) \log P(a_i). \quad (8.8)$$

Величина $H(a)$ називається *ентропією*. Цей термін запозичений з термодинаміки, де є аналогічний за своєю формою вираз, який характеризує невизначеність стану фізичної системи.

В теорії інформації ентропія $H(a)$ також характеризується невизначеністю ситуації до передавання повідомлення, оскільки заздалегідь невідомо, яке саме повідомлення з ансамблю повідомлень джерела буде передано. Чим більша ентропія, тим сильніша невизначеність, і тим більшу інформацію несе одне повідомлення джерела.

Властивості ентропії джерела незалежних дискретних повідомлень. Ентропія як міра невизначеності скінченної системи подій (повідомлень) має такі основні властивості.

1. У тому разі, коли події (повідомлення) є рівноймовірними, зі збільшенням числа можливих подій m , що складають скінченну систему подій (повну групу), невизначеність цієї системи подій збільшується.

Коли всі повідомлення рівноймовірні $[P(a_i) = 1/m]$, ентропія системи дорівнює:

$$H(a) = -\sum_{i=1}^m P(a_i) \log P(a_i) = -m \frac{1}{m} \log(1/m) = \log m$$

і збільшується за логарифмічним законом.

2. Якщо ймовірність якого-небудь повідомлення дорівнює одиниці $[P(a_i) = 1]$, то ентропія скінченної системи дорівнює нулеві.

Враховуючи, що:

$$\sum_{i=1}^m P(a_i) = 1,$$

одержуємо:

$$H(a) = -[1 \cdot \log 1 + (m+1) \cdot 0 \cdot \log 0].$$

Очевидно, що $1 \cdot \log 1 = 0$, оскільки $\log 1 = 0$. Добуток виду $0 \cdot \log 0$ є невизначеністю типу $0 \cdot (\infty)$, тому що $\log 0 = -\infty$. Ця невизначеність дорівнює нулеві:

$$\lim_{x \rightarrow 0} x \log x = \lim_{x \rightarrow 0} \frac{\log x}{x^{-1}} = \lim_{x \rightarrow 0} \frac{\frac{d(\log x)}{dx}}{\frac{d(x^{-1})}{dx}} = \lim_{x \rightarrow 0} \frac{\frac{d(c \ln x)}{dx}}{x^{-2}} = \lim_{x \rightarrow 0} cx = 0.$$

Отже, при $P(a_i) = 0$ ентропія системи повідомлень $H(a)$ дорівнює нулеві.

3. Невизначеність скінченної системи повідомлень при даному m максимальна при рівноймовірних повідомленнях. Доведення цієї властивості в загальному виді є досить складним. Тому обмежимося прикладом, коли загальне число $m = 2$. При цьому:

$$\sum_{i=1}^2 P(a_i) = 1; \quad P(a_1) = P; \quad P(a_2) = 1 - P;$$

$$H(a) = -P \log P - (1 - P) \log(1 - P).$$

Ентропія скінченної системи, що складається з двох подій, досягає свого максимального значення, яке дорівнює одиниці, при $P = 0.5$, тобто при однаковій імовірності подій, оскільки при цьому $(1 - P) = 1 - 0.5 = 0.5$.

Залежність $H(a)$ від P показана на рис.8.2. Максимум має місце при $P = 0.5$, тобто коли ситуація є найбільш невизначеною. При $P = 0$ або $P = 1$, що відповідає передачі одного з повідомлень a_1 чи a_2 , невизначеність відсутня. В цих випадках ентропія $H(a)$ дорівнює нулю.

4. Ентропія має властивість адаптивності, тобто невизначеність сполученої системи подій дорівнює сумі ентропій незалежних скінченних систем подій, що складають сполучену систему.

Розглянемо такі дві незалежні скінченні системи подій:

$$a = \begin{pmatrix} a_1, a_2, \dots, a_n \\ P_1, P_2, \dots, P_n \end{pmatrix}; \quad b = \begin{pmatrix} b_1, b_2, \dots, b_m \\ q_1, q_2, \dots, q_m \end{pmatrix},$$

де P_i, q_i - імовірності подій a_i і b_i відповідно.

Ентропія цих скінченних систем подій відповідно:

$$H(a) = -\sum_{i=1}^n P_i \log P_i; \quad H(b) = -\sum_{i=1}^m q_i \log q_i.$$

(8.9)

1.

$$\sum_{i=1}^n \sum_{j=1}^n$$

$l=1$

(8.10)

що і потрібно було довести.

Ентропія джерела залежних дискретних повідомлень. Джерела незалежних повідомлень є найпростішим типом джерел. У реальних умовах все значно ускладнюється через наявність статистичних зв'язків між повідомленнями.

Статистичний зв'язок очікуваного повідомлення з попереднім повідомленням кількісно оцінюється спільною $P(a_k, a_i)$ або умовною ймовірністю $P(a_i / a_k)$.

Кількість інформації, що міститься в повідомленні a_i за умови, що відомо попереднє повідомлення a_k визначається виразом:

$$I(a_i / a_k) = -\log P(a_i / a_k).$$

Середня кількість інформації визначається умовною ентропією, яка обчислюється як математичне сподівання:

$$H_2(a) = H(a_i / a_k) = -\sum_{i=1}^n \sum_{k=1}^m P(a_i, a_k) \log P(a_i / a_k).$$

При цьому, коли зв'язок поширюється на три повідомлення, одержимо:

$$H_3(a_i / a_k, a_l) = -\sum_{i=1}^n \sum_{k=1}^m \sum_{l=1}^m P(a_i, a_k, a_l) \log P(a_i / a_k, a_l).$$

У загальному випадку для n залежних повідомлень:

Враховуючи, що при рівноймовірних повідомленнях:

$$H_0(a) = \log m = H_{\max}, \quad (8.11)$$

а при нерівноймовірних, але незалежних:

$$H_1(a) = -\sum_{i=1}^m P_i \log P_i, \quad (8.12)$$

запишемо значення ентропії в порядку її зменшення при збільшенні статистичних зв'язків між повідомленнями, включаючи також формули (8.11) і (8.12), якщо повідомлення незалежні, так:

$$H_0(a) = \log m = H_{\max};$$

$$H_1(a) = -\sum_{i=1}^m P(a_i) \log P(a_i);$$

$$H_2(a) = H(a_i / a_k) = -\sum_{i=1}^n \sum_{k=1}^m P(a_i, a_k) \log P(a_i / a_k);$$

$$H_3(a) = -\sum_{i=1}^n \sum_{k=1}^m \sum_{l=1}^m P(a_i, a_k, a_l) \log P(a_i / a_k, a_l);$$

.....

$$H_n(a) = -\sum_{i=1}^n \cdots \sum_{k=1}^m \sum_{l=1}^m P(a_i, a_k, \dots, a_l) \log P(a_i / a_k, \dots, a_l).$$

Надлишковість джерела повідомлення. Важливою властивістю умовної ентропії джерела залежних повідомлень є те, що при незмінній кількості повідомлень в ансамблі джерела його ентропія зменшується зі збільшенням частоти повідомлень, між якими існує статистичний взаємозв'язок. Відповідно до цієї властивості наявність статистичних зв'язків між повідомленнями завжди спричинює зменшення кількості інформацій одного повідомлення.

Зменшення ентропії джерела зі збільшенням статистичного взаємозв'язку можна розглядати як зниження інформаційної ємності повідомлень. Те саме повідомлення за наявності взаємозв'язку містить в середньому менше інформації, ніж за його відсутності. Інакше кажучи, якщо джерело створює послідовність повідомлень,

які мають статистичний зв'язок, і характер цього зв'язку відомий, то частина повідомлень, що видається джерелом, є надлишковою, тобто її можна відновити за відомими статистичними зв'язками. З'являється можливість передавання повідомлень у скороченому вигляді без втрати інформації, що міститься в них. Наприклад, передаючи телеграму, ми виключаємо з тексту сполучники, розділові знаки, оскільки вони легко відновлюються при читанні телеграми за відомими правилами побудови слів і фраз.

Отже, будь-яке джерело має надлишок.

Кількісне визначення надлишковості можна отримати з наступних міркувань.

Щоб передати кількість інформації I , джерело без надлишковості має видати в середньому повідомлень

$$k_0 = I / H_0,$$

а з надлишковістю

$$k_n = I / H_n.$$

Оскільки $H_0 > H_n$, то $k_n > k_0$, тобто джерело з надлишковістю для передавання тієї самої кількості інформації має використовувати більшу кількість повідомлень.

Надлишкова кількість повідомлень:

$$\Delta k = k_n - k_0,$$

а надлишок джерела:

$$\chi = \frac{k_n - k_0}{k_n} = \frac{\Delta k}{k_n} = 1 - H_n / H_0. \quad (8.13)$$

Величина надлишку лежить у межах $0 \leq \chi < 1$ і є неспадною функцією n

Для російської мови:

$$H_0 = 5 \text{ дв.од./літеру};$$

$$H_1 = 0.5 \text{ дв.од./літеру};$$

$$H_2 = 3.52 \text{ дв.од./літеру};$$

.....

$$H_8 = 2 \text{ дв.од./літеру}.$$

Звідси надлишок російської мови має порядок 60%.

Коефіцієнт

$$r = H_n / H_0 \quad (8.14)$$

називається *коефіцієнтом стискання*. Він показує, до якої величини можна стиснути повідомлення, що передаються, якщо усунути надлишок.

Джерело, що має надлишок, передає зайву кількість повідомлень. Це збільшує тривалість передавання і знижує ефективність використання каналу зв'язку. Стискання повідомлень можна здійснити за допомогою відповідного кодування.

Але надлишок джерела не завжди є негативною властивістю. Наявність взаємозв'язку між літерами текстів дає можливість відновлювати його при спотворенні окремих літер, тобто використовувати для підвищення вірогідності.

Статистичні властивості джерел повідомлення. Використання ентропії в якості усередненої величини, що кількісно характеризує інформаційні властивості джерела, яке видає послідовності дискретних повідомлень, являється доцільним при умові, що імовірнісні співвідношення для цих послідовностей зберігаються незмінними. Джерело називають *стаціонарним*, коли розподіл імовірностей повідомлень не залежить від їх місця в послідовності повідомлень, що створюються цим джерелом, тобто:

$$P(a_1 / a_1, \dots, a_k) = P(a_{l+n} / a_{i+n}, \dots, a_{k+n}), \quad (8.15)$$

де n - будь-яке ціле число.

За аналогією зі стаціонарним випадковим процесом статистичні характеристики послідовності повідомлень стаціонарного джерела не залежать від вибору початку відліку.

Серед стаціонарних джерел повідомлень важливе місце займають *ергодичні* джерела, які відрізняються тим, що з імовірністю, близькою до одиниці, будь-яка достатньо довга послідовність повідомлень такого джерела повністю характеризує його статистичні властивості. Важливою особливістю ергодичних джерел є те, що статистичний взаємозв'язок між повідомленнями завжди розповсюджується тільки на скінченне число попередніх повідомлень.

Існують стаціонарні джерела, які можуть працювати в різних режимах, що відрізняються один від одного своїми статистичними характеристиками. В цьому випадку джерело не є ергодичним, так як при роботі в одному режимі навіть тривала послідовність повідомлень вже не може в цілому характеризувати властивості джерела.

Умовна ентропія стаціонарного джерела знаходиться як результат усереднення в усіх режимах роботи:

$$H_n(a) = \sum_j P_j H_{nj}(a), \quad (8.16)$$

де P_j і $H_{nj}(a)$ - імовірність й умовна ентропія j -го режиму роботи.

Розглянемо умовну ентропію при заданій послідовності попередніх повідомлень:

$$H_n(a/h_a) = - \sum_{l=1}^m P(a_l/h_a) \log(a_l/h_a). \quad (8.17)$$

Тут символом h_α позначена послідовність $n-1$ попередніх повідомлень $a_1, \dots, a_k$, причому індексом α нумеруються з'єднання із m повідомлень по $n-1$, тобто всього $N = m^{n-1}$ послідовностей h_α .

Послідовність h_α можна трактувати як стан джерела, в якому воно знаходиться після її передачі. Подібні випадкові послідовності (які володіють ергодичними властивостями) відомі в математиці як дискретні ланки А.А. Маркова.

В марковському ергодичному джерелі ймовірність передачі того чи іншого повідомлення однозначно визначається станом джерела. Після передачі повідомлення джерело переходить в новий стан, який залежить від попереднього стану і переданого повідомлення.

Достатньо довгі ергодичні послідовності повідомлень, з високим ступенем імовірності, які містять всі відомості про статистичні характеристики джерела, називаються *типовими*. Чим довша послідовність, тим більша ймовірність того, що вона являється типовою. В *типових послідовностях частота появи окремих повідомлень або груп повідомлень як зазвичай мало відрізняється від їх імовірності*. Звідси витікає важлива властивість типових послідовностей, яка полягає в тому, що *типові послідовності однакової довжини приблизно рівноймовірні*. Це легко показати для послідовностей незалежних повідомлень.

Припустимо, що ансамбль повідомлень складається з m повідомлень $a_1, a_2, \dots, a_m$, і нас цікавить ймовірність того, що в послідовності довжиною в n повідомлень число повідомлень a_1 рівне k_1 , число повідомлень a_2 рівне k_2 і т.д., a_m відповідає k_m , причому $k_1 + k_2 + \dots + k_m = n$. При незалежних повідомлення ця ймовірність, очевидно, рівна:

$$P_1^{k_1} P_2^{k_2} \dots P_m^{k_m}, \quad (8.18)$$

де P_i - ймовірність повідомлення a_i . Так як у всіх типових послідовностях за визначенням виконується умова $k_1 \approx nP_1$, то ймовірності типових послідовностей приблизно однакові й рівні:

$$P_n = (P_1^{k_1} P_2^{k_2} \dots P_m^{k_m})^n. \quad (8.19)$$

В цьому випадку кількість інформації в будь-якій з типових послідовностей:

$$J_n \approx \log \frac{1}{P_n}. \quad (8.20)$$

З другої сторони, величину J_n можна виразити через ентропію джерела $H(a)$:

$$J_n \approx nH(a). \quad (8.21)$$

Використовуючи вирази (8.20) і (8.21), ентропію джерела можна визначити як:

$$H(a) \approx \frac{1}{n} \log \frac{1}{P_n}. \quad (8.22)$$

Для загального випадку, в тому числі і для залежних повідомлень, в теорії інформації доводиться наступна теорема. Всі ергодичні послідовності, що містять достатньо велике число повідомлень n , можуть бути розбиті на дві групи:

- типові послідовності з імовірностями P_n , для яких задовольняється нерівність:

$$\left| H(a) - \frac{1}{n} \log \frac{1}{P_n} \right| < \varepsilon, \quad (8.23)$$

де $H(a)$ - ентропія джерела і ε - як завгодно мала величина;

- нетипові послідовності, сумарна імовірність яких δ як завгодно мала.

Величини ε і δ необмежено зменшуються з ростом n , що дозволяє завжди вибрати таке значення n , при якому всі послідовності джерела, за винятком дуже малої їх частини, можуть бути віднесені до рівноймовірних типових послідовностей.

Важливим наслідком теореми є можливість встановлення залежності між числом варіантів всіх можливих типових послідовностей M_n і ентропією джерела. Для достатньо довгих послідовностей величини ε і δ малі. Тоді на основі (8.23):

$$M_n \approx \frac{1}{P_n} \approx 2^{nH(a)}, \quad (8.24)$$

Що стосується нетипових послідовностей, то внаслідок їх малої імовірності при великому n вони в багатьох випадках взагалі не враховуються.

8.3. Взаємна інформація

Ентропія ансамблю характеризує середню кількість повної інформації, що міститься в повідомленні. Визначимо тепер інформацію, що міститься в одному ансамблі відносно другого; наприклад, в прийнятому сигналі відносно переданого повідомлення. Для цього розглянемо об'єднання двох дискретних ансамблів, A і B , взагалі кажучи, залежних. Його можна інтерпретувати як пару ансамблів повідомлень, або як ансамблі повідомлення і сигналу, за допомогою якого повідомлення передається, або як ансамблі сигналів на вході і виході каналу зв'язку і так далі. Нехай $P(a_k, b_l)$ - спільна ймовірність реалізацій a_k і b_l . Спільною ентропією ансамблів A і B будемо називати:

$$H(A, B) = M \left\{ \log \frac{1}{P(a_k, b_l)} \right\}. \quad (8.25)$$

Введемо також поняття умовної ентропії:

$$H(A / B) = M \left\{ \log \frac{1}{P(a_k / b_l)} \right\}, \quad (8.26)$$

де $P(a_k | b_l)$ - умовна ймовірність a_k , якщо має місце b_l ; тут математичні очікування беруться по об'єднаному ансамблю AB . Зокрема, для джерел без пам'яті:

$$H(A / B) = \sum_k \sum_l P(a_k, b_l) \log \frac{1}{P(a_k / b_l)}. \quad (8.27)$$

Із теорети множення ймовірностей
 $P(a, b) = P(a)P(b | a) = P(b)P(a | b)$ випливає, що

$$H(A, B) = H(A) + H(B / A) = H(B) + H(A / B). \quad (8.28)$$

Для умовної ентропії справедлива подвійна нерівність:

$$0 \leq H(A/B) \leq H(A). \quad (8.29)$$

При цьому рівність $H(A|B) = 0$, як видно з (8.26), має місце в тому випадку, коли при кожному значенні b_l умовна імовірність однієї реалізації $P(a_k | b_l) = 1$, а для решти реалізацій $P(a_k | b_l) = 0$. Це означає, що, знаючи реалізацію B , можна точно встановити реалізацію A . Іншими словами, B містить повну інформацію про A .

Другий крайній випадок коли $H(A|B) = H(A)$ має місце, якщо $P(a_k | b_l) = P(a_k)$ при всіх a і b . Остання рівність означає, що події A і B незалежні. В цьому випадку знання реалізації B не зменшує невизначеності A , тобто B не містить ніякої інформації про A .

В загальному випадку ентропія $H(A|B)$ менша від безумовної $H(A)$ і знання реалізації B знижує в середньому початкову невизначеність A . Природно назвати різницю $H(A) - H(A|B)$ кількістю інформації, що міститься в B відносно A . Її називають також взаємною інформацією між A і B та позначають $I(A, B)$:

$$I(A, B) = H(A) - H(A|B). \quad (8.30)$$

Підставляючи в (8.30) (8.26) і $H(A) = M \left\{ \log \frac{1}{P(a)} \right\}$ виразимо взаємну інформацію через розподіл імовірностей:

$$I(A, B) = M \left\{ \log \frac{1}{P(a_k)} \right\} - M \left\{ \log \frac{1}{P(a_k | b_l)} \right\} = M \left\{ \log \frac{P(a_k | b_l)}{P(a_k)} \right\}. \quad (8.31)$$

Якщо скористатися теоремою множення $P(a_k, b_l) = P(b_l)P(a_k | b_l)$, то можна записати $I(A, B)$ в симетричній формі:

$$I(A, B) = M \left\{ \log \frac{P(a_k, b_l)}{P(a_k)P(b_l)} \right\}. \quad (8.32)$$

Взаємна інформація вимірюється в тих же одиницях, що й ентропія, наприклад в бітах. Величина $I(A, B)$ показує, скільки в середньому отримуємо біт інформації про реалізацію ансамблю A , спостерігаючи реалізацію ансамблю B .

Сформулюємо основні властивості взаємної інформації:

$$1. \quad I(A, B) \geq 0, \quad (8.33)$$

причому рівність має місце тоді і тільки тоді, коли A і B незалежні між собою. Це випливає із визначення (8.30) та нерівності (8.29).

$$2. \quad I(A, B) = I(B, A), \quad (8.34)$$

тобто B містить стільки ж інформації відносно A , скільки A містить відносно B . Ця властивість випливає із симетрії виразу (8.32). Тому можна також записати:

$$I(A, B) = H(B) - H(B | A). \quad (8.35)$$

$$3. \quad I(A, B) \leq H(A), \quad \text{й} \quad (8.36)$$

причому рівність має місце, коли за реалізацією B можна однозначно відновити A . Це випливає із (8.29) і (8.30).

$$4. \quad I(A, B) \leq H(B), \quad (8.37)$$

причому рівність має місце, коли за реалізацією A можна точно встановити реалізацію B . Це випливає із (8.34) та (8.36).

5. Покладаючи в (8.30) $B=A$ і враховуючи, що $H(A | A) = 0$, отримаємо:

$$I(A, A) = H(A).$$

Це дозволяє інтерпретувати ентропію джерела як його *власну інформацію*, тобто інформацію, що міститься в ансамблі A про саму себе.

Отримані співвідношення дозволяють поглянути на сутність ентропії і з других точок зору. Так, із (8.37) видно, що ентропія ансамблю A представляє собою максимальну кількість інформації, яка може міститись в A відносно будь-якого іншого ансамблю B . Із (8.30) випливає, що для того, щоб відновити реалізацію ансамблю A в точності, необхідно передати в середньому кількість інформації, рівну ентропії A .

Нехай A – ансамбль дискретних повідомлень, а B – ансамбль дискретних сигналів, в які перетворюються повідомлення A . Тоді $I(A, B) = H(A)$ в тому і тільки в тому випадку, коли перетворення $A \rightarrow B$ зворотне. При незворотному перетворенні $I(A, B) < H(A)$ і різницю $H(A) - I(A, B) = H(A|B)$ можна назвати *втратою інформації* при перетворенні $A \rightarrow B$. Її називають також *ненадійністю*. Таким чином, інформація не втрачається тільки при зворотних перетвореннях.

Якщо T – середній час передачі одного повідомлення, то поділивши формули (8.25) – (8.37) на T і позначаючи

$$H'(A|B) = \frac{1}{T} H(A|B), \quad I'(A, B) = \frac{1}{T} I(A, B)$$

і. т. д., отримаємо відповідні рівності для ентропій і кількості інформації, розрахованих не на одне повідомлення, а на одиницю часу. Величина $I'(A, B)$ називається швидкістю передачі інформації від A до B (чи навпаки).

Якщо, наприклад, U – ансамбль сигналів на вході дискретного каналу, а Z – ансамбль сигналів на його виході, то швидкість передачі інформації по каналу:

$$I'(U, Z) = H'(U) - H'(U|Z) = H'(Z) - H'(Z|U). \quad (8.38)$$

Ці співвідношення наглядно ілюструє рис. 8.3. Тут $H'(U)$ – продуктивність джерела сигналу U , який передається, а $H'(Z)$ – “продуктивність” каналу, тобто повна власна інформація в прийнятому сигналі за одиницю часу. Величина $H'(U|Z)$ представляє собою швидкість “витоку” інформації при проходженні через канал, а $H'(Z|U)$ – швидкість передачі побічної інформації,

що не має відношення до U і створюється присутніми в каналі завадами.

Співвідношення між $H'(U | Z)$ і $H'(Z | U)$ залежить від властивостей каналу. Так, наприклад, при передачі телефонного сигналу по каналу з вузькою смугою перепуску, недостатньою для задовільного відтворення сигналу, та з низьким рівнем завад втрачається частина корисної інформації, але майже не отримується некорисної.

В цьому випадку $H'(U | Z) \gg H'(Z | U)$. Якщо ж сигнал відтворюється точно, але в паузах ясно прослуховуються “наводки” від сусіднього телефонного каналу, то майже не втрачаючи корисної інформації, можна отримати багато додаткової, як правило, некорисної інформації і $H'(U | Z) \ll H'(Z | U)$.

8.4. Швидкість і пропускна спроможність дискретного каналу без перешкод

Швидкість передавання інформації R дорівнює кількості інформації, переданої в середньому за одиницю часу, і вимірюється в двійкових одиницях за загальною тривалістю T (дв. од./с).

Для ергодичних послідовностей повідомлень, коли допускається усереднення за часом, швидкість, дв. од./с:

$$R = \lim_{T \rightarrow \infty} \frac{I(a_T)}{T}, \quad (8.39)$$

де $I(a_T)$ - кількість інформації в послідовності повідомлень a_T , загальна тривалість яких рівна T .

Кількість інформації, що створюється джерелом повідомлень у середньому за одиницю часу, називається продуктивністю джерела R_D .

Враховуючи, що при $T \rightarrow \infty$:

$$I(a_T) = nH(a), \quad T = n\bar{\tau},$$

де n - число повідомлень; $\bar{\tau}$ - середня тривалість повідомлення, одержимо:

$$R_{\mathcal{D}} = \lim_{n \rightarrow \infty} \frac{nH(a)}{n\bar{\tau}} = \frac{H(a)}{\bar{\tau}}, \quad (8.40)$$

де $\bar{\tau} = \sum_{i=1}^n \tau_i P(\tau_i) = \sum_{i=1}^n \tau_i P(a_i)$ - математичне сподівання τ_i ;

$P(\tau_i) = P(a_i)$ - ймовірність повідомлення a_i тривалістю τ_i .

Для повідомлень однакової тривалості найбільшу продуктивність має джерело з максимальною ентропією $H_0 = \log m$, тобто:

$$R_{\mathcal{D} \max} = \frac{1}{\tau} \log m. \quad (8.41)$$

Видає джерелом інформація з каналів зв'язку передається за допомогою сигналів $u(t)$, які мають іншу природу та, у загальному випадку, інші статистичні характеристики.

При цьому швидкість визначається так:

$$R = \lim_{T \rightarrow \infty} \frac{I(u_T)}{T} = \frac{H(u_T)}{\bar{\tau}}. \quad (8.42)$$

Максимально можлива швидкість передавання інформації в каналах зв'язку при фіксованих обмеженнях називається пропускною спроможністю каналу, дв. од./с:

$$C = \max R = \max \frac{H(u)}{\bar{\tau}} = \frac{\max H(u)}{\min \bar{\tau}} = \frac{\log m}{\tau}, \quad (8.43)$$

де $\bar{\tau} = \tau$ (тривалості сигналів однакові).

Для двійкових дискретних каналів (при $m = 2$) пропускна спроможність:

$$C = \frac{\log m}{\tau} = \frac{\log_2 2}{\tau} = \frac{1}{\tau}. \quad (8.44)$$

Для двійкових дискретних каналів пропускна спроможність збігається зі швидкістю телеграфування, бод:

$$C = \frac{1}{\tau} = V.$$

При телеграфуванні мінімальна смуга перепуску:

$$F = F_m,$$

де $F_M = 1/(2\tau)$ - частота маніпуляції; $\tau = 1/(2F_M)$.

При цьому пропускна спроможність (межа Найквіста):

$$C = \frac{1}{\tau} = \frac{1}{[1/2F_m]} = 2F_m. \quad (8.45)$$

Поняття пропускної спроможності застосовується не тільки до всього каналу в цілому, але й до окремих його ланок. Суттєвим тут є те, що пропускна спроможність C_1 будь якої ланки не перевищує пропускної спроможності C_2 другої ланки, якщо вона міститься всередині першої. Співвідношення $C_1 < C_2$ зумовлене можливістю появи додаткових обмежень, накладених на ланки каналу при його розширенні, які знижують пропускну спроможність.

8.5. Оптимальні статистичні кодування повідомлень

Статистичним оптимальним називається кодування, при якому найкраще використовується пропускна спроможність каналу без завад.

Для дискретних каналів без завад Шенноном була доведена така теорема:

якщо продуктивність джерела $R_D \leq C - \varepsilon$, де ε - як завгодно мала величина, то завжди існує спосіб кодування, який дозволяє передавати по каналу всі повідомлення джерела. Передачу всіх повідомлень при $R_D > C$ здійснити неможливо.

Суть теореми зводиться до того, що якою б не була великою надлишковістю джерела, всі його повідомлення можуть бути передані по каналу. Зворотнє твердження

теорему легко доводиться від протилежного. Припустимо, що $R_D > C$, тоді для передачі всіх повідомлень джерела по каналу необхідно, щоб швидкість передачі інформації R була не меншою від R_D . Тоді маємо $R \geq R_D > C$, що неможливо, так як за визначенням пропускна спроможність $C = R_{\max}$.

Для раціонального використання пропускної спроможності каналу необхідно застосовувати відповідні способи кодування. При оптимальному кодуванні фактична швидкість передачі інформації по каналу R наближається до пропускної спроможності C , що досягається шляхом узгодження джерела з каналом. Повідомлення кодуються таким чином, щоб вони найкраще відповідали обмеженням, які накладаються на сигнали, які передаються каналом зв'язку. Тому структура оптимального коду залежить як від статистичних характеристик джерела, так і від особливостей каналу.

Розглянемо основні принципи оптимального кодування на практиці узгодження джерела незалежних повідомлень з двійковим каналом без завад.

При цьому процес кодування міститься в однозначному перетворенні повідомлень джерела в двійкові кодові комбінації. Тоді ентропія кодових комбінацій дорівнюватиме ентропії джерела:

$$H(u) = H(a),$$

а швидкість:

$$R = \frac{H(u)}{\bar{\tau}_k} = \frac{H(a)}{\bar{\tau}_k}, \quad (8.46)$$

де $\bar{\tau}_k$ - середня тривалість кодової комбінації, яка визначається як математичне очікування:

$$\bar{\tau}_k = \sum_{i=1}^m \bar{\tau}_{ki} P(a_i) = \tau_0 \sum_{i=1}^m n_i P(a_i);$$

τ_0 - тривалість одного елемента коду; n_i - кількість елементів в комбінації, що присвоюється повідомленню a_i .

Враховуючи, що при незалежних нерівномірних повідомленнях

$$H(a) = -\sum_{i=1}^m P(a_i) \log P(a_i),$$

отримаємо:

$$R = \frac{\sum_{i=1}^m P(a_i) \log P(a_i)}{\tau_0 \sum_{i=1}^m n_i P(a_i)}. \quad (8.47)$$

Чисельник цього виразу визначається виключно статистичними властивостями джерела, а величина τ_0 - характеристиками каналу.

Як видно з виразу (8.47) для того, щоб швидкість передавання досягла свого максимального значення $C = \frac{4}{\tau_0}$, необхідно виконати умову:

$$n_i = -\log P(a_i) = I(a_i), \quad (8.48)$$

що відповідає мінімуму τk і максимуму R . Очевидно, вибір $n_i < I(a_i)$ не має сенсу, тому що при цьому $R > C$, що суперечить теоремі Шеннона для сигналів без завад.

Код Шеннона-Фано. Одним із кодів, які задовольняють умові (8.48), є код Шеннона-Фано. Сутність кодування полягає в наступному. Нехай маємо джерело повідомлення, яке виробляє 4 повідомлення: a_1, a_2, a_3, a_4 з імовірностями:

$$P(a_1) = 0.5; \quad P(a_2) = 0.5; \quad P(a_1) = 0.25; \quad P(a_3) = P(a_4) = 0.125.$$

Усі повідомлення записуються до кодової табл. 8.1. у порядку зменшення їх імовірностей. Потім вони поділяються на дві групи так, щоб суми їх імовірностей були однаковими. В даному випадку в першу групу входить повідомлення з імовірністю

$P(a_1) = 0.5$, а у другу – повідомлення a_1, a_2 та a_4 з сумарною імовірністю, яка також дорівнює 0.5.

Комбінаціям, що відповідають повідомленням першої групи, надається як перший символ код „0”, а комбінаціям другої групи - „1”. Кожна з двох груп знову поділяється на дві групи за тим же правилом надання символів „0” і „1”.

В ідеальному випадку після першого ділення ймовірності кожної групи мають дорівнювати 0.5, а після другого ділення – 0.25 і т.д. Процес триває до тих пір, поки в групі не залишиться по одному повідомленню.

Таблиця 8.1

a_i	$P(a_i)$	Групи			Комбі- нації	n_i	$I(a_i)$
		I	II	III			
a_1	0,5	0			0	1	1
a_2	0,25	1	0		10	2	2
a_3	0,125	1	1	0	110	3	3
a_4	0,125	1	1	1	111	3	3

При заданому розподілі ймовірностей повідомлень код стає нерівномірним, його комбінації мають різну кількість елементів n_i .

Швидкість передавання дорівнює пропускну здатності:

$$R = C = 1/\tau_0.$$

В цьому легко переконатися, підставивши числові значення:

$$R = \frac{-0.5 \log 0.5 - 0.25 \log 0.25 - 2 \cdot 0.125 \log 0.125}{(0.5 \cdot 1 + 0.25 \cdot 2 + 2 \cdot 0.125 \cdot 3) \tau_0} = \frac{1.75}{1.75 \tau_0} = \frac{1}{\tau_0} = C.$$

Важливою властивістю коду Шеннона-Фано є те, що незважаючи на його нерівномірність, тут *не потрібні розмежувальні знаки*. Це зумовлено тим, що короткі комбінації не є початком більш довгих комбінацій. Вказану властивість легко перевірити на прикладі будь-якої послідовності:

$$\underbrace{10}_{a_2} \quad \underbrace{0}_{a_1} \quad \underbrace{10}_{a_2} \quad \underbrace{110}_{a_3} \quad \underbrace{111}_{a_4}$$

Таким чином, *основний принцип оптимального кодування* зводиться до того, що найбільш імовірним повідомленням мають надаватися короткі комбінації, а повідомленням з малою імовірністю – більш довгі комбінації.

Основний недолік, зумовлений повною відсутністю надлишковості, полягає в тому, що *оптимальні коди прийнятні тільки для каналів, в яких вплив завад незначний*.

8.6. Дискретні канали з завадами

Наявність завад в каналах призводить до руйнування та незворотної втрати частини інформації, що надходить від джерела повідомлення.

Оскільки в дискретному каналі з завадами отриманому сигналу V_j може відповідати передавання одного з декількох сигналів U_i , то після отримання V_j залишається деяка невизначеність відносно переданого сигналу. Відповідність між U та V має випадковий характер, тому ступінь невизначеності характеризується умовною апостеріорною ймовірністю:

$$P(u_i / v_j) < 1.$$

Кількість інформації, яка необхідна для усунення невизначеності, що залишилася, дорівнює тій частині інформації, яку знищено в результаті дії завад. Кількість отриманої інформації визначається як різниця:

$$I(u_i, v_j) = \log \frac{1}{P(u_i)} - \log \frac{1}{P(u_i / v_j)} = \log \frac{P(u_i / v_j)}{P(u_i)}. \quad (8.49)$$

Для оцінки середньої кількості отриманої інформації при передаванні одного повідомлення значення $I(u_i, v_j)$ необхідно усереднити по всьому ансамблю $\mathbf{U}$ та $\mathbf{V}$:

$$I(u, v) = \bar{I}(u_i, v_j) = \sum_{i=1}^{m_u} \sum_{j=1}^{m_v} P(u_i, v_j) \log \frac{P(u_i / v_j)}{P(u_i)} = \sum_{j=1}^{m_v} P(v_j) \sum_{i=1}^{m_u} P(u_i, v_j) \log \frac{P(u_i / v_j)}{P(u_i)} \quad (8.50)$$

де $P(u_i, v_j) = P(v_j)P(u_i / v_j) = P(u_i)P(v_j / u_i)$ - суцільна ймовірність переданого та отриманого сигналів; m_u - кількість сигналів в ансамблі на вході; m_v - кількість сигналів в ансамблі на виході каналу.

У загальному випадку $m_v \neq m_u$.

Величина $I(u, v)$ характеризує у середньому кількість інформації, яка містить отриманий сигнал відносно переданого сигналу $\mathbf{U}$ і тому $I(u, v)$ називають також *середньою взаємною інформацією між $\mathbf{U}$ та $\mathbf{V}$* .

Зазвичай вираз для $I(u, v)$ надається у вигляді формули:

$$I(u, v) = H(u) - H(u / v),$$

де $H(u) = - \sum_{i=1}^{m_u} \sum_{j=1}^{m_v} P(u_i, v_j) \log P(u_i) = - \sum_{i=1}^{m_u} P(u_i) \log P(u_i)$ ентропія джерела сигналів;

$$H(u/v) = -\sum_{i=1}^{m_u} \sum_{j=1}^{m_v} P(u_i, v_j) \log P(u_i / v_j) = -\sum_{j=1}^{m_v} P(v_j) \sum_{i=1}^{m_u} P(u_i / v_j) \log P(u_i / v_j),$$

умовна ентропія або ненадійність.

Вираз для $I(u, v)$ показує, що середня кількість отриманої інформації дорівнює середній кількості переданої інформації $H(u)$ мінус середня кількість інформації $H(u/v)$, яка загубилась у каналі внаслідок дії завад.

Інша форма запису:

$$I(u, v) = H(v) - H(u/v),$$

де $H(v) = -\sum_{j=1}^{m_v} P(v_j) \log P(v_j)$ - ентропія каналу виходу;

$$H(v/u) = -\sum_{i=1}^{m_u} P(u_i) \sum_{j=1}^{m_v} P(v_j / u_i) \log P(u_i / v_j) \text{ умовна ентропія або ентропія}$$

шуму (визначає непотрібну частину інформації, яка міститься в отриманих сигналах внаслідок дії завад).

Поняття швидкості передавання та пропускної спроможності, які отримані для дискретного каналу без завад, можна поширити і на канали з завадами:

$$R = \lim_{T \rightarrow \infty} \frac{I(u_T, v_T)}{T},$$

де u_T, v_T - відповідно послідовності сигналів тривалістю T , що передаються і отримуються.

Швидкість передавання можна подати у такому вигляді:

$$R = \frac{I(u, v)}{\bar{\tau}} = \frac{1}{\tau} [H(u) - H(u/v)] = \frac{1}{\tau} [H(v) - H(v/u)]$$

Пропускна спроможність:

$$C = \max R.$$

Для каналів з сигналами однакової тривалості пропускна спроможність:

$$C = \frac{\max I(u, v)}{\tau} = \frac{\max [H(u) - H(u/v)]}{\tau},$$

де максимум є по всіх можливих ансамблях сигналів $\mathbf{U}$.

Для двійкового симетричного каналу:

$$C = \frac{1}{\tau} [1 + P_0 \log P_0 + (1 - P_0) \log(1 - P_0)] \quad (8.51)$$

або $C = v [1 + P_0 \log P_0 + (1 - P_0) \log(1 - P_0)]$, де P_0 - повна ймовірність помилки;

Для симетричного каналу при $m_u = m_v > 2$:

$$C = \frac{1}{\tau} \left[\log m + P_0 \log \frac{m-1}{P_0} - (1 - P_0) \log \frac{1}{1 - P_0} \right].$$

Залежність C/v від P відповідно (8.51) зображена на рис. 8.4. При $P = 0.5$ пропускна спроможність двійкового каналу $C = 0$, оскільки при такій імовірності помилки послідовність вихідних двійкових символів можна одержати зовсім не передаючи сигнали по каналу, а вибираючи їх навмання, тобто при $P = 0.5$ послідовність на вході і на виході каналу незалежні. Випадок $C = 0$ називають *обривом каналу*. Те, що пропускна спроможність при $P = 1$ в двійковому каналі така ж, як і при $P = 0$ (канал без завад), пояснюється тим, що при $P = 1$ достатньо всі вихідні символи інвертувати (тобто замінити 0 на 1 і 1 на 0), щоб правильно відновити вхідний сигнал

Теорема Шеннона для дискретного каналу з завадами. Пропускна спроможність каналу характеризує потенціальні можливості передачі інформації. Вони розкриваються в фундаментальній теоремі теорії інформації, відомій як основна теорема кодування К. Шеннона. Стосовно до дискретного джерела вона формулюється так:

якщо індуктивність джерела повідомлень R_D менша від пропускну здатності каналу C , тобто $R_D < C$, то існує спосіб кодування (перетворення повідомлення в сигнал на вході) і декодування (перетворення сигналу в повідомлення на виході каналу), при якому ймовірність помилкового декодування та ненадійність можуть бути як завгодно малі. Якщо ж $R_D > C$, то таких способів не існує.

Таким чином, згідно до теореми К. Шеннона скінченна величина C - це граничне значення швидкості безпомилкової передачі інформації по каналу. Ця теорема, на жаль, неконструктивна, тобто вона не вказує на конкретний спосіб кодування, існування якого доводиться. Тим не менше, значення теореми важко переоцінити, оскільки вона в корені змінила світогляд на принципові можливості техніки зв'язку. До К. Шеннона вважалося, що в каналі з шумами можна забезпечити як завгодно малу ймовірність помилки тільки при необмеженому зменшенні швидкості передачі інформації. Такий, скажемо, шлях підвищення вірності зв'язку за рахунок повторення символів в каналі без пам'яті.

8.7. Інформація в неперервних повідомленнях. Епсилон-ентропія.

Джерело неперервних повідомлень за скінченний час T може видати будь-яку інформацію з безкінечної множини реалізацій деякого ансамблю повідомлень. Ймовірність появи деякої конкретної реалізації рівна нулю. Якщо спробувати визначити ентропію та продуктивність такого джерела шляхом граничного переходу для неперервних повідомлень, то вони виявляться безмежними.

Зміст цього результату заключається в тому, що для передачі неперервного повідомлення з абсолютною точністю треба було б передати безмежно велику кількість інформації, що, зрозуміло, неможливо зробити за скінченний час, користуючись каналом зі скінченною пропускну здатністю. Точно так неперервне повідомлення не можна абсолютно точно запам'ятати (записати) при наявності як завгодно слабкої завади.

Тим не менш, як відомо, неперервні повідомлення (наприклад, телефонні, телевізійні) успішно передаються каналами зв'язку з завадами та записуються, наприклад, на магнітну плівку. Це пояснюється тим, що на практиці ніколи не вимагається точне відтворення переданого чи записаного повідомлення. А для передачі навіть з дуже високою, але обмеженою точністю вимагається скінченна кількість інформації так, як і при передачі дискретних повідомлень. Зрозуміло, ця кількість інформації тим більша, чим вища точність, з якою вимагається передати (відтворити)

неперервне повідомлення. Нехай допустима точність вимірюється деяким параметром $\mathcal{E}$. Та мінімальна кількість інформації, яку необхідно передати каналом зв'язку для відтворення неперервного повідомлення з неточністю не більшою від допустимої, академік А.Н. Колмогоров запропонував називати $\mathcal{E}$ - ентропією (епсілон-ентропією).

Критерій $\mathcal{E}$, який визначає точність, що вимагається, може бути яким завгодно. Будемо називати два варіанти повідомлення, які розрізняються не більш ніж на $\mathcal{E}$, *еквівалентними*. Це значить, що якщо послано одне повідомлення, а прийняте інше, еквівалентне йому, то за переданим критерієм передане повідомлення вважається прийнятим вірно. Так, в системі телефонного зв'язку, якщо необхідно передати лише зміст мови, то один і той же текст, який розбірливо прочитаний двома різними дикторами (наприклад, чоловіком та жінкою), представляє собою еквівалентні повідомлення, не дивлячись на те, що вони різко відрізняються одне від одного навіть за спектром. Критерієм еквівалентності повідомлень тут є розбірливість мови. При художніх мовних передачах такий критерій не являється прийнятним, оскільки в цих випадках істотні й більш тонші характеристики повідомлення.

Ентропія неперервних повідомлень. При передаванні неперервних повідомлень передані сигнали $s(t)$ є функціями часу, які належать до деякої множини, а отримані сигнали $x(t)$ будуть їх реалізаціями.

Всі дійсні сигнали мають спектри, загальна енергія яких зосереджена в обмеженій смузі частот F . Згідно з теоремою В.А. Котельникова такі сигнали визначаються своїми значеннями в точках відліку, які вибирають через інтервали:

$$\Delta t = \frac{1}{2F}.$$

У каналі на сигнал накладаються завади, внаслідок чого кількість різних рівнів сигналу в точках відліку буде скінченою. Отже, сукупність значень неперервного сигналу еквівалентна деякій дискретній скінченій сукупності. Це дає змогу визначити необхідну кількість інформації та пропускну спроможність при передаванні неперервних повідомлень, що базуються на результатах, які отримано для дискретних повідомлень.

Визначимо спочатку кількість інформації, яка міститься в одному відліку сигналу P_{x_j} відносно переданого сигналу S_i :

$$I(s_i, x_j) = \lim_{\Delta s \rightarrow 0} \log \frac{p(s_i / x_j)}{p(s_i) \Delta s} = \log \frac{p(s_i / x_j)}{p(s_i)},$$

де $p(s_i)$, $p(s_i / x_j)$ - щільність ймовірності.

Середня кількість інформації:

$$I(s, x) = \iint_{S \times X} p(s, x) I(s, x) ds dx = \iint_{S \times X} p(s, x) \log \frac{p(s / x)}{p(s)} dx ds,$$

де $p(s, x)$ - сумісна щільність ймовірності; S, X - області можливих значень s_i та x_j .

Вираз для $\overline{I(s, x)}$ можна подати в двох виглядах:

$$\overline{I(s, x)} = H(s) - H(s / x);$$

$$\overline{I(s, x)} = H(x) - H(x / s),$$

(8.52)

де

$$H(s) = - \int_S p(s) \log p(s) ds \quad \text{- диференціальна ентропія сигналу } S;$$

$$H(x) = - \int_X p(x) \log p(x) dx \quad \text{- диференціальна ентропія сигналу } X;$$

$$H(s / x) = - \int_S \int_X p(x, s) \log p(s / x) ds dx \quad \text{- умовна диференціальна ентропія сигналу } S;$$

$$H(x/s) = - \int \int_{S \times X} p(s, x) \log p(x/s) ds dx - \text{умовна диференціальна}$$

ентропія сигналу або ентропія шуму.

Величина $H(s)$ характеризує інформаційні властивості сигналів і за формулою запису аналогічно ентропії дискретних повідомлень. Так як у виразі (8.52) входить диференціальний розподіл імовірностей $P(s)$, то $H(s)$ називають диференціальну ентропію сигналу S . Вираз для $H(s/x)$ представляє собою умовну диференціальну ентропію сигналу S .

Багато властивостей ентропії неперервного розподілу аналогічні властивостям ентропії дискретного сигналу.

Якщо неперервний сигнал S обмежений інтервалом $s_1 \leq S < s_2$, то ентропія $H(s)$ максимальна і дорівнює $\log(s_2 - s_1)$, коли сигнал має рівномірний розподіл:

$$p(s) = 1/(s_2 - s_1).$$

Якщо обмежене середньоквадратичне значення неперервного сигналу:

$$\int_{-\infty}^{\infty} s^2 p(s) ds = \sigma^2,$$

то ентропія максимальна при нормальному розподілі щільності ймовірностей:

$$P(s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-s^2/(2\sigma^2)};$$

$$H(s)_{\max} = \int_{-\infty}^{\infty} p(s) \log \frac{1}{P(s)} ds = \int_{-\infty}^{\infty} p(s) \log(\sqrt{2\pi}\sigma) ds + \\ + \int_{-\infty}^{\infty} p(s) \frac{s^2 \log e}{2\sigma^2} ds = \log(\sqrt{2\pi e}\sigma).$$

де $e = 2.71$.

Величина диференціальної ентропії (на відміну від ентропії дискретних сигналів) залежить від розмірності неперервного сигналу, завдяки чому вона не є мірою кількості інформації, хоч і характеризує ступінь невизначеності, яка властива джерелу. Тільки різниця ентропій кількісно визначає середню інформацію.

8.8. Швидкість та пропускну спроможність неперервного каналу. Формула Шеннона

Для визначення середньої кількості $\overline{I(s, x)}$, яка передається сигналом на інтервалі T , необхідно розглянути $n = 2FT$ відліків неперервного сигналу на вході каналу $s_1, s_2, \dots, s_n$ і на виході $x_1, x_2, \dots, x_n$. При цьому можна записати:

$$\overline{I(s, x)} = \begin{cases} H_T(s) - H_T(s/x), \\ H_T(x) - H_T(s), \end{cases} \quad (8.54)$$

де

$$H_T(x) = - \int \int \dots \int_{x_1 x_2 \dots x_n} p(x_1, x_2, \dots, x_n) \log p(x_1, x_2, \dots, x_n) dx_1 \dots dx_n,$$

$$H_T(x/s) = - \int \int \dots \int_{s_1 s_2 \dots s_n} \int \int \dots \int_{x_1 x_2 \dots x_n} p(s_1, s_2, \dots, s_n; x_1, x_2, \dots, x_n) \times \\ \times \log p(s_1, s_2, \dots, s_n; x_1, x_2, \dots, x_n) ds_1 ds_2 \dots ds_n dx_1 dx_2 \dots dx_n.$$

Ентропія $H_T(s)$ та $H_T(s/x)$ описуються аналогічними виразами, тільки всюди необхідно поміняти місцями змінні X та S .

Швидкість передавання неперервним каналом знаходять як межу:

$$R = \lim_{T \rightarrow \infty} \frac{\overline{I(s, x)}}{T};$$

$$C = \max R = \max \lim_{T \rightarrow \infty} \frac{1}{T} \overline{I(s, x)},$$

де максимум обчислюють по всіх можливих ансамблях вхідних сигналів S .

При належних сигналах S і заваді ω одержуємо формулу Шеннона []:

$$C = F \log \left(\frac{P_c}{P_{\omega}} + 1 \right), \quad (8.55)$$

де P_c - потужність сигналу; P_{ω} - потужність завад; F - ширина спектра сигналу.

Формулу отримано в припущенні, що S і ω мають нормальний розподіл, і тому сигнал S також повинен мати нормальний розподіл щільності ймовірності.

Таким чином для отримання максимальної швидкості передавання необхідно застосувати сигнали з нормальним розподілом щільності ймовірності та рівномірним спектром.

Формула Шеннона відіграє важливу роль у теорії та техніці зв'язку.

При рівномірному спектрі:

$$P_{\omega} = N_0 F,$$

тоді

$$C = F \log \left(\frac{P_c}{N_0 F} + 1 \right).$$

Зі збільшенням F пропускна спроможність монотонно зростає і наближається, як вказано на рис.8.5, а, до величини дв.од./с.:

$$C_{\max} = \frac{P_c}{N_0} \log e = 1.44 \frac{P_c}{N_0}.$$

Формулу Шеннона (8.55), яку виведено для рівномірних спектрів сигналу і шуму, можна поширити також на випадок нерівномірних спектрів. З цією метою в межах деякої частоти f виберемо достатньо вузьку смугу Δf , в якій енергетичні спектри сигналу $G_c(t)$ та $G_{ш}(t)$ будуть постійними. При цьому для цієї смуги:

$$\Delta C = \Delta f \log \left(\frac{\Delta P_c}{\Delta P_{ш}} + 1 \right) = \Delta f \log \left(\frac{G_c(f)}{G_{ш}(f)} + 1 \right),$$

$$\text{де } \Delta P_c = \Delta f G_c(f), \quad \Delta P_{ш} = \Delta f G_{ш}(f).$$

Повна пропускна спроможність обчислюється як інтеграл по всіх частотах спектра сигналу:

$$C = \int_{f_1}^{f_2} \log \left(\frac{G_c(f)}{G_{ш}(f)} + 1 \right) df.$$

Можна показати, що при заданому спектрі шуму $G_{ш}(t)$ та обмеженій потужності сигналу максимум C має місце при виконанні умови:

$$G_c(f) + G_{ш}(f) = \text{const}, \tag{8.56}$$

тобто потужність сигналу має збільшуватись на тих частотах, де зменшується потужність шуму і навпаки.

Як показано в [], при виконанні умови (8.56) найбільш зменшується пропускна спроможність, якщо спектр шуму є рівномірним, тобто є спектром білого шуму. Таким чином білий шум, що найбільш зменшує пропускну спроможність, є найнебезпечнішим видом завад.

Продуктивність джерела неперервних повідомлень. За відсутності будь-яких обмежень, накладених на неперервні повідомлення, кількість інформації, яка в них міститься, дорівнює безмежності:

$$I(u) = -\log P(u) = \lim_{\Delta u \rightarrow 0} [-\log \Delta u P(u)] = \infty. \quad (8.57)$$

Тому джерело таких повідомлень володіє безмежною продуктивністю. Для того, щоб кількість інформації та продуктивність джерела прийняли певний зміст і стали скінченими величинами, необхідно розглядати неперервне повідомлення $u(t)$ із врахуванням точності його оцінки. Остання, зокрема, визначається похибкою приборів, за допомогою яких вимірюється чи реєструється неперервне повідомлення. Зазвичай похибка кількісно оцінюється середньоквадратичним відхиленням наближеного неперервного повідомлення $u^*(t)$ від його точного значення $u(t)$:

$$\overline{\varepsilon_0^2} = \overline{[u^*(t) - u(t)]^2} = \int \int_{U U^*} p(u, u^*) (u^* - u)^2 du du^*. \quad (8.58)$$

Неважко зрозуміти, що чим менше $\overline{\varepsilon_0^2}$, тим менша кількість інформації в середньому міститься в $u^*(t)$ відносно $u(t)$ і тим вища продуктивність джерела.

Кількість інформації на виході джерела при $\overline{\varepsilon_0^2} > 0$ визначається так []:

$$J(u, u^*) = \int \int_{U U^*} p(u, u^*) \log \frac{p(u / u^*)}{p(u)} du du^*. \quad (8.59)$$

Для обмеженої похибки $\overline{\varepsilon_0^2} \leq \varepsilon$ завжди можна знайти такий спосіб відтворення $u(t)$ через $u^*(t)$, а отже такий розподіл $p'(u, u^*)$, при якому вираз (8.59)

досягає найменшого значення. Розподіл $p'(u, u^*)$ є найбільш вигідним, так як він дозволяє при заданій похибці відтворювати $u(t)$, використовуючи мінімальну кількість інформації. Найменше значення $J(u, u^*)$ при $\overline{\varepsilon_0^2} \leq \varepsilon$ називається *епсілон-ентропією*.

$$(H_\varepsilon(u) = \min_{\overline{\varepsilon_0^2} \leq \varepsilon} J(u, u^*)). \quad 8.60)$$

Тоді продуктивність джерела неперервних повідомлень:

$$R_D = \lim_{T \rightarrow \infty} \frac{H_\varepsilon(u)}{T}.$$

Для неперервного каналу з пропускною спроможністю C , на вхід якого підключається джерело, що має продуктивність R_D , Шенноном була доказана наступна теорема.

Якщо при заданій похибці оцінки повідомлень джерела $\overline{\varepsilon_0^2}$ його продуктивність $R_D < C$, то існує спосіб кодування, який дозволяє передавати всі неперервні повідомлення джерела з похибкою у відтворенні на виході каналу, що як завгодно мало відрізняється від $\overline{\varepsilon_0^2}$.

Інакше кажучи, додаткова неточність у відтворенні повідомлення $v(t)$ на виході каналу, зумовлена дією завад, може бути зроблена дуже незначною, тобто $[\overline{v(t) - u(t)}]^2 = \overline{\varepsilon_0^2} + \overline{\delta^2}$, де $\overline{\delta^2}$ - як завгодно мала величина.

Швидкість передачі інформації по каналу в кінці кінців визначається швидкістю потоку інформації на виході приймача:

$$R = \lim_{T \rightarrow \infty} \frac{1}{T} [H_T(v) - H_T(u^*)],$$

де $H_T(v)$ - ентропія прийнятого повідомлення $v(t)$, $H_T(\omega^*)$ - ентропія шуму на виході приймача. Якщо вважати, що повідомлення $v(t)$ і завада $\omega^*(t)$ на виході приймача мають нормальний розподіл і рівномірний спектр, то

$$R = F_m \log \left(\frac{P_c^*}{P_{ш}^*} + 1 \right). \quad (8.61)$$

Тут F_m - ширина спектра частот повідомлення, яке передається, зазвичай рівна смузі перепуску приймача на низькій частоті; P_c^* - середня потужність прийнятого повідомлення $v(t)$; $P_{ш}^*$ - середня потужність шуму на виході приймача.

Пропускна спроможність каналів зі змінними параметрами.

Характеристики системи зв'язку в значній мірі залежать від параметрів каналу зв'язку, який використовується для передачі повідомлень. Досліджуючи пропускну спроможність каналів, ми вважали, що їх параметри зберігаються постійними. Однак більшість реальних каналів мають змінні параметри. Параметри каналу, як правило, змінюються в часі випадковим чином. Як вже відмічалось, випадкові зміни коефіцієнта передачі каналу μ викликають завмирання сигналу, що еквівалентно дії мультиплікативної завади. Еквівалентна потужність цієї завади рівна $P_c D[\mu]$. Тоді сумарна потужність завад в каналі зі змінними параметрами при наявності адитивної завади буде рівною:

$$P_{\Pi\Sigma} = P_{\Pi} + P_c D[\mu]. \quad (8.62)$$

Звідси випливає, що завмирання сигналу приводять до збільшення потужності завад, а отже, і до зниження пропускну спроможності каналу (8.55).

Розглянемо, яким чином обчислюється пропускна спроможність каналу зі завмираннями при передачі неперервних повідомлень. В цих умовах необхідно знайти такий розподіл імовірностей сигналу, який при заданих статистичних властивостях коефіцієнта передачі μ забезпечувало б максимальну швидкість передачі інформації.

Рішення задачі в загальному вигляді викликає значні труднощі. Однак у випадку повільних завмирань, коли швидкість завмирань набагато менша від швидкості зміни неперервного повідомлення, можна з достатньою точністю

передбачити за поточними значеннями коефіцієнта μ його наступні значення. При такому допущенні максимум швидкості передачі, як і раніше, має місце для повідомлень з нормальним розподілом, що дає можливість користуватися формулою (8.55). Підставляючи в (8.55) потужність $\mu^2 P_0$ сигналу, який приймається, де P_0 - передана потужність сигналу, отримаємо пропускну спроможність при фіксованому μ :

$$C(\mu) = F \log \left(\frac{\mu^2 P_0}{P_{\Pi}} + 1 \right). \quad (8.63)$$

Усереднивши $C(\mu)$ за всіма значеннями μ знайдемо середнє значення пропускну спроможності каналу при повільних завмираннях:

$$C = \int_0^{\infty} p(\mu) C(\mu) d\mu = \int_0^{\infty} p(\mu) F \log \left(\frac{\mu^2 P_0}{P_{\Pi}} + 1 \right) d\mu, \quad (8.64)$$

де $p(\mu)$ - щільність ймовірності коефіцієнта передачі μ . В багатьох каналах зміни коефіцієнта μ підчиняються закону Релея. На рис.8.6 показано залежності C / F від середнього відношення P_0 / P_{Π} для каналу з постійними параметрами 1 та каналу з релеєвським завмиранням 2. Із аналізу кривих рис.8.6 випливає, що повільні релеєвські завмирання зменшують пропускну спроможність каналу не більш ніж на 17%.

8.9. Визначення щільності розподілу ймовірностей і втрат інформації на виході нелінійного елемента.

Багато з елементів каналів зв'язку є нелінійними з нелінійними статистичними характеристиками різних типів: зона нечутливості, насичення, складні нелінійності, що поєднують зону нечутливості та ін. Наявність нелінійностей у каналі зв'язку приводить до ускладнень синтезу та до втрат інформації.

Визначимо щільність розподілу ймовірностей випадкового сигналу $\eta(t)$ на виході нелінійності з заданою статистичною характеристикою. При цьому будемо

вважати, що випадковий сигнал $\xi(t)$ на вході нелінійності $F[\xi(t)]$ має щільність $\omega_{\xi}(x)$ та функцію розподілу $W_{\xi}(x)$.

1. Статистична характеристика нелінійності типу зони нечутливості (рис.8.7, а):

$$F[\xi(t)] = \begin{cases} 0, & |\xi(t)| < b_n; \\ k\xi(t) - kb_n \operatorname{sign} \xi(t); & |\xi(t)| > b_n, \end{cases}$$

де b_i - половина ширини зони нечутливості.

Статистична характеристика, зворотна до зони нечутливості, представлена на рис.8.8, а. Випадковий процес $\eta(t)$ на виході нелінійності розподілений у цьому випадку з щільністю (рис.8.8, б):

$$\omega_{\eta}(x) = \begin{cases} \frac{1}{k} \omega_{\xi}\left(\frac{x - kb_i}{k}\right), & -\infty < x < 0; \\ S_0 \delta(x), & x = 0; \\ \frac{1}{k} \omega_{\xi}\left(\frac{x + kb_i}{k}\right), & 0 < x < \infty, \end{cases} \quad (8.65)$$

де постійна S_0 визначається через задану щільність на вході за формулою:

$$S_0 = \int_{-b_i}^{b_i} \omega_{\xi}(x) dx. \quad (8.66)$$

2. Статистична характеристика нелінійності типу насичення (рис.8.7, б):

$$F[\xi(t)] = \begin{cases} k\xi(t), & |\xi(t)| \leq b; \\ c \operatorname{sign} \xi(t), & |\xi(t)| > b, \quad b = \frac{c}{k}. \end{cases}$$

Для визначення щільності вихідного сигналу на рис.8.8, в зображена характеристика, зворотна насиченню.

Випадковий процес $\eta(t)$ на виході нелінійності має щільність (рис.8.8, г):

$$\omega_{\eta}(x) = \begin{cases} 0, |x| > c; \\ S_{-c} \delta(x + c), x = -c; \\ \frac{1}{k} \omega_{\xi}\left(\frac{x}{k}\right), -c < x < c; \\ S_c \delta(x - c), x = c, \end{cases} \quad (8.67)$$

де $\delta(x)$ - дельта-функція, а постійні S_{-c} і S_c визначаються через задану щільність на вході за формулами:

$$S_{-c} = \int_{-\infty}^{-c/k} \omega_{\xi}(x) dx = W_{\xi}(-c/k), c/k = b;$$

$$S_c = \int_{c/k}^{\infty} \omega_{\xi}(x) dx = 1 - W_{\xi}(c/k), \quad \int_{-\infty}^{c/k} + \int_{c/k}^{\infty} \equiv 1. \quad (8.68)$$

3. Посднання зони нечутливості та насичення з відсутністю лінійної ділянки.

Статистична характеристика нелінійності має вигляд (рис.8.8, д):

$$F[\xi(t)] = \begin{cases} 0, |\xi(t)| \leq b; \\ c \operatorname{sign} \xi(t), |\xi(t)| > b. \end{cases}$$

Статистична характеристика, зворотна до характеристики нелінійності (8.8, д), зображена на рис.8.8, е.

Випадковий процес $\eta(t)$ на виході нелінійності розподілений з щільністю (рис.8.8, ж):

$$\omega_{\eta}(x) = \begin{cases} 0, x \neq 0, x \neq \pm c; \\ S_{-c} \delta(x+c), x = -c; \\ S_0 \delta(x), x = 0; \\ S_c \delta(x-c), x = c, \end{cases} \quad (8.69)$$

де $S_{-c} = \text{const}$, $S_0 = \text{const}$, $S_c = \text{const}$, причому

$$S_{-c} = \int_{-\infty}^{-b} \omega_{\xi}(x) dx, \quad S_0 = \int_b^{-b} \omega_{\xi}(x) dx, \quad S_c = \int_b^{\infty} \omega_{\xi}(x) dx,;$$

$$S_{-c} + S_0 + S_c = 1.$$

4. Релейна характеристика. Нелінійність та зворотна до неї статистична характеристика представлені на рис.8.9, а та рис.8.9, б відповідно.

Випадковий процес $\eta(t)$ на виході нелінійності має щільність (рис.8.9, в) виду:

$$\omega_{\eta}(x) = \begin{cases} S_{-c} \delta(x+c), x = -c; \\ S_c \delta(x-c), x = c; \\ 0, |x| \neq c, \end{cases} \quad (8.70)$$

$$\text{де } S_{-c} = \int_{-\infty}^0 \omega_{\xi}(x) dx; \quad S_c = \int_0^{\infty} \omega_{\xi}(x) dx; \quad S_{-c} + S_c = 1.$$

5. Нелінійність загального виду, яка поєднує зони нечутливості і насичення. Статистична характеристика нелінійності (рис.8.7, в):

$$F[\xi(t)] = \begin{cases} 0, |\xi(t)| \leq b_i; \\ k \xi(t) - k b_i \operatorname{sign} \xi(t), b_i < |\xi(t)| \leq b_i + c/k; \\ c \operatorname{sign} \xi(t), |\xi(t)| > b_i + c/k, \end{cases} \quad (8.71)$$

звідки розглянуті нелінійності (1 - 4) отримують як частинні випадки: 1) $b_i = 0$; 2) $c = \infty$; 3) $k = \infty$; 4) $b_i = 0, k = \infty$.

Тоді в загальному випадку $\eta(t)$ на вході нелінійності має щільність:

$$\omega_{\eta}(x) = \begin{cases} 0, |x| > c; \\ S_{-c} \delta(x+c), x = -c; \\ \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right), -c < x < 0; \\ S_0 \delta(x), x = 0; \\ \frac{1}{k} \omega_{\xi} \left(\frac{x + kb_i}{k} \right), 0 < x < c; \\ S_c \delta(x-c), x = c, \end{cases} \quad (8.72)$$

де постійні:

$$S_{-c} = \int_{-\infty}^{-b_i - c/k} \omega_{\xi}(x) dx; \quad S_0 = \int_{-b_i}^{b_i} \omega_{\xi}(x) dx; \quad S_c = \int_{b_i + c/k}^{\infty} \omega_{\xi}(x) dx.$$

Визначення втрат інформації на виході нелінійного елемента

Визначимо ентропію випадкового сигналу $\eta(t)$ та втрати інформації на виході нелінійності з заданою статистичною характеристикою.

1. Щільність розподілу ймовірностей випадкового сигналу нелінійності типу зони нечутливості визначається виразами (8.65), (8.66). Тому ентропія сигналу $\eta(t)$:

$$\begin{aligned}
h(\eta) &= - \int_{-\infty}^{\infty} \omega_{\eta}(x) \log_2 \omega_{\eta}(x) = \\
&= - \int_{-\infty}^0 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right) \log_2 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right) dx - \\
&\quad - \int S_0 \delta(x) \log_2 S_0 \delta(x) dx - \\
&\quad - \int_0^{\infty} \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right) \log_2 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right) dx.
\end{aligned}
\tag{8.73}$$

У виразі (8.73) позначимо:

$$\begin{aligned}
J_2 &= \int_{(0)} S_0 \delta(x) \log_2 S_0 \delta(x) dx = \\
&= S_0 \int_{(0)} \delta(x) \log_2 S_0 dx + S_0 \int_{(0)} \delta(x) \log_2 \delta(x) dx = \\
&= S_0 \log_2 S_0 + 0 = S_0 \log_2 S_0,
\end{aligned}
\tag{8.74}$$

$$\text{де } S_0 = \int_{-b_i}^{b_i} \omega_{\xi}(x) dx = W_{\xi}(b_i) - W_{\xi}(-b_i);$$

$$\begin{aligned}
J_1 &= \int_{-\infty}^0 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right) \log_2 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_i}{k} \right) dx = \left| \frac{\frac{x}{k} - b_i = u}{dx = kdu} \right|_{x=0, u=-b_i} = \\
&= \int_{-\infty}^{-b_i} \omega_{\xi}(x) \log_2 \frac{1}{k} \omega_{\xi}(x) dx;
\end{aligned}
\tag{8.75}$$

$$J_3 = \int_0^\infty \frac{1}{k} \omega_\xi \left(\frac{x + kb_i}{k} \right) \log_2 \frac{1}{k} \omega_\xi \left(\frac{x + kb_i}{k} \right) dx = \left| \begin{array}{l} \frac{x}{k} + b_i = u \\ dx = k du \\ x = 0, u = +b_i \end{array} \right| =$$

$$= \int_{b_i}^\infty \omega_\xi(x) \log_2 \frac{1}{k} \omega_\xi(x) dx. \quad (8.76)$$

Тоді ентропія з врахуванням виразів (8.74) – (8.76) приймає вигляд:

$$h(\eta) = -(J_1 + J_2 + J_3) =$$

$$= - \int_{-\infty}^{-b_i} \omega_\xi(x) \log_2 \frac{1}{k} \omega_\xi(x) dx - S_0 \log_2 S_0 - \int_{b_i}^\infty \omega_\xi(x) \log_2 \frac{1}{k} \omega_\xi(x) dx =$$

$$= - \int_{-\infty}^{-b_i} \omega_\xi(x) \log_2 \omega_\xi(x) dx - \int_{b_i}^\infty \omega_\xi(x) \log_2 \omega_\xi(x) dx -$$

$$- \log_2 \frac{1}{k} \left[\int_{-\infty}^{-b_i} \omega_\xi(x) dx + \int_{b_i}^\infty \omega_\xi(x) dx \right] - S_0 \log_2 S_0,$$

звідки ентропія сигналу $\eta(t)$, перерахована до входу нелінійності,

$$h(\eta) = - \int_{-\infty}^{-b_i} \omega_\xi(x) \log_2 \omega_\xi(x) dx - \int_{b_i}^\infty \omega_\xi(x) \log_2 \omega_\xi(x) dx -$$

$$- [W_\xi(b_i) - W_\xi(-b_i)] \log_2 [W_\xi(b_i) - W_\xi(-b_i)] \quad (8.77)$$

В той же час ентропія сигналу $\xi(t)$ на вході нелінійності без врахування зони нечутливості:

$$h(\xi) = - \int_{-\infty}^\infty \omega_\xi(x) \log_2 \omega_\xi(x) dx. \quad (8.78)$$

З врахуванням виразів (8.77), (8.78) втрати інформації в системі синхронізації, зумовлені наявністю зони нечутливості:

$$\begin{aligned}
\Delta J &= h(\xi) - h(\eta) = - \int_{-\infty}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + \\
&+ \int_{-\infty}^{-b_i} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + \int_{b_i}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + \\
&+ [W_{\xi}(b_i) - W_{\xi}(-b_i)] \log_2 [W_{\xi}(b_i) - W_{\xi}(-b_i)] = \\
&= \int_{-b_i}^{b_i} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + [W_{\xi}(b_i) - W_{\xi}(-b_i)] \log_2 [W_{\xi}(b_i) - W_{\xi}(-b_i)]
\end{aligned}
\tag{8.79}$$

2. Нелінійності типу насичення.

Щільність розподілу імовірностей випадкового сигналу визначається виразами (8.67) і (8.68).

Тоді ентропія сигналу $\eta(t)$:

$$\begin{aligned}
h(\eta) &= - \int_{-\infty}^{\infty} \omega_{\eta}(x) \log_2 \omega_{\eta}(x) dx = - \int_{(-c)} S_{-c} \delta(x+c) \log_2 S_{-c} \delta(x+c) dx - \\
&- \int_{-c}^{-c} \frac{1}{k} \omega_{\xi}\left(\frac{x}{k}\right) \log_2 \frac{1}{k} \omega_{\xi}\left(\frac{x}{k}\right) dx - \int_{(-c)} S_c \delta(x-c) \log_2 S_c \delta(x-c) dx.
\end{aligned}
\tag{8.80}$$

Так як перший інтеграл у формулі (8.80):

$$\begin{aligned}
J_1 &= S_{-c} \int_{(-c)} \delta(x+c) \log_2 S_{-c} \delta(x+c) dx = \\
&= S_{-c} \int_{(-c)} \delta(x+c) [\log_2 S_{-c} + \log_2 \delta(x+c)] dx = \\
&= S_{-c} \int_{(-c)} \delta(x+c) \log_2 S_{-c} dx = S_{-c} \log_2 S_{-c},
\end{aligned}$$

то за аналогією третій інтеграл:

$$J_3 = \int_{(-c)} S_c \delta(x-c) \log_2 S_c \delta(x-c) dx = S_c \log_2 S_c$$

Другий інтеграл у виразі (8.80) шляхом заміни $x/k = u$, $dx = kdu$ приводить до вигляду:

$$J_2 = \log_2 \frac{1}{k} \int_{-c/k}^{c/k} \omega_\xi(u) du + \int_{-c/k}^{c/k} \omega_\xi(u) \log_2 \omega_\xi(u) du.$$

Тоді

$$h(\eta) = (-J_1 + J_2 + J_3) = -S_{-c} \log_2 S_{-c} - \\ - \log_2 \frac{1}{k} \int_{-c/k}^{c/k} \omega_\xi(x) dx + \int_{-c/k}^{c/k} \omega_\xi(x) \log_2 \omega_\xi(x) dx - S_c \log_2 S_c$$

і ентропія сигналу $\eta(t)$, перерахована до входу нелінійності:

$$h(\eta) = -S_{-c} \log_2 S_{-c} - \int_{-c/k}^{c/k} \omega_\xi(x) \log_2 \omega_\xi(x) dx - S_c \log_2 S_c, \quad (8.81)$$

де $S_{-c} = W_\xi(-c/k)$, $S_c = 1 - W_\xi(c/k)$, $W_\xi(x)$ - інтегральна функція розподілу сигналу $\xi(t)$.

Перепишемо вираз (8.81) в наступному вигляді:

$$h(\eta) = -W_\xi(-c/k) \log_2 W_\xi(-c/k) - \int_{-c/k}^{c/k} \omega_\xi(x) \log_2 \omega_\xi(x) dx - \\ - [1 - W_\xi(c/k)] \log_2 [1 - W_\xi(c/k)].$$

Знайдемо втрати інформації в системі синхронізації, що зумовлені наявністю нелінійності типу зони насичення:

$$\Delta J = - \int_{-\infty}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + W_{\xi}(-c/k) \log_2 W_{\xi}(-c/k) + \\ + \int_{-c/k}^{c/k} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + [1 - W_{\xi}(c/k)] \log_2 [1 - W_{\xi}(c/k)]$$

або з врахуванням того, що

$$\int_{-\infty}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx = \int_{-\infty}^{-c/k} + \int_{-c/k}^{c/k} + \int_{c/k}^{\infty} ; \\ \Delta J = - \int_{-\infty}^{-c/k} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + W_{\xi}(-c/k) \log_2 W_{\xi}(-c/k) + \\ + \int_{c/k}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + [1 - W_{\xi}(c/k)] \log_2 [1 - W_{\xi}(c/k)].$$

(8.82)

3. Релейна характеристика.

Щільність сигналу $\eta(t)$ на виході нелінійності визначається виразом (8.70).

З врахуванням виразу (8.70) ентропія сигналу $\eta(t)$:

$$h(\eta) = - \int_{-\infty}^{\infty} \omega_{\eta}(x) \log_2 \omega_{\eta}(x) dx = \int_{(-c)} S_{-c} \delta(x+c) \log_2 S_{-c} \delta(x+c) dx \\ - \int_{(c)}^{(-c)} S_c \delta(x-c) \log_2 S_c \delta(x-c) dx = -S_{-c} \log_2 S_{-c} - S_c \log_2 S_c,$$

де

$$S_{-c} = \int_{-\infty}^0 \omega_{\xi}(x) dx = W_{\xi}(0),$$

$$S_c = \int_0^{\infty} \omega_{\xi}(x) dx = 1 - \int_{-\infty}^0 \omega_{\xi}(x) dx = 1 - W_{\xi}(0).$$

Запишемо:

$$\begin{aligned} h(\eta) &= -S_{-c} \log_2 S_{-c} - S_c \log_2 S_c = \\ &= -W_{\xi}(0) \log_2 W_{\xi}(0) - [1 - W_{\xi}(0)] \log_2 [1 - W_{\xi}(0)] \end{aligned}$$

Тоді втрати інформації:

$$\begin{aligned} \Delta J &= h(\xi) - h(\eta) = \\ &= - \int_{-\infty}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + W_{\xi}(0) \log_2 W_{\xi}(0) + [1 - W_{\xi}(0)] \log_2 [1 - W_{\xi}(0)]. \end{aligned} \quad (8.83)$$

Вираз (8.83) визначає втрати інформації, зумовлені наявністю характеристики на рис.8.9, а.

4. Нелінійність загального вигляду, що поєднує зону нечутливості і насичення.

Щільність розподілу імовірностей випадкового сигналу $\eta(t)$ на виході нелінійності визначається виразом (8.72).

З врахуванням виразу (8.72) ентропія сигналу $\eta(t)$:

$$\begin{aligned} h(\eta) &= - \int_{-\infty}^{\infty} \omega_{\eta}(x) \log_2 \omega_{\eta}(x) dx = \int_{(-c)} S_{-c} \delta(x+c) \log_2 S_{-c} \delta(x+c) dx - \\ &- \int_{-c}^0 \frac{1}{k} \omega_{\xi}\left(\frac{x - kb_H}{k}\right) \log_2 \frac{1}{k} \omega_{\xi}\left(\frac{x - kb_H}{k}\right) dx - \int_{(0)} S_0 \delta(x) \log_2 S_0 \delta(x) dx - \\ &- \int_0^c \frac{1}{k} \omega_{\xi}\left(\frac{x + kb_H}{k}\right) \log_2 \frac{1}{k} \omega_{\xi}\left(\frac{x + kb_H}{k}\right) dx - \int_{(c)} S_c \delta(x-c) \log_2 S_c \delta(x-c) dx, \end{aligned} \quad (8.84)$$

де

$$S_{-c} = \int_{-\infty}^{-b_H - c/k} \omega_{\xi}(x) dx; \quad S_0 = \int_{-b_H}^{b_H} \omega_{\xi}(x) dx;$$

$$S_c = \int_{b_H + c/k}^{\infty} \omega_{\xi}(x) dx.$$

(8.85)

У виразі (8.85):

$$J_1 = \int_{(-c)} S_{-c} \delta(x+c) \log_2 S_{-c} \delta(x+c) dx = S_{-c} \log_2 S_{-c};$$

$$J_2 = \int_{-c}^0 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_H}{k} \right) \log_2 \frac{1}{k} \omega_{\xi} \left(\frac{x - kb_H}{k} \right) dx = \left| \frac{x}{k} - b_H = u; \right.$$

$$\left. dx = kdu \right| = \int_{-b_H - c/k}^{-b_H} \omega_{\xi}(u) \log_2 \frac{1}{k} \omega_{\xi}(u) du;$$

$$J_3 = \int_{(0)} S_0 \delta(x) \log_2 S_0 \delta(x) dx = S_0 \log_2 S_0;$$

$$J_4 = \int_0^c \frac{1}{k} \omega_{\xi} \left(\frac{x + kb_H}{k} \right) \log_2 \frac{1}{k} \omega_{\xi} \left(\frac{x + kb_H}{k} \right) dx = \left| \frac{x}{k} + b_H = u, dx = kdu \right| =$$

$$= \int_{b_H}^{b_H + c/k} \omega_{\xi}(u) \log_2 \frac{1}{k} \omega_{\xi}(u) du;$$

$$J_5 = \int_{(-c)} S_c \delta(x-c) \log_2 S_c \delta(x-c) dx = S_c \log_2 S_c.$$

Тоді

$$\begin{aligned}
h(\eta) = & -(J_1 + J_2 + J_3 + J_4 + J_5) = -S_{-c} \log_2 S_{-c} - \\
& - \int_{-b_H - c/k}^{-b_H} \omega_\xi(u) \log_2 \frac{1}{k} \omega_\xi(u) du - S_0 \log_2 S_0 - \\
& - \int_{b_H}^{b_H + c/k} \omega_\xi(u) \log_2 \frac{1}{k} \omega_\xi(u) du - S_c \log_2 S_c.
\end{aligned} \tag{8.86}$$

Перетворивши інтеграли у виразі (8.86) отримаємо:

$$\begin{aligned}
h(\eta) = & -S_{-c} \log_2 S_{-c} - \int_{-b_i - c/k}^{-b_i} \omega_\xi(x) \log_2 \omega_\xi(x) dx - S_0 \log_2 S_0 - \\
& - \int_{b_i}^{b_i + c/k} \omega_\xi(x) \log_2 \omega_\xi(x) dx - S_c \log_2 S_c.
\end{aligned} \tag{8.87}$$

Із виразу (8.87) отримуємо:

$$\begin{aligned}
S_{-c} &= W_\xi(-b_i - c/k), \quad S_0 = W_\xi(b_i) - W_\xi(-b_i), \\
S_c &= 1 - W_\xi(b_i + c/k).
\end{aligned}$$

Перерахована ентропія вихідного сигналу приймає вигляд:

$$\begin{aligned}
h(\eta) = & -W_\xi(-b_H - c/k) \log_2 W_\xi(-b_H - c/k) - \\
& - \int_{-b_H - c/k}^{-b_H} \omega_\xi(x) \log_2 \omega_\xi(x) dx - \\
& - [W_\xi(b_H) - W_\xi(-b_H)] \log_2 [W_\xi(b_H) - W_\xi(-b_H)] - \\
& - \int_{b_H}^{b_H + c/k} \omega_\xi(x) \log_2 \omega_\xi(x) dx - [1 - W_\xi(b_H + c/k)] \log_2 [1 - W_\xi(b_H + c/k)]
\end{aligned} \tag{8.88}$$

Визначимо тепер із врахуванням виразу (8.88) втрати інформації:

$$\begin{aligned}
\Delta J &= h(\xi) - h(\eta) = - \int_{-\infty}^{\infty} \omega_{\xi}(x) \log_2(x) dx + \\
&+ W_{\xi}(-b_H - c/k) \log_2 W_{\xi}(-b_H - c/k) + \\
&+ \int_{-b_H - c/k}^{-b_H} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + [W_{\xi}(b_H) - W_{\xi}(-b_H)] \log_2 [W_{\xi}(b_H) - W_{\xi}(-b_H)] + \\
&+ \int_{b_H}^{b_H + c/k} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + [1 - W_{\xi}(b_H + c/k)] \log_2 [1 - W_{\xi}(b_H + c/k)] =
\end{aligned}$$

або

$$\begin{aligned}
\Delta J &= - \int_{-\infty}^{c/k - b_H} - \int_{-c/k - b_H}^{-b_H} - \int_{-b_H}^{b_H} - \int_{b_H}^{c/k + b_H} - \int_{c/k + b_H}^{\infty} + \\
&+ W_{\xi}(-b_H - c/k) \log_2 W_{\xi}(-b_H - c/k) + \\
&+ \int_{-b_H - c/k}^{-b_H} [W_{\xi}(b_H) - W_{\xi}(-b_H)] \log_2 [W_{\xi}(b_H) - W_{\xi}(-b_H)] + \\
&+ \int_{b_H}^{b_H + c/k} [1 - W_{\xi}(b_H + c/k)] \log_2 [1 - W_{\xi}(b_H + c/k)] =
\end{aligned}
\tag{8.89}$$

$$\begin{aligned}
&= \int_{-\infty}^{-c/k - b_H} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx - \int_{-b_H}^{b_H} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx - \int_{b_H + c/k}^{\infty} \omega_{\xi}(x) \log_2 \omega_{\xi}(x) dx + \\
&+ W_{\xi}(-b_H - c/k) \log_2 W_{\xi}(-b_H - c/k) + [W_{\xi}(b_H) - W_{\xi}(-b_H)] \log_2 [W_{\xi}(b_H) - W_{\xi}(-b_H)] + \\
&+ [1 - W_{\xi}(b_H + c/k)] \log_2 [1 - W_{\xi}(b_H + c/k)]
\end{aligned}$$

Втрати інформації в пристрої зв'язку, зумовлені наявністю нелінійності з характеристикою загального вигляду (рис.8.7, в), визначаються виразом (8.89).

Як частинні випадки, із виразу (8.89) можна отримати співвідношення (8.79) і (8.82), що визначають втрати інформації, зумовлені наявністю насичення і зони нечутливості відповідно.

8.10. Ефективність систем передачі інформації

Пропускна спроможність каналу зв'язку C визначає максимальну швидкість передачі інформації, тобто є тією границею, якої можна досягти при ідеальному кодуванні. Природно, що в реальних каналах швидкість передачі R завжди буде меншою від C . Степінь залежності R від C залежить від того, наскільки раціонально вибрана і ефективно використовується та чи інша система зв'язку. Найбільш загальною оцінкою ефективності системи зв'язку є коефіцієнт використання каналу:

$$\eta = \frac{R}{C}. \quad (8.90)$$

Для дискретних систем зв'язку можна записати $\eta = \eta_1 \eta_2$, де η_1 і η_2 - ефективність системи кодування і ефективність системи модуляції. Вводячи надлишковість повідомлення $X_1 = 1 - \eta_1$ та надлишковість сигналу $X_2 = 1 - \eta_2$, отримаємо:

$$\eta = 1 - X, \quad (8.91)$$

де $X = X_1 + X_2 - X_1 X_2$ - повна надлишковість системи.

При передачі неперервних повідомлень маємо:

$$\eta = \frac{F_m \log \left(\frac{P_c^*}{P_\emptyset^*} + 1 \right)}{F \log \left(\frac{P_c}{P_\emptyset} + 1 \right)}. \quad (8.92)$$

В ряді практичних випадків зручними оцінками ефективності зв'язку є коефіцієнт використання потужності сигналу:

$$\eta_P = \frac{R}{P_c / N_0}, \quad (8.93)$$

де N_0 - інтенсивність завади, і коефіцієнт використання смуги частот каналу:

$$\eta_F = \frac{R}{F}. \quad (8.94)$$

Коефіцієнт η_P є основним показником ефективності для тих систем, в яких потужність сигналу жорстко обмежена (наприклад, для космічних систем). В системах провідного зв'язку більш важливим показником є коефіцієнт η_F .

Згідно до виразу (8.91) для коефіцієнта використання каналу η ефективність системи зв'язку повністю визначається величиною її надлишковості. Звідси задача підвищення ефективності зв'язку зводиться до задачі зменшення надлишковості повідомлення та сигналу.

Надлишковість повідомлення, як ми бачили, зумовлена тим, що елементи повідомлення не є рівноймовірними, і між ними існує статистичний зв'язок. При кодуванні можна перерозподілити ймовірності вихідного повідомлення так, щоб розподіл ймовірностей символів коду наблизився до оптимального (до рівномірного в дискретному випадку або до нормального при передачі неперервних повідомлень). Такий перерозподіл дозволяє усунути надлишковість, що залежить від розподілу ймовірностей елементів повідомлення. Прикладом подібного кодування є код Шеннона-Фано, розглянутий раніше. Якщо перейти від кодування окремих символів повідомлення до кодування цілих груп символів, то можна усунути взаємозв'язок між ними і тим самим зменшити надлишковість. Загальна ідея такого методу кодування, який називають методом укрупнення, полягає в наступному. Вихідне повідомлення розбивається на відрізки по K символів в кожному. Такі відрізки можуть розглядатися як укрупнені елементи повідомлення. Можна показати, що ймовірності зв'язки між такими укрупненими елементами слабкіші, ніж між елементами вихідного повідомлення. Очевидно, що чим більше K (крупніші відрізки), тим слабшим буде зв'язок між ними. Далі укрупнені елементи кодуються з врахуванням їх розподілу ймовірностей.

Слід відмітити, що при укрупненні елементів відбувається перетворення, що полягає в переході до коду з більш високою основою $m_1 = m^k$, де m - початковий стан.

Своєрідним прикладом методу укрупнення повідомлень є стенографічний текст. Кожний стенографічний знак в цьому тексті виражає ціле слово або навіть групу слів.

Що стосується сигналу, то його надлишковість залежить від способу модуляції та від виду переносника. Процес модуляції зазвичай супроводжується розширенням смуги частот сигналу в порівнянні зі смугою частот повідомлення, яке передається. Це розширення смуги і є надлишковим. Частотна надлишковість також збільшується при переході від синусоїдального переносника до переносника імпульсного чи шумоподібного. З точки зору підвищення ефективності передачі слід було б вибирати такі способи модуляції, які мають малу надлишковість. До таких систем, частково, відноситься односмугова передача, при якій сигнали, що передаються, не містять частотної надлишковості – вони є просто копіями повідомлень, які передаються.

Однак, говорячи про ефективність системи зв'язку, не можна забувати про її завадостійкість. Усунення надлишковості підвищує ефективність передачі, але знижує при цьому достовірність (завадостійкість) і, навпаки, збереження чи введення надлишковості дозволяє забезпечити високу достовірність передачі. Наприклад, при телеграфній передачі тексту усунення надлишковості приводить до того, що стає важче виправляти помилки в повідомленні, і в кінці-кінців знижується завадостійкість. При збереженні надлишковості в тексті завадостійкість буде вищою.

При кодуванні в ряді випадків надлишковість навмисно вводиться з метою підвищення достовірності передачі. Прикладом такого кодування є коректуючі коди.

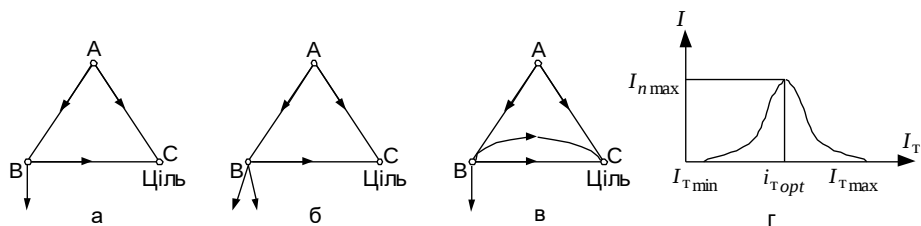
Аналогічна ситуація має місце й у відношенні надлишковості сигналу. Частотна надлишковість при різних видах модуляції використовується по-різному. При частотній модуляції, наприклад, ми можемо отримати більший виграш в завадостійкості, ніж при амплітудній модуляції, а при кодовій імпульсній модуляції цей виграш ще більший. Частотна надлишковість шумового переносника дозволяє знизити вплив завади і зосереджених завад.

Отже, при оцінці різних систем зв'язку необхідно враховувати, в крайньому випадку, два показника: ефективність і завадостійкість, сукупність цих двох показників складає достатньо повну характеристику системи.

Найбільш досконалою системою вважається така, яка забезпечує найбільшу ефективність при заданій завадостійкості або, навпаки, забезпечує найбільшу завадостійкість при заданій ефективності [2].

Контрольні питання

1. Який критерій використовується для кількісної міри оцінки інформації? Чому застосовується логарифмічна міра кількості інформації?
2. Що називається ентропією і як вона визначається для джерел незалежних і залежних повідомлень?
3. Які причини появи надлишковості в повідомленні?
4. Які джерела повідомлення називаються стаціонарними і ергодичними?
5. Що називається пропускнуою спроможністю каналу? Чому вона рівна для двійкового каналу без завад?
6. Поясніть принцип оптимального статистичного кодування.
7. Як визначається швидкість передачі і пропускна спроможність каналу з завадами?
8. Які висновки випливають із теореми Шеннона для дискретного каналу з завадами?
9. Як визначається кількість інформації, що передається в неперервних каналах?
10. Проаналізуйте формулу Шеннона для пропускнуої спроможності неперервного каналу.
11. Сформулюйте теорему Шеннона для неперервного каналу і поясніть її зміст.
12. Як обчислюється пропускна спроможність каналів зі змінними параметрами?
13. Що називається ефективністю системи зв'язку і як вона визначається кількісно?
14. Поясніть залежність між завадостійкістю і ефективністю зв'язку. Наведіть приклади.
15. Як впливають нелінійності в каналі зв'язку на передачу інформації?
16. Що таке епсілон-ентропія?
17. В чому полягає структурна міра інформації?
18. Як визначити коефіцієнт стиснення?
19. В чому полягають особливості оцінки інформації?



д

Рис. 8.1. До визначення цінності інформації:
а, б, в - шляхи досягнення мети; г - кількість інформації в тезаурусі;
д - система оцінки цінності за Р.Л. Стратоновичем

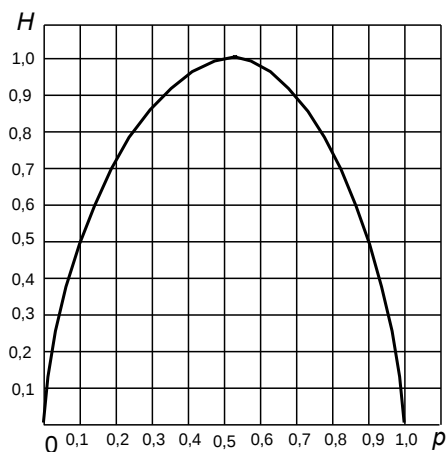


Рис. 8.2. Залежність ентропії від імовірності p

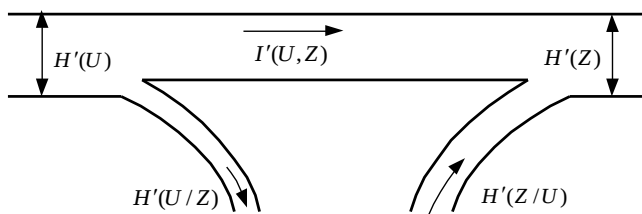


Рис. 8.3. Ілюстрація передачі інформації по каналу з завадами

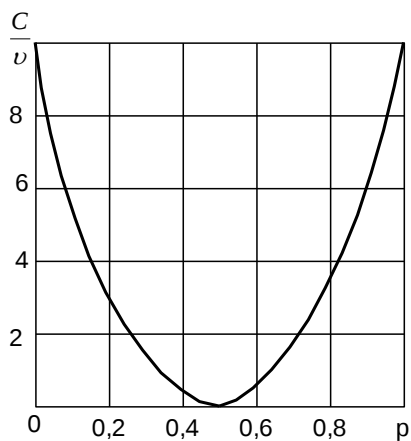


Рис. 8.4. Залежність пропускної здатності двійкового симетричного каналу без пам'яті від імовірності помилкового прийому символу

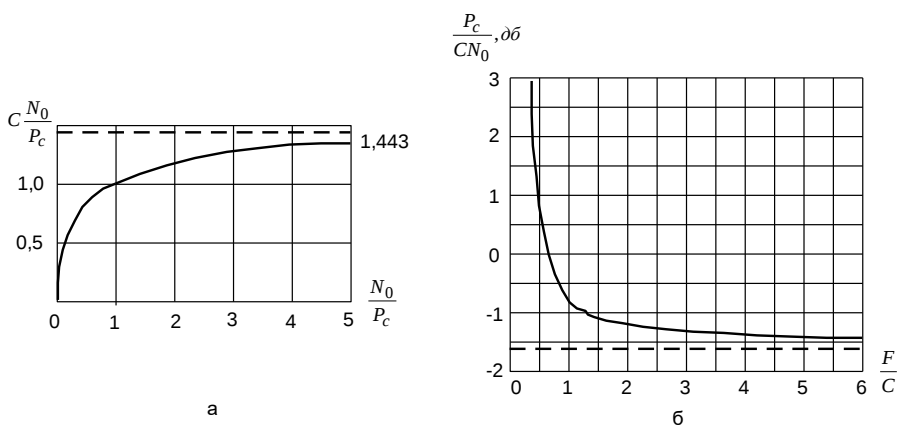


Рис. 8.5. Залежність пропускної здатності C від ширини перепуску F в неперервному каналі:

$$a - C \frac{N_0}{P_c} = f\left(\frac{N_0}{P_c}\right); \quad б - \frac{P_c}{C N_0} = f\left(\frac{N_0}{P_c}\right)$$

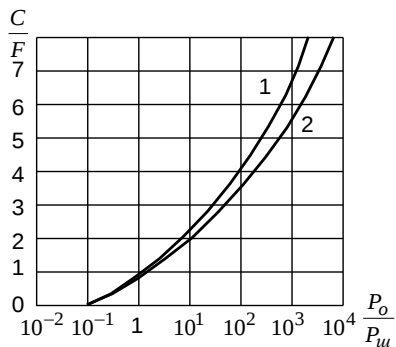


Рис. 8.6. Залежність пропускної здатності від середнього відношення сигналу до завади для каналів з постійними параметрами (1) і релеєвськими завмираннями (2)

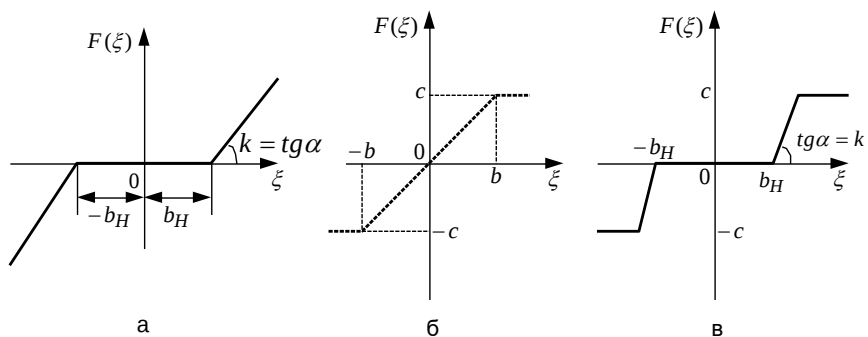


Рис. 8.7. Типові нелінійні статичні характеристики: а - зона нечутливості; б - зона насичення; в - зона загального вигляду.

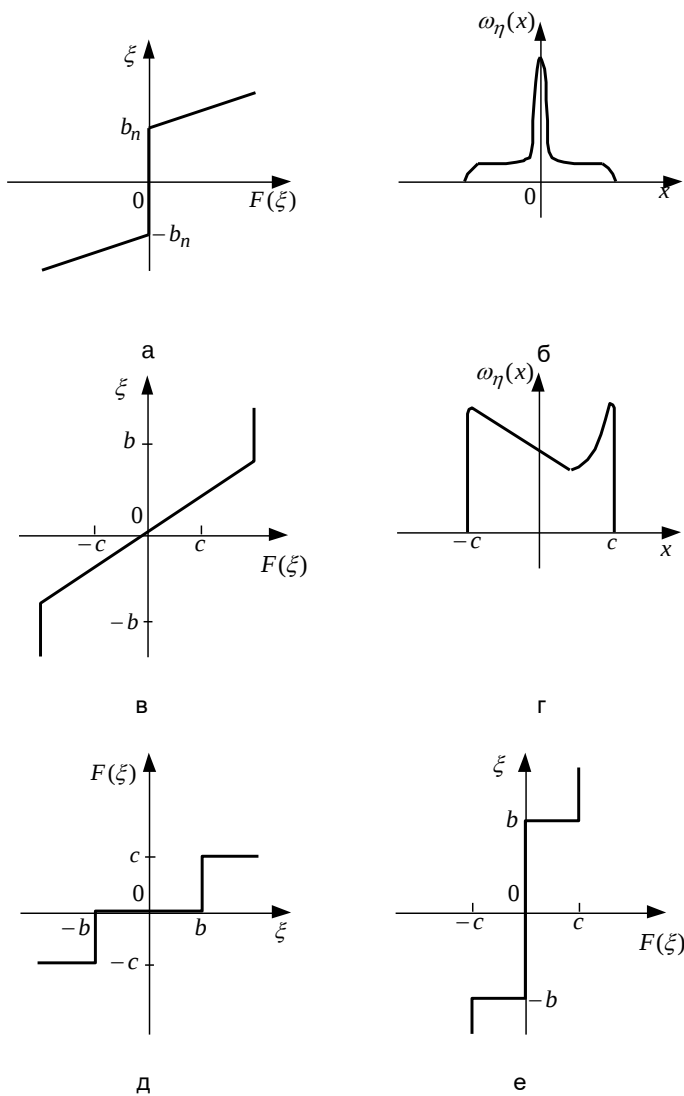


Рис. 8.8. Зворотні статичні характеристики нелінійностей (а, в, д) та відповідні щільності вихідного сигналу (б, г, е)

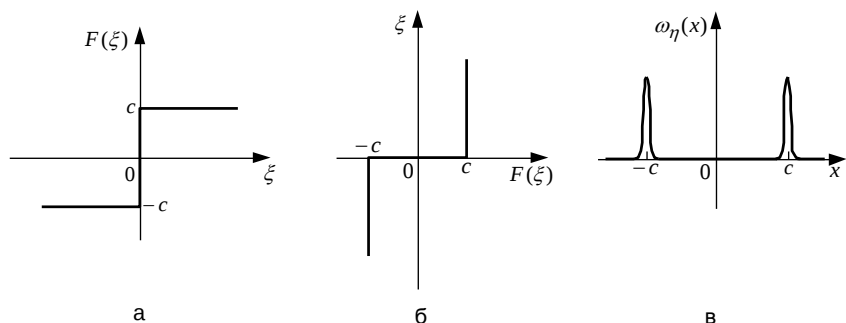


Рис. 8.9. Статична релейна характеристика (а), зворотня статична релейна характеристика (б) та щільність вихідного сигналу (в)

РОЗДІЛ 9. ОПТИМАЛЬНИЙ ПРИЙОМ НЕПЕРЕРВНИХ ПОВІДОМЛЕНЬ. ОПТИМАЛЬНА ФІЛЬТРАЦІЯ

9.1. Оптимальний прийом неперервних повідомлень

Нехай повідомлення являє собою деяке коливання $u(t)$, що може безперервно змінюватися з часом і приймати будь-яку форму. З такими повідомленнями ми зустрічаємося, наприклад, у телефонії, телеметрії, телебаченні. Визначасмо умови оптимального прийому безперервних повідомлень.

Для простоти аналізу будемо вважати, що функція $u(t)$ приймає значення в межах від $+1$ до -1 . Будемо також думати, що передане коливання є періодичним з періодом T (це завжди можна допустити, узявши T досить великим) і що спектр цього коливання обмежений частотами від i_1/T до i_2/T . При цих умовах функцію $u(t)$ можна представити у вигляді

$$u(t) = \sum_{i=i_1}^{i_2} \left(\lambda_{2i-1} \sqrt{2} \sin \frac{2\pi i}{T} t + \lambda_{2i} \sqrt{2} \cos \frac{2\pi i}{T} t \right) = \sum_{e=e_1}^{e_2} \lambda_e \varphi_e(t) \quad (9.1)$$

де λ_e – деякі параметри, що визначають передане коливання $u(t)$, $\varphi_e(t)$ – одиничні ортогональні функції, $e_1 = 2i_1 - 1$, $e_2 = 2i_2$.

Для передачі по каналу коливання $u(t)$ перетворюється в сигнал $s(u, t)$. Оскільки коливання (9.1) визначається $e_2 - e_1 + 1$ параметрами λ_1 , то сигнал буде залежати від цих параметрів: $s(u, t) = s(\lambda_{e_1}, \dots, \lambda_{e_2}, t)$. Прийнятий сигнал внаслідок накладання завади $\omega(t)$ дорівнює $x(t) = s(u, t) + \omega(t)$.

Задача полягає в тому, щоб за коливанням $x(t)$ відновити передане повідомлення $u(t)$ з якомога більшою точністю. Як ми вже встановили, найбільше, що може зробити приймач на основі аналізу прийнятого сигналу $x(t)$ це обчислити розподіл імовірностей $P(s/x)$ для всіх можливих реалізацій $u(t)$. При передачі безперервних коливань функція $P(s/x)$ є щільністю розподілу ймовірності.

Згідно (5.2) $P(s_i/x) \supset P(s_j/x)$, $j \neq i$ оптимальний приймач обчислює апостеріорний розподіл імовірностей $P(s/x)$ і видає на виході ту реалізацію повідомлення $u(t)$, при якій функція $P(s/x)$ буде максимальною. При цьому приймач не обов'язково повинний визначати апостеріорний розподіл імовірностей $P(s/x)$ у очевидному вигляді – він повинний видавати на виході результат, еквівалентний до цієї функції.

Відповідно до формули Байєса (5.3) вираз $P(s/x)$ для розглянутого випадку можна записати в наступному вигляді:

$$P(s/x) = kP(s)P(x/s), \quad (9.2)$$

де k -деяка постійна, яка може бути обчислена з умови нормування $\int_s P(s/x)dx = 1$.

Припускаємо, що всі можливі повідомлення $u(t)$ та відповідні їм сигнали $s(u, t)$ рівноімовірні, тобто $P(s) = \text{const}$ для всіх реалізацій $u(t)$, що лежать в інтервалі $(1, +1)$. При цьому відповідно до виразу

$$P(x / s_i) = \prod_{i=l_1}^{l_2} P(x_i / s_{il}) = (\pi N_0)^{-\frac{n}{2}} e^{\left[-\frac{1}{N_0} \sum_{i=l_1}^{l_2} (x_i - s_{il}) \right]}$$

Із виразу (9.2) маємо:

$$P(s / x) = k_1 P(s) \exp \left\{ -\frac{1}{N_0} \int_0^T [x(t) - s(u, t)]^2 dt \right\}. \quad (9.3)$$

Звідси максимуму апостеріорної ймовірності $P(s / x)$ відповідає мінімум за $u(t)$ величини:

$$\Delta^2 = \int_0^T [x(t) - s(u, t)]^2 dt. \quad (9.4)$$

Таким чином, оптимальний приймач повинен відтворювати повідомлення $u(t)$, при якому середньоквадратичне відхилення Δ^2 має мінімальне значення. За відсутності завад такий приймач відтворює повідомлення без спотворень (без помилок): $x(t) = s(u, t)$, $v(t) = u(t)$ і $\Delta^2 = 0$, а при наявності завад помилка буде мінімальною.

З виразу (9.4) випливає, що операції оптимальної фільтрації та детектування дають достатнє рішення задачі про витягнення максимальної інформації з прийнятого сигналу $x(t)$ щодо переданого повідомлення $u(t)$. Структурна схема приймача з оптимальним фільтром приведена на рис.9.1,а.

Запишемо вираз в іншому вигляді:

$$P(s / x) = k_1 P(s) \exp \left[-\frac{1}{N_0} \int_0^T x^2(t) dt \right] \exp \left[-\frac{1}{N_0} \int_0^T s^2(t) dt \right] \exp \left[\frac{2}{N_0} \int_0^T s(u, t) x(t) dt \right] \quad (9.5)$$

Перший експоненціальний множник, що не залежить від $\mathbf{U}$, може бути включений у постійну k_1 . Другий множник взагалі не представляє операції над $x(t)$, він може бути внесений у виді окремого множника, подібного до апіорної імовірності. В багатьох випадках цей множник, рівний $\exp(-E / N_0)$, де E – енергія сигналу, також можна включити в постійну k_1 . Отже, вираз (9.5) можна записати:

$$P(s / x) = k P(s) \exp[h(u)], \quad (9.6)$$

де

$$h(u) = \frac{2}{N_0} \int_0^T x(t) s(u, t) dt, \quad (9.7)$$

Звідси випливає, що при відомій апіорній імовірності визначення апостеріорної ймовірності зводиться до обчислення функції $h(u)$, тобто до обчислення взаємної кореляції між прийнятим сигналом $x(t)$ і переданим (очікуваним) сигналом $s(u, t)$. Іншими словами, кореляційна обробка сигналу є оптимальною.

Функція $h(u)$ легко обчислюється, коли сигнал є повністю відомим. Досягається це за допомогою корелятора або погодженого фільтра. Така ситуація звичайно має місце при передачі дискретних повідомлень.

При передачі безперервних повідомлень сигнал $s(u, t)$ не є повністю відомим. Однак ми маємо деяку апіорну інформацію про цей сигнал. Відомі, наприклад, несуча частота, вид модуляції, ширина спектра сигналу і т.п. Відсутню частину інформації можна одержати в результаті спостереження над прийнятою реалізацією сигналу $x(t)$ за попередній проміжок часу. В результаті існує можливість на прийомі визначити оцінку сигналу $s(v, t)$ й обчислити функцію $h(v)$ для цієї оцінки:

$$h(v) = \frac{2}{N_0} \int_0^T x(t) s(v, t) dt. \quad (9.8)$$

Функцію $h(v)$ можна обчислити за допомогою фільтра зі змінними параметрами (рис.9.1,б) чи схеми слідкуючого корелятора (рис.9.1,в). Кожна з цих схем має основний інформаційний канал, на виході якого виходить оціночне значення $v(t)$ переданого повідомлення $u(t)$, і канал зворотнього зв'язку, за допомогою якого в схемі на рис.9.1, в формується опорний сигнал $s(v,t)$, а в схемі рис.9.1,б проводиться зміна параметрів фільтра. В схемі рис.9.1,б за допомогою управляючого елемента (УЕ) змінюють параметри фільтра так, щоб він був погоджений з безперервно змінюючимся очікуваним сигналом $s(t,v)$. В схемі ж рис.9.1,в за допомогою УЕ змінюється модульований параметр несучого коливання, формованого генератором Г. При частотній модуляції, наприклад, цим параметром є частота, при час-імпульсній модуляції – часовий зсув імпульсів і т.п. Фільтр нижніх частот (ФНЧ) у цій схемі виконує роль інтегратора на інтервалі спостереження Т, який зв'язаний з максимальною частотою F_m в спектрі переданого повідомлення співвідношенням $T = 1/2F_m$.

При різних видах модуляції принцип слідкуючого прийому залишається тим же. Вид модуляції визначає параметр, за яким повинне здійснюватися слідкування. Схеми слідкуючого прийому дозволяють практично реалізувати завадостійкість, близьку до потенційної, і тим самим знизити поріг завадостійкості до його граничного значення.

До цих пір ми припускали, що амплітуда і фаза несучого коливання сигналу незмінні за часом, а завада має постійну спектральну щільність N_0 . В каналах зі змінними параметрами амплітуда й фаза сигналу змінюються за часом. У цьому випадку схеми рис.9.1,б і рис.9.1,в повинні бути доповнені пристроєм автоматичного регулювання рівня сигналу й системного фазового автопідстроювання. При невідомому значенні N_0 чи його зміні в часі для здійснення оптимального прийому необхідний спеціальний пристрій вимірювання й вирівнювання інтенсивності завад. Зокрема, для зосереджених за спектром чи за часом завад з різко вираженою інтенсивністю, досить близькою до оптимальної, є схема зі стиранням ділянок спектру чи тривалості сигналу, вражених завадою.

Таким чином, у загальному випадку при оптимальному прийомі необхідно здійснювати спостереження за формою як сигналу, так і завади. Чим більше число параметрів при цьому буде враховано, тим вищою буде завадостійкість. При цьому схема приймача повинна бути адаптивною.

Оптимальний приймач являє собою пристрій, який відтворює передане повідомлення $u(t)$ з найменшою помилкою. Так як корисний сигнал $s(u,t)$ залежить від $u(t)$ нелінійно, то оптимальний приймач також повинен бути нелінійним пристроєм (нелінійним фільтром). Прикладом такого пристрою (фільтра) є розглянутий вище

слідкуючий приймач. Таким чином, теорію оптимального прийому можна розглядати як теорію оптимальної нелінійної фільтрації.

Загальна теорія нелінійної фільтрації охоплює різноманітні й дуже складні задачі. В даний час ця теорія розроблена лише для випадків, коли передане повідомлення $u(t)$ являє собою марківський чи нормальний випадковий процес.

9.2. Синтез пристроїв зв'язку при заданій структурній схемі

Задача синтезу пристрою зв'язку при заданій його структурній схемі полягає в тому, що при відомих характеристиках сигналу й завади потрібно визначити оптимальні значення параметрів пристрою зв'язку, при яких сумарна середня квадратична помилка мінімальна. Так як структурна схема пристрою відома, то неважко знайти вираз для сумарної середньої квадратичної помилки як функцію параметрів пристрою:

$$\sigma_e^2 = f(\beta_1, \beta_2, \dots, \beta_n), \quad (9.9)$$

де β_i - параметри пристрою.

Для розрахунку оптимальних значень параметрів необхідно дослідити на мінімум функцію (9.9). Для цього необхідно розв'язати наступну систему рівнянь:

$$\frac{\partial}{\partial \beta_i} \sigma_e^2 = 0, \quad i = \overline{1, n}. \quad (9.10)$$

Зазвичай мінімум сумарної середньоквадратичної помилки при оптимальних значеннях параметрів впливає з аналізу фізичного змісту задачі, тому немає необхідності в обчисленні другої похідної. Якщо друга похідна позитивна, то при знайдених оптимальних параметрах має місце мінімум сумарної середньої квадратичної помилки пристрою зв'язку.

В процесі синтезу пристрою зв'язку зазвичай вимагається обчислити оптимальне значення тільки настроюваних параметрів (коефіцієнт підсилення та одна чи дві постійних часу коректуючого пристрою) пристрою, число яких в пристроях невелике, що полегшує задачу оптимізації.

Приклад 9.1. Знайти оптимальне значення коефіцієнта підсилення пристрою зв'язку, передаточна функція якого в розімкнутому стані:

$$W_p(s) = \frac{K}{(Ts + 1)s}. \quad (9.11)$$

Розв'язок. На пристрої діють сигнал і завада, спектральні щільності яких

$S_x(\omega) = N_x / \omega^2$; $S_i(\omega) = N_i$. Сумарна середня квадратична помилка системи:

$$\sigma_e^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[|W_e(j\omega)W_\delta(j\omega)|^2 N_x + |W_3(j\omega)|^2 N_i \right] d\omega = N_x \frac{1+KT}{2K} + N_i \frac{K}{2},$$

де $W_e(j\omega) = \frac{j\omega(1+j\omega T)}{(j\omega)^2 T + j\omega + K}$ - частотна характеристика

помилки системи; $W_\delta(j\omega) = 1/(j\omega)$ - частотна характеристика формуючого

фільтру сигналу; $W_3(j\omega) = \frac{K}{(j\omega)^2 T + j\omega + K}$ частотна характеристика

замкнутої системи.

Оптимальне значення коефіцієнта підсилення визначимо із умови (9.10), яка для даної задачі має вигляд:

$$\frac{\partial}{\partial K} \sigma_e^2 = -\frac{N_x}{2K^2} + \frac{N_i}{2} = 0.$$

Тоді оптимальний коефіцієнт підсилення $K_{\text{opt}} = \sqrt{N_x / N_i}$.

На рис.9.2,а показані криві зміни складових середньої квадратичної помилки пристрою зв'язку в залежності від коефіцієнта підсилення для $N_x = 0.76 \cdot 10^{-4}$ рад²·с; $N_i = 0.305 \cdot 10^{-4}$ рад²·с; $T = 0.1$ с.

9.3. Фізичний зміст оптимальної смуги перепуску

Розглянемо пристрій зв'язку з частотною характеристикою (рис. 9.2,б):

$$W(j\omega) = \begin{cases} e^{j0} & \text{їдє } \omega \leq \omega_i; \\ 0 & \text{їдє } \omega > \omega_i, \end{cases} \quad (9.12)$$

де ω_i - ширина смуги перепуску пристрою зв'язку.

Якщо сигнал і завада некорельовані, то дисперсія помилки системи з частотною характеристикою (9.12) визначається виразом:

$$\sigma_e^2 = \frac{1}{\pi} \int_{\omega_{\text{г}}}^{\infty} S_x(\omega) d\omega + \frac{1}{\pi} \int_0^{\omega_{\text{г}}} S_{\text{г}}(\omega) d\omega = \sigma_{ex}^2 + \sigma_{en}^2, \quad (9.13)$$

де σ_e - середня квадратична помилка; $S_x(\omega)$ - спектральна щільність сигналу; $S_{\text{г}}(\omega)$ - спектральна щільність завади; σ_{ex} - середня квадратична помилка відносно сигналу; $\sigma_{e\text{г}}$ - середня квадратична помилка через дію завади.

Із виразу (9.13) випливає, що середня квадратична помилка пристрою зв'язку залежить від ширини смуги перепуску. Смугою перепуску пристрою зв'язку, при якій середня квадратична помилка приймає мінімальне значення, називають *оптимальною*.

На рис.9.2,б показані графіки спектральних щільностей сигналу й завади. Так як ці графіки не перекриваються, то оптимальна смуга перепуску системи рівна граничній частоті спектра сигналу. В цьому випадку середня квадратична помилка пристрою зв'язку рівна нулю, так як всі складові спектра сигналу відтворюються системою, і ні одна складова спектру завади не проходить на її вихід.

Якщо графіки спектральних щільностей сигналу й завади перекриваються, то обидві складові середньої квадратичної помилки системи у виразі (9.13) не дорівнюють нулю. Із рис.9.3,а видно, що перша складова (середня квадратична помилка відтворення сигналу) визначається тією частиною спектральної щільності сигналу, яка розташована за смугою перепуску системи (площа 1 під графіком спектральної щільності сигналу). З розширенням смуги перепуску ця складова помилки зменшується (рис.9.3,б). Друга складова середньої квадратичної помилки системи, зумовлена завадою, залежить від тієї частини спектральної щільності завади, яка співпадає зі смугою перепуску системи (площа 2 на рис.9.3,а). З розширенням смуги перепуску середньоквадратична помилка через дію завади зростає (рис.9.3,б). Оптимальна смуга перепуску системи відповідає мінімальній середньоквадратичній помилці. Для її обчислення про диференціюємо

вираз (9.13) за смугою перепуску $\omega_{\text{г}}$ і отриманий результат прирівняємо до нуля. В результаті отримаємо:

$$-\frac{1}{\pi} S_x(\omega_{\text{г}}) + \frac{1}{\pi} S_{\text{г}}(\omega_{\text{г}}) = 0. \quad (9.14)$$

Із цього виразу випливає, що оптимальна смуга перепуску системи – це частота, на якій виконується рівність:

$$S_x(\omega_{\text{г}}) = S_{\text{г}}(\omega_{\text{г}}). \quad (9.15)$$

Так як частотні характеристики систем зв'язку відрізняються від ідеальної характеристики (9.12), то оптимальна смуга перепуску, знайдена із виразу (9.15), виходить наближеною. Як показує практика, помилки при цьому не перевищують 10%-20%.

Мінімум сумарної середньої квадратичної помилки системи визначається не тільки смугою перепуску, але й виглядом її частотної характеристики. Тому в загальному випадку синтез системи полягає в знаходженні її оптимальних частотних характеристик із умови мінімуму середніх квадратичних помилок при заданих статистичних характеристиках сигналу і завади.

9.4. Синтез оптимальних пристроїв (систем) зв'язку

Сформулюємо постановку задачі синтезу оптимальних пристроїв (систем) зв'язку. На вхід пристрою, який проектується, діють стаціонарні сигнал і завада, автокореляційні функції яких відомі. Математичні очікування сигналу й завади рівні нулю. Бажаний вихідний сигнал пристрою зв'язку, що синтезується, визначається заданою частотною характеристикою. Необхідно знайти передаточну функцію пристрою, при якій сумарна середня квадратична помилка системи (рис.9.1,а) мінімальна:

$$\sigma_e = \sqrt{e^2(t)} = \min, \quad (9.16)$$

де $e(t) = y_{\text{ж}}(t) - y(t)$ - помилка системи; $y_{\text{ж}}(t)$ - бажаний вихідний сигнал системи; $y(t)$ - вихідний сигнал системи.

Розглянемо задачу оптимальної фільтрації. В цьому випадку:

$$W_{\text{ж}}(j\omega) = 1, \text{ тобто } y_{\text{ж}}(t) = x(t). \quad (9.17)$$

Згідно (9.16), дисперсія помилки:

$$\sigma^2 = \overline{[x(t) - y(t)]^2} = \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy}^2. \quad (9.18)$$

Дисперсію вихідного сигналу пристрою зв'язку, що синтезується, σ_y^2 знайдемо аналогічно дисперсії помилки системи в

$$\sigma_y^2 = \int_{-\infty}^{\infty} \omega(\lambda) \int_{-\infty}^{\infty} \omega(\eta) R_f(\lambda - \eta) d\eta d\lambda, \quad (9.19)$$

де $\omega(t)$ - імпульсна перехідна функція системи; $R_f(\tau)$ - автокореляційна функція сумарного вхідного сигналу $f(t) = x(t) + n(t)$; $n(t)$ - завада.

Таким чином отримаємо:

$$\sigma_{xy}^2 = \int_{-\infty}^{\infty} \omega(\eta) R_{xf}(\lambda) d\lambda, \quad (9.20)$$

де $R_{xf}(\tau)$ - взаємна кореляційна функція сигналу з сумарною дією.

Підставивши вирази (9.19) і (9.20) в (9.18) визначимо:

$$\sigma_e^2 = \sigma_x^2 + \int_{-\infty}^{\infty} \omega(\lambda) \int_{-\infty}^{\infty} \omega(\eta) R_f(\lambda - \eta) d\eta d\lambda - 2 \int_{-\infty}^{\infty} \omega(\eta) R_{xf}(\lambda) d\lambda. \quad (9.21)$$

Синтез оптимальної системи зводиться до знаходження імпульсної перехідної функції з (9.21). Для рішення цієї задачі дамо варіацію імпульсної перехідної функції:

$$\omega(t) = \omega_0(t) + \delta\omega(t), \quad (9.22)$$

де $\omega_0(t)$ - шукана оптимальна імпульсна перехідна функція пристрою зв'язку, що проектується; δ - варіація імпульсної перехідної функції.

Імпульсну перехідну функцію, мінімізуючи дисперсію помилки (9.21), визначимо із умови:

$$\left. \frac{\partial \sigma^2}{\partial \delta} \right|_{\delta=0} = 0. \quad (9.23)$$

Підставимо формулу (9.22) в (9.21). Тоді оптимальна імпульсна перехідна функція системи з врахуванням (9.23) повинна задовольняти рівнянню:

$$2 \int_{-\infty}^{\infty} \omega(\lambda) \left[\int_{-\infty}^{\infty} \omega_0(\eta) R_f(\lambda - \eta) d\eta - R_{xf}(\lambda) \right] d\lambda = 0. \quad (9.24)$$

Так як неоптимальна імпульсна перехідна функція $\omega(t)$ - функція довільна, то рівняння (9.24) виконується тільки в тому випадку, коли:

$$\int_{-\infty}^{\infty} \omega_0(\eta) R_f(t - \eta) d\eta - R_{xf}(t) = 0. \quad (9.25)$$

де змінна λ замінена на t .

Вираз (9.25) називають *рівнянням Вінера-Хопфа*.

Середня квадратична помилка виділення сигналу із дії в установленому режимі – постійна величина, її значення визначається із виразу (9.21), в якому замість $\omega(t)$ необхідно підставити $\omega_0(t)$. Тоді з врахуванням рівняння (9.25):

$$\sigma_{e \min}^2 = \sigma_x^2 + \int_{-\infty}^{\infty} \omega_0(\lambda) \int_{-\infty}^{\infty} \omega_0(\eta) R_f(\lambda - \eta) d\eta d\lambda. \quad (9.26)$$

Рішення рівняння Вінера-Хопфа в часовій області являється складною задачею. Значно простіше розв'язати цю задачу в частотній області, тобто знайти оптимальну частотну характеристику пристрою зв'язку. З цієї метою застосуємо до рівняння (9.25) перетворення Фур'є. В результаті отримаємо:

$$W_0(j\omega) S(j\omega) = S_{xf}(\omega). \quad (9.27)$$

З цього рівняння знайдемо оптимальну частотну характеристику:

$$W_0(j\omega) = \frac{S_{xf}(\omega)}{S_f(\omega)} = \frac{S_{xf}(\omega)}{\psi(j\omega)\psi(-j\omega)}, \quad (9.28)$$

де $S_f(\omega)$ - спектральна щільність сумарного сигналу на вході системи, яка проектується; $S_{xf}(\omega)$ - взаємна спектральна щільність сигналу з сумарним сигналом; $\psi(j\omega)$ - функція, всі полюси якої на площині комплексного змінного $p = j\omega$ розташовані в лівій напівплощині; $\psi(-j\omega)$ - функція, всі полюси якої розташовані в правій напівплощині.

В загальному випадку:

$$\begin{aligned} S_{xf}(\omega) &= S_x(\omega) + S_{x\bar{i}}(\omega); \\ S_f(\omega) &= S_x(\omega) + S_{x\bar{i}}(\omega) + S_{\bar{i}}(\omega) + S_{i\bar{x}}(\omega). \end{aligned} \quad (9.29)$$

Якщо сигнал і завада некорельовані, то:

$$\begin{aligned} S_{xf}(\omega) &= S_x(\omega); \\ S_f(\omega) &= S_x(\omega) + S_i(\omega). \end{aligned} \quad (9.30)$$

Із виразів (9.28) – (9.30) випливає, що оптимальна частотна характеристика виділяє складові сигналу на частотах, на яких його спектральна щільність достатньо велика, і послаблює складові сигналу на частотах з максимальною спектральною щільністю завади.

Із врахуванням оптимальної частотної характеристики мінімальне значення дисперсії помилки для некорельованих сигналу й завади має вигляд:

$$\sigma_{e\min}^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{S_x(\omega)S_i(\omega)}{S_x(\omega) + S_i(\omega)} d\omega. \quad (9.31)$$

Таким чином, якщо спектри сигналу й завади не перекриваються, то середня квадратична помилка може бути рівною нулю.

Спектральна щільність є парною функцією відносно частоти ω , тому полюси характеристики (9.28) розташовані на площині $p = j\omega$ як зліва, так і справа від уявної осі. Тому знайдена оптимальна частотна характеристика (9.28) відповідає фізично нереалізовуваній нестійкій системі. Надалі синтез оптимальної системи зводиться до визначення реалізуємої оптимальної частотної характеристики, найбільш близької до отриманої нереалізуємої.

Реалізуєма оптимальна частотна характеристика визначається виразом:

$$W_0(j\omega) = \frac{1}{\psi(j\omega)} \left[\frac{S_{xf}(\omega)}{\psi(-j\omega)} \right]^+, \quad (9.32)$$

де операція $[\dots]^+$ визначає виділення доданків, полюси яких на площині комплексного змінного $p = j\omega$ розташовані зліва від уявної осі.

У випадку, що зустрічається в інженерній практиці, коли функція

$$\frac{S_{xf}(\omega)}{\psi(-j\omega)} \quad (9.33)$$

являється дрібно раціональною функцією відносно частоти, виділення доданків з полюсами, розташованими зліва від уявної осі, здійснюється шляхом розкладу (9.33) на прості дробі:

$$\frac{S_{xf}(\omega)}{\psi(-j\omega)} = \sum_{i=1}^n \frac{\alpha_i}{\omega - \omega_i}.$$

(9.34)

де

$$\alpha_i = \left(\omega - \frac{1}{\lambda_i} \right) \frac{S_{xf}(\omega)}{\psi(-j\omega)};$$

(9.35)

λ_i - полюси виразу (9.33).

Дисперсія помилки в системі з частотною характеристикою (9.32) буде більшою від значення, яке визначається за виразом (9.31), і знайденого без врахування фізичної реалізації оптимальної системи. Однак серед реалізовуваних пристроїв зв'язку система з характеристикою (9.32) забезпечує найменше значення середньоквадратичної помилки виділення сигналу з його суміші з завадою.

Згідно (9.32) передаточна функція оптимальної системи, яка проектується:

$$W_0(s) = W_0(j\omega)|_{j\omega=s}.$$

(9.36)

Тоді передаточна функція послідовного коректуючого пристрою, що вмикається в ланку сигналу помилки:

$$W_k(s) = W_{po}(s) / W_{pu}(s),$$

(9.37)

де $W_{po}(s)$ - передаточна функція розімкненої оптимальної системи; $W_{pu}(s)$ - передаточна функція вихідної частини, складена із функціонально необхідних елементів пристрою (системи).

Приклад 9.2. Знайти передаточну функцію послідовного коректуючого пристрою в оптимальній системі автопідстроювання частоти, структурна схема якої зображена на рис.9.4,а, для випадку, коли спектральні щільності сигналу й завади визначаються виразами:

$$S_x(\omega) = \frac{N_x}{T_x^2 \omega^2 + 1}; \quad S_{\bar{x}}(\omega) = N_{\bar{x}}.$$

(9.38)

Розв'язок. Інерційністю підсилювача і частотного дискримінатора системи знехтуємо. Тоді передаточна функція вихідної системи автопідстроювання частоти визначається виразом:

$$W_{pu}(s) = \frac{K_{pu}}{T_{\tilde{a}}s + 1}, \quad (9.39)$$

де $T_{\tilde{a}}$ - постійна часу гетеродину; $k_{\tilde{a}\tilde{i}} = k_{\tilde{a}}k_o k_{\tilde{a}}$.

Спектральна щільність суміші сигналу з завадою:

$$S_f(\omega) = S_x(\omega) + S_{\tilde{i}}(\omega) = B \frac{1 + j\omega T_0}{1 + j\omega T_x} \cdot \frac{1 - j\omega T_0}{1 - j\omega T_x}, \quad (9.40)$$

$$\text{де } B = N_x + N_{\tilde{i}}; \quad T_0^2 = T_x^2 \frac{N_x}{N_x + N_{\tilde{i}}}.$$

Таким чином:

$$\psi(j\omega) = \sqrt{B} \frac{1 + j\omega T_0}{1 + j\omega T_x}; \quad \psi(-j\omega) = \sqrt{B} \frac{1 - j\omega T_0}{1 - j\omega T_x}. \quad (9.41)$$

У відповідності з виразом (9.32):

$$\begin{aligned} W_0(j\omega) &= \frac{1}{\psi(j\omega)} \left[\frac{\alpha_1}{1 + j\omega T_x} + \frac{\alpha_2}{1 - j\omega T_0} \right]^+ \frac{N_x}{\sqrt{B}} = \\ &= \frac{1}{\psi(j\omega)} \cdot \frac{\alpha_1}{1 + j\omega T_x} \cdot \frac{N_x}{\sqrt{B}}. \end{aligned} \quad (9.42)$$

За методом невизначених коефіцієнтів:

$$\alpha_1 = T_x / (T_x + T_0), \quad (9.43)$$

У відповідності з (9.42) і (9.43) оптимальна частотна характеристика системи автопідстроювання частоти, що проектується:

$$W_0(j\omega) = \frac{k_0}{Tj\omega + 1}, \quad (9.44)$$

$$\text{де } k_0 = \alpha_1 \frac{N_x}{B}.$$

Для очевидності виразимо параметри оптимальної частотної характеристики через коефіцієнт, рівний відношенню рівнів спектральної щільності сигналу й завади. В результаті отримаємо:

$$k_0 = 1 - T / \sqrt{1 + \rho}, \quad T_0 = T / \sqrt{1 + \rho}, \quad (9.45)$$

$$\text{де } \rho = N_x / N_{\text{і}}.$$

На рис.9.4,6 показані залежності коефіцієнта підсилення K_0 й постійної змінної T_0 від коефіцієнта ρ , із яких видно, що зі зростанням ρ постійна часу оптимальної системи зменшується (її граничне значення рівне нулю), а коефіцієнт підсилення зростає (його граничне значення рівне одиниці). Таким чином, за відсутності завади оптимальна система є безінерційною ланкою (дисперсія помилки рівна нулю).

Дисперсію сумарної помилки в оптимальній системі автопідстроювання частоти визначимо за формулою

$$\begin{aligned} \sigma_e^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[|1 - W_0(j\omega)|^2 S_x(\omega) + |W_0(j\omega)|^2 S_{\text{і}}(\omega) \right] d\omega = \\ &= \frac{T_0 - (1 - k_0)^2 T_x}{2T_x(T_0 + T_x)} N_x + \frac{k_0^2}{2T_0} N_{\text{і}}. \end{aligned}$$

Приймаючи до уваги введені значення ρ :

$$\sigma_e^2 = \frac{N_x}{2T_x \sqrt{1 + \rho}} \cdot \left[1 + \frac{\rho}{(\sqrt{1 + \rho} + 1)^2} \right]. \quad (9.46)$$

Обчислимо дисперсію помилки за формулою (9.31):

$$\sigma_{e\min}^2 \frac{N_x}{2T_x \sqrt{1+\rho}}.$$

Отримане значення менше від дисперсії помилки, розрахованої за оптимальною частотною характеристикою. Це пояснюється тим, що формула (9.31) знайдена без врахування фізичної реалізованості системи.

Передаюча функція розімкненої оптимальної системи автопідстроювання частоти:

$$W_{p0}(s) = \frac{K_{p0}}{Ts + 1},$$

$$\text{де } k_{p0} = k_0 / (1 - k_0); \quad T = T_0 / (1 - k_0).$$

Передаюча функція послідовності коректуючого пристрою у відповідності до виразу (9.37):

$$W_k(s) = k_k \frac{T_{\tilde{a}}s + 1}{Ts + 1},$$

де $k_{\tilde{f}} = k_{p0} / k_{pu}$ - коефіцієнт передачі коректуючого пристрою.

Коректуючий пристрій в оптимальній системі автопідстроювання частоти підключають між підсилювачем і дискримінатором, він може бути реалізоване *RC* - ланкою з відставанням по фазі.

Як висновок відмітимо, що теорія лінійної фільтрації Колмогорова-Вінера для стаціонарних процесів, викладена вище, володіє одним істотним недоліком, який полягає в тому, що практична реалізація таких оптимальних фільтрів супроводжується великими труднощами. А якщо врахувати, що реальні модульовані сигнали не є стаціонарними, то задача побудови оптимального фільтра на базі лінійної теорії стає практично нереалізовуваною. Недоліком цієї теорії є те, що лінійний фільтр оптимальний для сигналу $s(u, t)$, а не для самого повідомлення $u(t)$. Тому при проектуванні нестационарних оптимальних систем фільтр Колмогорова-Вінера не застосовується. Рішення нестационарних задач оптимальних систем зв'язку базується на оптимальних фільтрах Калмана.

9.5. Синтез оптимальних фільтрів Калмана методом простору станів

Ефективним шляхом синтезу оптимальних нестационарних фільтрів є запропонована Калманом і Б'юсі методика, яка використовує метод простору станів. Відповідно до методики, процес, який виділяється (повідомлення), представляється у

вигляді компоненти чи лінійної комбінації компонент векторного процесу $\mathbf{x}(t)$.

Складові процесу $\mathbf{x}(t)$ визначаються рішенням системи лінійних диференціальних рівнянь, збурюваннями в яких являються білі шуми. Описаний векторний процес є частинним випадком багатовимірною марківського процесу. Тому підхід Калмана-Б'юсі до синтезу оптимальних лінійних фільтрів може бути названий марківською теорією оптимальної лінійної фільтрації. Будемо, однак, для зручності, а також зважаючи на розповсюджену термінологію, називати його синтезом методом простору станів.

Зазначений підхід до синтезу оптимальних лінійних фільтрів тісно поєднується з загальною теорією оптимальної нелінійної фільтрації марківських процесів, розробленою Р.Л. Стратоновичем [62, 64]. Оскільки задача оптимальної лінійної фільтрації є більш простою в порівнянні з нелінійною фільтрацією, відповідний до неї математичний апарат розроблений на даний час більш докладно.

Опис повідомлень. Розгляд синтезу оптимальних лінійних фільтрів методом простору станів почнемо з пояснення можливості представлення повідомлення у вигляді компоненти або лінійної комбінації компонентів марківського процесу.

Нехай, наприклад, процес $\lambda(t)$, який підлягає виділенню за допомогою оптимального фільтра, є рішенням лінійного стохастичного диференціального рівняння n -го порядку:

$$A(p)\lambda(t) = k_1 B(p)n(t), \quad (9.47)$$

де

$$A(p) = p^n + a_{n-1}p^{n-1} + \dots + a_0;$$

$$B(p) = b_m p^m + b_{m-1} p^{m-1} + \dots + b_1 p + 1; n(t) - \text{білий шум}; m \leq n - 1. \text{ Процес}$$

$S(t)$ відповідно до (9.47) можна розглядати як результат впливу білого шуму $n(t)$ на формуючий фільтр з операторним коефіцієнтом передачі:

$$K_f(p) = k_1 B(p) / A(p). \quad (9.48)$$

Представлення повідомлення $S(t)$ у вигляді рішення рівняння (9.47) є достатньо універсальним. Так, якщо при синтезі оптимального лінійного фільтра задана

спектральна щільність $S_\lambda(\omega)$ повідомлення $\lambda(t)$, яка описується дробово-раціональною функцією, то процес $S(t)$ можна розглядати як ефект на виході формуючого фільтра (9.48) в установленому режимі й відобразити диференціальним рівнянням виду (9.47). Для цього достатньо прирівняти:

$$S_\lambda(\omega) = |K_f(j\omega)|^2 = |k_1 B(j\omega) / A(j\omega)|^2,$$

визначити з цієї умови комплексний коефіцієнт передачі $K_f(j\omega)$ і відновити за $K_f(j\omega)$ диференціальне рівняння формуючого фільтра.

Заміною змінних рівняння (9.47) можна представити у вигляді системи лінійних диференціальних рівнянь першого порядку, у праві частини яких входить білий шум, але не його похідні. Якщо $B(p) = 1$, то вказаний перехід виконується за допомогою заміни змінних:

$$\left. \begin{aligned} dx_1(t)/dt &= x_2(t), \\ dx_2(t)/dt &= x_3(t) \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ dx_{n-1}/dt &= x_n(t) \end{aligned} \right\}, \quad (9.49)$$

де $x_1(t) = \lambda(t)$.

Рівняння (9.47) в нових змінних записується у вигляді:

$$\frac{dx_n}{dt} = -a_{n-1}x_n - \dots - a_0x_1 + k_1n(t). \quad (9.50)$$

Сукупність компонент $x_1(t), x_2(t), \dots, x_n(t)$, задовольняючих рівнянням (9.49), (9.50), утворить вектор-функцію, яка описує багатовимірний марківський процес $\mathbf{x}(t)$ в n -вимірному просторі (просторі станів). Систему рівнянь (9.49), можна записати у векторній формі:

$$d\mathbf{x}/dt = \mathbf{F}\mathbf{x}(t) + \mathbf{H}n(t), \quad (9.51)$$

де $\mathbf{x}(t)$ - n -вимірний вектор-стовпець із компонентами $x_1(t), x_2(t), \dots, x_n(t)$, $\mathbf{F}$ і $\mathbf{H}$ - відповідно матриця і вектор-стовпець вигляду:

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & 1 \\ -a_0 & \dots & -a_1 & \dots & -a_{n-1} \end{bmatrix},$$

$$\mathbf{H} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ k_1 \end{bmatrix}. \quad (9.52)$$

Повідомлення $\lambda(t) = x_1(t)$ при такому підході являється першою компонентою векторного процесу $\mathbf{x}(t)$ і зв'язано з ним співвідношенням:

$$\lambda(t) = x_1(t) = \mathbf{c}_\lambda \mathbf{x}(t), \quad (9.53)$$

де $\mathbf{c}_\lambda$ - n -вимірний вектор-рядок виду $\mathbf{c}_\lambda = [1 \quad 0 \quad 0 \quad \dots \quad 0]$.

Якщо в рівнянні (9.47) коефіцієнти $b_1, b_2, \dots, b_m$ відрізняються від нуля, то представлення його у вигляді системи рівнянь першого порядку з білими шумами в

правих частинах ускладнюється. Дійсно, в цьому випадку при заміні змінних (9.49) в рівняннях з'являються похідні білого шуму. Існує декілька шляхів, які дозволяють позбутися від похідних білого шуму на випадок, коли коефіцієнти $b_1, b_2, \dots, b_m$ не дорівнюють нулю. Один із них полягає в наступному. Вводяться змінні:

$$\left. \begin{aligned} \lambda(t) &= x_1(t), \\ dx_1 / dt &= x_2(t) + c_1 x(t) \\ &\dots\dots\dots \\ dx_{n-1} / dt &= x_n(t) + c_{n-1} x(t) \end{aligned} \right\}, \quad (9.54)$$

де $c_1, c_2, \dots, c_{n-1}$ - постійні коефіцієнти, які обираються надалі так, щоб усунути похідні білого шуму. Із (9.54) випливає, що:

$$\frac{d^k x}{dt^k} = x_{k+1} + \sum_{i=0}^{k-1} c_{k-i} \frac{d^i \chi(t)}{dt^i} \text{ при } 1 \leq k \leq n-1, \quad (9.55)$$

$$\frac{d^n x}{dt^n} = \frac{dx_n}{dt} + \sum_{i=1}^{n-1} c_{n-i} \frac{d^i \chi(t)}{dt^i}.$$

Із врахуванням (9.54) і (9.55) рівняння (9.47) записується у вигляді:

$$\frac{dx_n}{dt} + \sum_{i=1}^{n-1} c_{n-i} \frac{d^i \chi(t)}{dt^i} + \sum_{k=0}^{n-1} a_k x_{k+1} + \sum_{i=0}^{n-2} \frac{d^i \chi(t)}{dt^i} \sum_{k=1}^{n-1-i} a_{k+i} c_k = \sum_{i=0}^m b_i \frac{d^i \chi(t)}{dt^i}. \quad (9.56)$$

Похідні білого шуму в рівнянні (9.56) усуваються при виконанні наступних рівностей:

$$c_{n-i} + \sum_{k=1}^{n-1-i} a_{k+i} c_k = b_i, \quad i = 1, 2, \dots, n-1. \quad (9.57)$$

Послідовно розв'язуючи рівняння (9.57) починаючи з $i = n-1$, визначимо шукані коефіцієнти $c_1, c_2, \dots, c_{n-1}$.

При виконанні рівностей (9.57) система рівнянь (9.54), (9.56) прийме вигляд:

$$\left. \begin{aligned} dx_1 / dt &= x_2(t) + c_1 n(t), \\ &\dots\dots\dots, \\ dx_n / dt &= -\sum_{k=0}^{n-1} a_k x_{k+1} + \left(k_1 - \sum_{k=1}^{n-1} a_k c_k \right) n(t), \end{aligned} \right\} \quad (9.58)$$

де $x_1(t) = \lambda(t)$. Систему рівнянь (9.58) також можна записати у вигляді векторного рівняння (9.51), в якому матриця **F** визначається виразом (9.52), а векторний стовпець **H** дорівнює:

$$\mathbf{H} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ k_1 - \sum_{k=1}^{n-1} a_k c_k \end{bmatrix}.$$

Повідомлення $\lambda(t)$ в даному випадку також визначається через векторний процес $\mathbf{x}(t)$ рівністю (9.53).

Існує другий спосіб складання системи стохастичних диференціальних рівнянь першого порядку з діями у вигляді білого шуму. В цьому випадку вводиться нова змінна $x_1(t)$, пов'язана з $\lambda(t)$ співвідношенням:

$$\lambda(t) = B(p)x_1(t) \dots \quad (9.59)$$

Після підстановки (9.59) в рівняння (9.47) воно записується у вигляді:

$$A(p)x_1(t) = k_1 n(t). \quad (9.60)$$

Рівнянню (9.60) відповідає система диференціальних рівнянь першого порядку:

$$\left. \begin{aligned} dx_1 / dt &= x_2(t), \\ &\dots\dots\dots \\ dx_{n-1} / dt &= x_n(t), \\ dx_n / dt &= -\sum_{k=0}^{n-1} a_k x_{k+1} + k_1 n(t). \end{aligned} \right\} \quad (9.61)$$

Ця система може бути записана у вигляді векторного рівняння:

$$dx/dt = Fx(t) + H(t)n(t),$$

де $\mathbf{x}(t)$ - n -вимірний векторний процес; матриця $\mathbf{F}$ і вектор-стовпець $\mathbf{H}$ визначаються виразами (9.10). На відміну від попереднього випадку, процес $\lambda(t)$ при даному способі заміни змінних виражається, як впливає з (9.59), (9.61), через декілька компонент векторного процесу $\mathbf{x}(t)$. Так, наприклад, якщо $B(p) = 1 + T_1 p$, то

$$\lambda(t) = x_1(t) + T_1 x_2(t). \quad (9.62)$$

Зв'язок між повідомленнями $\lambda(t)$ і n -вимірним вектором $\mathbf{x}(t)$ встановлюється виразом:

$$\lambda(t) = \mathbf{c}_\lambda \mathbf{x}(t),$$

де $\mathbf{c}_\lambda = [1 \quad T_1 \quad 0 \quad \dots \quad 0]$ - вектор-рядок, який містить n компонент.

Другий спосіб складання системи стохастичних диференціальних рівнянь дозволяє отримати в векторному рівнянні більш простий порівняно з першим способом

вираз для вектора-стовпця $\mathbf{H}$. Однак це досягається ціною ускладнення вектора-рядка $\mathbf{C}_\lambda$.

Проведений розгляд показує, як можна представити повідомлення у вигляді окремої компоненти чи лінійної комбінації декількох компонент векторного марківського процесу, якщо це повідомлення описується лінійним диференціальним рівнянням n -го порядку достатньо загального вигляду (9.47) або задано дробово-раціональною спектральною щільністю.

На практиці виникають також задачі, в яких процес (повідомлення), що виділяється, описується диференціальним рівнянням чи системою диференціальних рівнянь зі змінними коефіцієнтами. При синтезі згладжуючих ланок складних пристроїв зв'язку на входи оптимізуемого фільтра надходять в суміші з шумом декілька компонент процесу, який фільтрується, або фільтруємих процес в суміші з різними завадами. Щоб охопити ці, а також ряд інших можливих випадків, загальна постановка задачі синтезу оптимального лінійного фільтра методом простору станів формується наступним чином.

Постановка задачі. Передбачається, що векторний процес $\mathbf{x}(t)$, який описує повідомлення $\lambda(t)$ в n -вимірному просторі станів, задовольняє векторному диференціальному рівнянню:

$$d\mathbf{x}/dt = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{H}(t)\mathbf{n}(t) \quad (9.63)$$

де $\mathbf{x}(t)$ - вектор-функція, що містить n компонент; $\mathbf{F}$ - квадратна матриця розміром $n \times n$, яка залежить у загальному випадку від часу; $\mathbf{n}(t)$ - вектор білих шумів, який складається з l компонент; $\mathbf{H}(t)$ - матриця розміром $n \times l$.

Кореляційна матриця процесу $\chi(t)$ має вигляд:

$$\mathbf{R}_x(t, u) = M \{ \mathbf{n}(t) \mathbf{n}^T(u) \} = \mathbf{Q}(t) \delta(t - u), \quad (9.64)$$

де $\mathbf{Q}(t)$ - квадратна матриця розміром $l \times l$. Як видно із (9.64), шум $\chi(t)$ в загальному випадку передбачається нестационарним.

Нагадаємо, що в найпростішому випадку повідомлення $\lambda(t)$ є однією з компонент векторного процесу $\mathbf{x}(t)$. Однак, як видно, наприклад, з (9.62) можливі й більш складні ситуації, коли повідомлення $\lambda(t)$ є комбінацією декількох компонентів процесу $\mathbf{x}(t)$. У задачах одночасної фільтрації декількох повідомлень компоненти вектора $\mathbf{x}(t)$ можуть бути зв'язані з різними повідомленнями. Обмежимося надалі обговоренням фільтрації одного повідомлення, який виражається через процес $\mathbf{x}(t)$ співвідношенням:

$$\lambda(t) = \mathbf{c}_\lambda \mathbf{x}(t),$$

де $\mathbf{c}_\lambda$ вектор-рядок, що містить n елементів.

Фільтр, який оптимізується, в загальному випадку має декілька входів, на які можуть надходити в суміші з шумом повідомлення $\lambda(t)$ та його похідні. Сукупність процесів на m входах фільтра утворює векторний процес, який ми спостерігаємо:

$$\varphi(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{n}(t), \quad (9.65)$$

де $\varphi(t)$ - вектор-функція, яка складається із m елементів; $\mathbf{C}(t)$ - прямокутна матриця розміром $m \times n$; $\mathbf{n}(t)$ - m -вимірний вектор білих шумів (шумів вимірювання) з кореляційною матрицею:

$$\mathbf{R}(t, u) = \mathbf{R}(t)\delta(t - u), \quad (9.66)$$

тут $\mathbf{R}(t)$ - позитивно-визначена квадратна матриця розміром $m \times m$. В багатьох випадках процес $\varphi(t)$ скалярний і являє собою адитивну суміш повідомлення $\lambda(t)$ і шуму. В цьому випадку $\mathbf{C}(t)$ - матриця розміром $1 \times n$, тобто вектор-рядок, що співпадає з $\mathbf{c}_\lambda$.

Процес $\varphi(t)$, що спостерігається, надходить на вхід фільтра, починаючи з моменту $t = t_0$. Значення кореляційної матриці процесу $\mathbf{x}(t)$ в момент $t = t_0$ вважається відомим. Процес $\mathbf{x}(t)$ і перешкода $\mathbf{n}(t)$ вважаються незалежними. Фільтр повинний забезпечити виділення повідомлення $\lambda(t)$ з мінімальною середньоквадратичною помилкою в будь-який момент часу $t > t_0$.

Рівняння оптимального фільтра. Рівняння, які визначають структуру й характеристики оптимального фільтра, при розглянутій постановці задачі можуть бути отримані різними способами. Оскільки висновок їх виявляється досить громіздким, обмежимося тут обговоренням кінцевих результатів.

Структура оптимального фільтра описується векторним диференціальним рівнянням:

$$d\mathbf{x}_0 / dt = \mathbf{G}_0(t, t)[\varphi(t) - \mathbf{C}(t)\mathbf{x}_0(t)] + \mathbf{F}(t)\mathbf{x}_0(t), \quad (9.67)$$

де $\mathbf{x}_0(t)$ - векторний процес на виході оптимального фільтра, який відтворює з деякою помилкою процес $\mathbf{x}(t)$; $\mathbf{G}_0(t, t)$ - матриця змінних коефіцієнтів, яка утворюється з матриці $\mathbf{G}_0(t, u)$ при $t = u$, $\mathbf{G}_0(t, u)$ - матриця імпульсних перехідних функцій оптимального фільтра, названа також більш коротко перехідною матрицею.

Початкове значення вектора $\mathbf{x}_0(t_0)$ в рівнянні (9.67) зазвичай вважається нульовим.

Поняття перехідної матриці лінійного фільтра вже зустрічалося в розділі 7. Нагадаємо, що у випадку, коли дія й реакція на виході фільтра описуються скалярними функціями, перехідна матриця містить єдиний компонент, що є звичайною імпульсною перехідною функцією фільтра. Якщо фільтр є багатовимірним, то компонента перехідної матриці, що знаходиться на перетині i -го рядка та j -го стовпця, являється імпульсною перехідною функцією, яка зв'язує зміну i -ї компоненти вихідного процесу при дії у вигляді одиничного дельта-імпульсу за його j -ю компонентою. Останнє еквівалентне одиничному початковому значенню i -ї компоненти процесу на виході фільтра при нульових початкових значеннях решти компонент.

Матриця $\mathbf{G}_0(t, t)$, що входить в (9.67), визначається виразом:

$$\mathbf{G}_0(t, t) = \mathbf{E}(t)\mathbf{C}^T(t)\mathbf{R}^{-1}(t), \quad (9.68)$$

де $\mathbf{R}^{-1}(t)$ - матриця, зворотна до матриці $\mathbf{R}(t)$; $\mathbf{E}(t)$ - матриця дисперсій помилок фільтрації:

$$\mathbf{E}(t) = M \left\{ [\mathbf{x}(t) - \mathbf{x}_0(t)] [\mathbf{x}(t) - \mathbf{x}_0(t)]^T \right\} \quad (9.69)$$

Матриця $\mathbf{E}(t)$ - квадратна симетрична розміром $n \times n$. Її елементами є дисперсії та взаємні дисперсії помилок фільтрації окремих складових вектора $\mathbf{x}(t)$. Іноді $\mathbf{E}(t)$ називають матрицею коваріацій.

Зміна помилок фільтрації в часі описується наступним рівнянням :

$$d\mathbf{E}/dt = \mathbf{F}(t)\mathbf{E}(t) + \mathbf{E}(t)\mathbf{F}^T(t) - \mathbf{E}(t)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)\mathbf{C}(t)\mathbf{E}(t) + \mathbf{H}(t)\mathbf{Q}\mathbf{H}^T(t). \quad (9.70)$$

Рівняння (9.70) є матричним нелінійним диференціальним рівнянням Ріккати. Для його розв'язку необхідно задати початкове значення $\mathbf{E}(t_0)$ матриці дисперсій помилок.

Якщо в момент $t = t_0$ процес $\mathbf{x}_0(t_0)$ на виході фільтра дорівнює нулю, то, як видно

із (9.69), матриця $\mathbf{E}(t_0)$ дорівнює матриці дисперсій компонент фільтрованого процесу $\mathbf{x}(t)$ в момент $t = t_0$, тобто $\mathbf{E}(t_0) = M\{\mathbf{x}(t_0)\mathbf{x}^T(t_0)\}$.

Сукупність виразів (9.67), (9.68), (9.70) повністю визначає структуру й параметри оптимального фільтра. Структурна схема такого фільтра, побудована за рівнянням (9.67), зображена на рис. 9.5. Подвійними лініями на цьому рисунку позначені зв'язки між векторними процесами. При рішенні конкретних задач рівняння (9.67) доцільно представити у вигляді системи диференціальних рівнянь першого порядку. За цією системою рівнянь легко будується, як показано нижче на прикладах, структура оптимального фільтра.

Найбільш складним етапом синтезу оптимальних фільтрів, що розглядається методом, є розв'язок рівняння (9.70). Як правило він вимагає застосування ЦОМ.

Одним із можливих шляхів розв'язку рівняння (9.70) є заміна нелінійного диференціального рівняння з матрицею розміром $n \times n$ системою лінійних диференціальних рівнянь, еквівалентних лінійному матричному диференціальному рівнянню з матрицею розміром $2n \times 2n$. Вказаний перехід виконується шляхом представлення $\mathbf{E}(t)$ у вигляді:

$$\dot{\mathbf{v}}_1(t) = \mathbf{E}(t)\mathbf{v}_2(t), \quad (9.71)$$

де $\mathbf{v}_1(t)$ і $\mathbf{v}_2(t)$ - n -вимірні вектори, що задовольняють рівнянням:

$$d\mathbf{v}_1/dt = \mathbf{F}\mathbf{v}_1 + \mathbf{H}\mathbf{Q}\mathbf{H}^T\mathbf{v}_2, \quad (9.72)$$

$$d\mathbf{v}_2/dt = \mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\mathbf{v}_1 - \mathbf{F}^T\mathbf{v}_2. \quad (9.73)$$

Щоб впевнитися в справедливості виконуваного переходу, представимо (9.72). враховуючи (9.71), у вигляді:

$$\dot{\mathbf{E}}(t)\mathbf{v}_2(t) + \mathbf{E}(t)\dot{\mathbf{v}}_2(t) = \mathbf{F}\mathbf{E}(t)\mathbf{v}_2(t) + \mathbf{H}\mathbf{Q}\mathbf{H}^T\mathbf{v}_2(t). \quad (9.74)$$

Замінюючи $\mathbf{v}_2(t)$ в (9.74) у відповідності з (9.73), отримуємо:

$$\left[\dot{\mathbf{E}}(t) - \mathbf{F}\mathbf{E}(t) - \mathbf{E}(t)\mathbf{F}^T + \mathbf{E}(t)\mathbf{C}^T \mathbf{R}^{-1} \mathbf{C}\mathbf{E}(t) - \mathbf{H}\mathbf{Q}\mathbf{H}^T \right] \mathbf{v}_2(t) = 0. \quad (9.75)$$

При виконанні рівняння Ріккати рівняння (9.75) задовольняється тотожно, що й доводить правомірність виконуваного переходу.

Виразимо матрицю дисперсій $\mathbf{E}(t)$ через перехідну матрицю $\mathbf{T}(t, t_0)$ системи рівнянь (9.72), (9.73). Матриця $\mathbf{T}(t, t_0)$ встановлює наступний зв'язок:

$$\begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{bmatrix} = \mathbf{T}(t, t_0) \begin{bmatrix} \mathbf{v}_1(t_0) \\ \mathbf{v}_2(t_0) \end{bmatrix}. \quad (9.76)$$

Представляючи матрицю $\mathbf{T}(t, t_0)$ розміром $2n \times 2n$ у вигляді блочної матриці, елементи якої являються матрицями розміром $n \times n$:

$$\mathbf{T}(t, t_0) = \begin{bmatrix} \mathbf{T}_{11}(t, t_0), & \mathbf{T}_{12}(t, t_0) \\ \mathbf{T}_{21}(t, t_0), & \mathbf{T}_{22}(t, t_0) \end{bmatrix},$$

Запишемо на основі (9.76):

$$\begin{aligned} \mathbf{v}_1(t) &= \mathbf{T}_{11}(t, t_0)\mathbf{v}_1(t_0) + \mathbf{T}_{12}(t, t_0)\mathbf{v}_2(t_0), \\ \mathbf{v}_2(t) &= \mathbf{T}_{21}(t, t_0)\mathbf{v}_1(t_0) + \mathbf{T}_{22}(t, t_0)\mathbf{v}_2(t_0), \end{aligned}$$

Звідси випливає, що

$$\mathbf{E}(t)[\mathbf{T}_{21}(t, t_0)\mathbf{v}_1(t_0) + \mathbf{T}_{22}(t, t_0)\mathbf{v}_2(t_0)] = \mathbf{T}_{11}(t, t_0)\mathbf{v}_1(t_0) + \mathbf{T}_{12}(t, t_0)\mathbf{v}_2(t_0). \quad (9.77)$$

При множенні обох частин рівності (9.77) справа на $\mathbf{v}_2^{-1}(t_0)$, враховуючи, що $\mathbf{v}_1(t_0)\mathbf{v}_2^{-1}(t_0) = \mathbf{E}(t_0)$, після перетворення отримаємо:

$$\mathbf{E}(t)[\mathbf{T}_{21}(t, t_0)\mathbf{E}(t_0) + \mathbf{T}_{22}(t, t_0)] = \mathbf{T}_{11}(t, t_0)\mathbf{E}(t_0) + \mathbf{T}_{12}(t, t_0). \quad (9.78)$$

Як видно з (9.78), для визначення матриці $\mathbf{E}(t)$ необхідно знати початкове значення $\mathbf{E}(t_0)$ і перехідну матрицю $\mathbf{T}(t, t_0)$. Щоб знайти матрицю $\mathbf{T}(t, t_0)$, запишемо систему рівнянь (9.72), (9.73) у вигляді:

$$d\mathbf{v}/dt = \mathbf{A}\mathbf{v}(t), \quad (9.79)$$

де $\mathbf{v}(t)$ - складовий вектор, компонентами якого є вектори $\mathbf{v}_1(t)$ і $\mathbf{v}_2(t)$; $\mathbf{A}$ - блочна матриця виду:

$$\mathbf{A} = \begin{bmatrix} \mathbf{F}(t) & \mathbf{H}(t)\mathbf{Q}(t)\mathbf{H}^T(t) \\ \mathbf{C}^T(t)\mathbf{R}^{-1}\mathbf{C}(t) & -\mathbf{F}^T(t) \end{bmatrix}. \quad (9.80)$$

Існують різні методи обчислення перехідної матриці для системи, описуваної рівнянням (9.79). Якщо розмір $\mathbf{A}$ невеликий, і вхідні в неї матриці $\mathbf{F}, \mathbf{C}, \mathbf{H}, \mathbf{Q}$ постійні в часі, то перехідну матрицю $\mathbf{T}(t, t_0)$ можна, наприклад, знайти за формулою:

$$\mathbf{T}(t, t_0) = \mathbf{T}(t - t_0) = \mathbf{T}(\tau) = L^{-1} \{ [\mathbf{sI} - \mathbf{A}]^{-1} \}, \quad (9.81)$$

де L^{-1} - операція зворотного перетворення Лапласа; $\mathbf{I}$ - одинична матриця; $[\mathbf{sI} - \mathbf{A}]^{-1}$ - матриця, зворотна до $\mathbf{sI} - \mathbf{A}$; $\mathcal{S}$ оператор перетворення Лапласа.

В тих випадках, коли необхідно знайти значення коефіцієнтів $\mathbf{G}_0(t, t_0)$ в структурній схемі оптимального фільтра тільки в установленому режимі, достатньо знайти встановлене значення матриці $\mathbf{E}(t)$. Якщо крім того, корисне повідомлення й завади являються стаціонарними процесами, то в рівнянні (9.70) слід покласти $d\mathbf{E}/dt = 0$ і розв'язати отримане матричне алгебраїчне рівняння. Остання задача значно простіша, ніж розв'язок повного рівняння (9.70).

9.6. Приклади синтезу оптимальних фільтрів

Розглянемо декілька прикладів синтезу лінійних фільтрів методом простору станів.

Приклад 9.3. Нехай повідомлення $\lambda(t)$ - стаціонарний випадковий процес, який описується диференціальним рівнянням першого порядку:

$$\frac{d\lambda}{dt} = -\frac{1}{T}\lambda(t) + \frac{1}{T}\chi(t).$$

Такий процес відповідає білому шуму $n(t)$, пропущеному через інерційну ланку з операторним коефіцієнтом передачі $K(p) = 1/(1 + pT)$. Функція кореляції шуму $n(t)$ записується у вигляді:

$$R_x(t, u) = S_c(0)\delta(t - u), \quad (9.82)$$

де $S_c(0)$ - спектральна щільність шуму $n(t)$ на нульовій частоті. Відмітимо, що спектральна щільність повідомлення $\lambda(t)$ при цьому рівна:

$$S_\lambda(\omega) = \frac{2\mu\sigma_\lambda^2}{\omega^2 + \mu^2}, \quad (9.83)$$

де $\mu = 1/T$, $\sigma_\lambda^2 = S_c(0)/2T$ - дисперсія процесу $\lambda(t)$. Так як $\lambda(t)$ описується рівнянням першого порядку, то векторний процес $\mathbf{x}(t)$ в даній задачі одномірний і складається з єдиної компоненти, яка співпадає з $\lambda(t)$. Відповідно, процес $\mathbf{x}(t)$ описується рівнянням:

$$\frac{d\lambda}{dt} = -\frac{1}{T}x(t) + \frac{1}{T}\chi(t). \quad (9.84)$$

Припустимо, що суміш повідомлення й завади на вході фільтра має вигляд:

$$\varphi(t) = x(t) + n(t), \quad (9.85)$$

де $n(t)$ - стаціонарний білий шум із спектральною щільністю $S_n(0)$ та функцією кореляції $R(\tau) = S_n(0)\delta(\tau)$.

Співставлення (9.82), (9.84), (9.85) із (9.64) і (9.65) показує, що в даній задачі матриці **F**, **H**, **C**, **Q**, **R** є постійними скалярами і дорівнюють:

$$\mathbf{F} = -1/T, \quad \mathbf{C} = 1, \quad \mathbf{H} = 1/T, \quad \mathbf{Q} = S_c(0), \quad \mathbf{R} = S_n(0). \quad (9.86)$$

Рівняння (9.67), що визначає структуру оптимального фільтра, конкретизується в даному випадку у вигляді:

$$dx_0/dt = \mathbf{G}_0(t,t)[\varphi(t) - x_0(t)] - x_0(t)/T, \quad (9.87)$$

де згідно (9.68) і (9.86)

$$\mathbf{G}_0(t,t) = \mathbf{E}(t)/S_n(0). \quad (9.88)$$

Рівнянню (8.87) відповідає структурна схема, показана на рис.9.6,а інтегратор, охоплений зворотнім зв'язком, можна замінити інерційною ланкою. Тоді структурна схема прийме вигляд, зображений на рис.9.6,б.

Рівняння (9.70) для дисперсій похибок в даній задачі скалярне, і має вигляд:

$$\frac{dE}{dt} = -\frac{2}{T}E(t) - \frac{1}{S_n(0)}\mathbf{E}^2(t) + \frac{1}{T^2}S_c(0).$$

Замінімо його у відповідності з (9.72), (9.73) системою лінійних диференціальних рівнянь:

$$\left. \begin{aligned} \frac{dv_1}{dt} &= -\frac{1}{T} \mathbf{v}_1 + \frac{S_c \mathbf{v}_2}{T^2}, \\ \frac{dv^2}{dt} &= \frac{\mathbf{v}_1}{S_n(0)} + \frac{\mathbf{v}_2}{T}. \end{aligned} \right\} \quad (9.89)$$

Із (9.89) випливає, що матриця $\mathbf{A}$ має вигляд:

$$\mathbf{A} = \begin{bmatrix} -1/T & S_c(0)/T^2 \\ 1/S_n(0) & 1/T \end{bmatrix}. \quad (9.90)$$

Припустимо, що $t_0 = 0$, і використаємо для обчислення перехідної матриці $\mathbf{T}(t,0)$ формулу (9.82). Матриця $[s\mathbf{I} - \mathbf{A}]$ у відповідності з (9.90) рівна:

$$\begin{bmatrix} s + 1/T & -S_c(0)/T^2 \\ -1/S_n(0) & s - 1/T \end{bmatrix}.$$

Зворотна до неї матриця, обчислена у відповідності з Д.2.10, записується у вигляді:

$$[s\mathbf{I} - \mathbf{A}]^{-1} = \frac{1}{s^2 - \gamma^2} \begin{bmatrix} s - 1/T & S_c(0)/T^2 \\ 1/S_n(0) & s + 1/T \end{bmatrix}, \quad (9.91)$$

де $\gamma = \sqrt{1 + S_c(0)/S_n(0)} = \sqrt{1 + \rho/T}$; $\rho = S_c(0)/S_n(0)$.

Відшуковуючи зворотне перетворення Лапласа для елементів матриці (9.91) отримуємо:

$$\mathbf{T}(t,0) = \begin{bmatrix} ch\gamma t - \frac{1}{\gamma T} sh\gamma t & \frac{S_c(0)}{T^2 \gamma} sh\gamma t \\ \frac{sh\gamma t}{S_n(0) \gamma} & ch\gamma t + \frac{sh\gamma t}{\gamma T} \end{bmatrix} \quad (9.92)$$

і, отже:

$$\begin{aligned} \mathbf{T}_{11} &= ch\gamma t - \frac{sh\gamma t}{\gamma T}, & \mathbf{T}_{12} &= \frac{S_c(0)sh\gamma t}{T^2\gamma}, \\ \mathbf{T}_{21} &= \frac{sh\gamma t}{S_n(0)\gamma}, & \mathbf{T}_{22} &= ch\gamma t + \frac{sh\gamma t}{\gamma T}. \end{aligned} \quad (9.93)$$

За формулою (9.78) із врахуванням (9.93) знаходимо, що дисперсія помилки рівна:

$$\mathbf{E}(t) = \frac{\mathbf{E}(0) \left[ch\gamma t - \frac{sh\gamma t}{\gamma T} \right] + \frac{S_c(0)sh\gamma t}{T^2\gamma}}{\mathbf{E}(0) \cdot \frac{sh\gamma t}{S_n(0)\gamma} + ch\gamma t + \frac{sh\gamma t}{\gamma T}}. \quad (9.94)$$

Значення $\mathbf{E}(0)$ дорівнює дисперсії σ_λ^2 процесу $\lambda(t)$ і визначається співвідношенням (9.84), тобто $\mathbf{E}(0) = \sigma_\lambda^2 = S_c(0)/2T$. При цьому вираз (9.94) для помилки приймає вигляд:

$$\mathbf{E}(t) = \sigma_\lambda^2 \frac{\sqrt{1 + \rho ch\gamma t + sh\gamma t}}{\sqrt{1 + \rho ch\gamma t + (1 + 0.5\rho)sh\gamma t}}. \quad (9.95)$$

Дисперсію помилки фільтрації в установленому режимі знайдемо, покладаючи в (9.95) $t \rightarrow \infty$:

$$E_{\text{оñò}} = \frac{S_c(0)/T}{1 + \gamma T} = \frac{2\sigma_\lambda^2}{1 + \sqrt{1 + \rho}},$$

Зміна в часі коефіцієнта передачі $G_0(t, t)$ в схемі оптимального фільтра (рис. 9.6) описується, як випливає із (9.88) і (9.95), виразом:

$$G_0(t, t) = \frac{1}{2T} \rho \frac{\sqrt{1 + \rho ch\gamma} + sh\gamma}{\sqrt{1 + \rho ch\gamma} + (1 + 0.5\rho)sh\gamma}. \quad (9.96)$$

На рис. 9.6, в суцільними лініями показана розрахована за формулою (9.96) для декількох значень ρ зміна нормованого коефіцієнта $G_0(t, t)T$ в часі. Як видно з рисунку, коефіцієнт $G_0(t, t)$ максимальний в перший момент часу, а потім поступово зменшується, прямує до встановленого значення $G_{0\infty}$. Така поведінка коефіцієнта $G_0(t, t)$ оптимального фільтра пояснюється наступним чином. В перший момент часу повідомлення $\lambda(t)$ не проходить на вихід фільтра, й виникає значна помилка. Щоб швидко усунути її, необхідно розширити смугу перепуску слідкуючої системи, що й досягається збільшенням коефіцієнта $G_0(t, t)$. По мірі обробки слідкуючої системи повідомлення $\lambda(t)$ вага помилки, викликаной неточним відтворенням повідомлення $\lambda(t)$, зменшується. Одночасно збільшується флуктуаційна складова помилки. Щоб обмежити її є доцільним звужити смугу перепуску системи, зменшивши коефіцієнт $G_0(t, t)$. Початкове та встановлене значення коефіцієнта $G_0(t, t)$, як випливає з (9.96), дорівнюють:

$$G_0(0) = \frac{\rho}{2T}, \quad G_{0\infty} = \frac{\rho/T}{1 + \sqrt{1 + \rho}}. \quad (9.97)$$

Їх співвідношення $G_0(0)/G_{0\infty} = (1 + \sqrt{1 + \rho})/2$, яке характеризує степінь зміни коефіцієнта $G_0(t, t)$, тим більше, чим більше відношення $\rho = S_c(0)/S_n(0)$ спектральних щільностей повідомлення й завади. Зі зростанням ρ зростає також значення коефіцієнта передачі $G_0(t, t)T$ в установленому режимі.

Якщо повідомлення $\lambda(t)$ є нестационарним процесом, то поведінка коефіцієнта $G_0(t, t)$ оптимального фільтра змінюється. Розглянемо в якості ілюстрації один частинний випадок. Нехай в умовах даної задачі повідомлення $\lambda(t)$ утворюється включенням на вхід формуючого фільтра білого шуму $\chi(t)$ в той же момент часу t_0 , коли суміш повідомлення й завади надходить для фільтрації на вхід оптимального фільтра.

Повідомлення $\lambda(t)$ представляє собою в цьому випадку процес з наростаючою дисперсією $\sigma_\lambda^2(t) = \sigma_{\lambda_0}^2 [1 - \exp(-2(t - t_0)/T)]$, величина якої прямує до усталеного значення $\sigma_{\lambda_0}^2 = S_c(0)/2T$. В момент $t = t_0$ повідомлення на виході формуючого фільтра відсутнє, й дисперсія $\sigma_\lambda^2(t_0) = 0$. Так як початкове значення $x_0(t_0)$ процесу на виході оптимального фільтра приймається рівним нулю, то при $\sigma_\lambda^2(t_0) = 0$ дисперсія помилки $E(t_0)$ в момент $t = t_0$ також рівна нулю.

Зміну коефіцієнта $G_0(t, t)$ в схемі оптимального фільтру для цього випадку можна обчислити за формулами (9.86), (9.94), покладаючи в (9.94) $E(t_0) = 0$. Відповідна крива при $\rho = 8$ показана на рис. 9.7 пунктиром. Як видно з рисунку, оптимальний коефіцієнт $G_0(t, t)$ в даному випадку збільшується по мірі зростання дисперсії повідомлення, прямує до попереднього встановленого значення.

Комплексний коефіцієнт передачі знайденого оптимального фільтра (рис. 9.6, б) в установленому режимі дорівнює:

$$\Phi(j\omega) = \frac{G_{\text{онд}} T / (1 + j\omega T)}{1 + G_{\text{онд}} T / (1 + j\omega T)} = k_{\text{ааа}} \frac{1}{1 + j\omega T_{\text{онд}}}, \quad (9.98)$$

$$\text{де } T_{\text{ааа}} = T / \sqrt{1 + \rho}, \quad k_{\text{ааа}} = \rho / (1 + \rho + \sqrt{1 + \rho}).$$

Представляє певний інтерес оцінити, який виграш досягається за рахунок змінності параметрів оптимального фільтру. Щоб оцінити його, співставимо помилки фільтрації при виділенні стаціонарного повідомлення оптимальним фільтром зі змінними параметрами й оптимальним фільтром (9.98) з постійними параметрами. Якщо початкові умови на виході фільтра нульові, то в перший момент для обох фільтрів дисперсії помилки рівні дисперсії повідомлення а, отже, однакові. В установленому режимі, як відмічалось вище, помилки фільтрації також співпадають. Виграш при використанні оптимального фільтра зі змінними проявляється в зменшенні помилки в перехідному режимі.

Виведемо співвідношення, яке описує зміну помилки фільтрації для оптимального фільтра (9.98). Дисперсія флуктуаційної складової σ_1^2 , зумовлена дією завади $n(t)$, змінюється за законом:

$$\sigma_1^2(t) = S_n(0) \int_0^t g^2(\tau) d\tau,$$

де $g(t)$ - імпульсна перехідна функція, яка відповідає комплексному коефіцієнту передачі (9.98) й рівна

$$g(t) = (k_{\text{ааа}} / \dot{O}_{\text{ааа}}) e^{-t/T_{\text{ааа}}}.$$

В результаті інтегрування отримаємо:

$$\sigma_1^2(t) = S_n(0) \frac{k_{\text{ааа}}^2}{2\dot{O}} (1 - e^{-2t/T_{\text{ааа}}}).$$

Комплексний коефіцієнт передачі, який пов'язує повідомлення $\lambda(t)$ і викликану його зміною складову помилки, рівний:

$$K(j\omega) = 1 - \Phi(j\omega).$$

Йому відповідає імпульсна перехідна функція

$$g_2(t) = \delta(t) - \frac{k_{\hat{a}\hat{e}\hat{a}}}{T_{\hat{a}\hat{e}\hat{a}}} e^{-t/T_{\hat{a}\hat{e}\hat{a}}}.$$

Так як повідомлення $\lambda(t)$ не є білим шумом і має функцію кореляції

$R_\lambda(\tau) = \sigma_\lambda^2 e^{-\mu|\tau|}$, то для обчислення дисперсії складової помилки, зумовленої неточним відтворенням повідомлення, застосуємо формулу

$$\sigma_2^2(t) = 2 \int_0^t g_2(\tau) g_1(\tau) d\tau,$$

де $g_1(\tau)$ - функція, зображення якої за Лапласом дорівнює добутку зображення $\sigma_\lambda^2/(s + \mu)$ кореляційної функції $R_\lambda(\tau)$ дії та передаточної функції $K(s) = 1 - O(s)$, тобто рівний $\sigma_\lambda^2 K(s)/(s + \mu)$.

Визначивши, що допоміжна функція $g_1(\tau) = \sigma_\lambda^2 e^{-|\tau|/T_{\hat{a}\hat{e}\hat{a}}}$ та виконавши інтегрування, отримаємо:

$$\sigma_2^2(t) = \sigma_\lambda^2 \left[1 - k_{\hat{a}\hat{e}\hat{a}} (1 - e^{-2t/T_{\hat{a}\hat{e}\hat{a}}}) \right]$$

Повідомлення й завада за умовою задачі незалежні. Тому зміна повної дисперсії помилки описується виразом:

$$\sigma^2(t) = E(t) = \sigma_1^2(t) + \sigma_2^2(t) = \sigma_\lambda^2 \left[1 - (1 - e^{-2t}) \frac{\rho}{(1 + \sqrt{1 + \rho})^2} \right].$$

Зміна нормованої помилки $E(t)/\sigma_\lambda^2$, розрахована за приведеною формулою для декількох значень ρ , показана на рис. 9.7, а пунктиром. На тому ж

малюнку суцільними лініями показані обчислені за формулою (9.95) аналогічні криві для оптимального фільтра зі змінними параметрами. Співставлення кривих, зображених на рис. 9.7, а, дозволяє зробити для розглянутого випадку фільтрації стаціонарного повідомлення наступні висновки:

1) виграш в зменшенні помилки фільтрації, який досягається за рахунок змінності параметрів, збільшується з ростом величини ρ відношення повідомлення/завада;

2) інтервал часу, в межах якого проявляється перевага оптимального фільтру зі змінними параметрами пропорційний до постійної часу T фільтру, що формує повідомлення а, отже, зворотно пропорційний ширині спектра повідомлення;

3) в цілому виграш, який ми отримуємо за рахунок змінності параметрів порівняно невеликий, і для розглянутих співвідношень повідомлення/завада не перевищує 15%.

Сказане не зменшує інтересу до оптимальних фільтрів зі змінними параметрами та методами їх синтезу. Якщо повідомлення чи завада описуються нестационарними випадковими процесами, то оптимальний фільтр навіть в установленому режимі має змінні параметри, й заміна його фільтром з постійними параметрами може привести до істотного погіршення якості фільтрації.

Для того, щоб проілюструвати синтез багатовимірних фільтрів методом просторів станів, розглянемо приклад.

Приклад 9.4. Припустимо, що повідомлення $\lambda(t)$ може бути представлене як процес, що формується на виході двох послідовно включених інтеграторів і описане стохастичним диференціальним рівнянням другого порядку:

$$d^2 \lambda / dt^2 = k_{u2} n(t), \quad (9.99)$$

де $n(t)$ - білий шум із спектральною щільністю $S_c(0)$. Спектральна щільність процесу $\lambda(t)$ дорівнює $S_\lambda(\omega) = k_{u2}^2 S_c(0) / \omega^4$. На вхід оптимізуемого фільтра надходить суміш:

$$\varphi(t) = \lambda(t) + n(t), \quad (9.100)$$

де $n(t)$ - білий шум із спектральною щільністю $S_n(0)$.

Необхідно знайти структуру і параметри оптимального фільтра, який мінімізує середній квадрат помилки відтворення процесу $\lambda(t)$ в будь-який момент часу t .

Для рішення поставленої задачі надамо рівняння (9.97) у вигляді системи рівнянь першого порядку:

$$dx_1/dt = x_2(t), \quad dx_2/dt = k_{u2}n(t),$$

де $x_1(t) = \lambda(t)$. Отриману систему рівнянь запишемо в матричній формі:

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ k_{u2} \end{bmatrix} \chi(t). \quad (9.101)$$

Із співставлення (9.99) і (9.63) випливає, що в даній задачі вектор $\mathbf{x}(t)$ містить дві компоненти $x_1(t) = \lambda(t)$ і $x_2(t)$, а матриці $\mathbf{F}, \mathbf{H}, \mathbf{Q}$ постійні й рівні:

$$\mathbf{F} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 0 \\ k_{u2} \end{bmatrix}, \quad \mathbf{Q} = S_c(0). \quad (9.102)$$

Рівняння (9.98) також запишемо в матричній формі, виразивши $\varphi(t)$ через вектор $\mathbf{x}(t)$:

$$\varphi(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + n(t). \quad (9.103)$$

Із порівняння (9.101) і (9.65) випливає, що в даному випадку матриці $\mathbf{C}$ і $\mathbf{R}$ конкретизуються у вигляді:

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathbf{R} = S_n(0). \quad (9.104)$$

Так як вектор $\mathbf{x}(t)$ в даній задачі містить дві компоненти, то матриця дисперсій помилок $\mathbf{E}(t)$ має розмір 2×2 :

$$E(t) = \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}. \quad (9.105)$$

Синтез фільтра зручно розпочати з визначення матриці коефіцієнтів $G_0(t, t)$. Так як згідно до (9.68) $G_0(t, t) = E(t)C^T R^{-1}$, то, враховуючи (9.103) і (9.102), отримуємо:

$$G_0(t, t) = \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} R^{-1} = \begin{bmatrix} e_{11} \\ e_{21} \end{bmatrix} R^{-1}. \quad (9.106)$$

Щоб знайти структуру оптимального фільтра для даної задачі підставимо в рівняння (9.62) матриці $F, C, G_0(t, t)$, які визначаються виразами (9.100), (9.102), (9.104), і запишемо його у вигляді системи рівнянь першого порядку:

$$\begin{aligned} dx_{10} / dt &= x_{20}(t) + R^{-1}e_{11}[\varphi(t) - x_{10}(t)], \\ dx_{20} / dt &= R^{-1}e_{21}[\varphi(t) - x_{10}(t)]. \end{aligned} \quad (9.107)$$

Структурна схема оптимального фільтра, описуваного рівняннями (9.105), показана на рис. 9.7, б. Залишається знайти коефіцієнти, що входять в неї, $e_{11}(t)$, $e_{21}(t)$. Для цього необхідно розв'язати рівняння (9.70). Обмежимося для простоти розгляданням встановленого режиму, в якому $dE(t)/dt = 0$. Обчислимо із врахуванням (9.72), (9.74) окремі рівняння, що додаються (9.70):

$$\begin{aligned} FE &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix} = \begin{bmatrix} e_{21} & e_{22} \\ 0 & 0 \end{bmatrix}, & EF^T &= \begin{bmatrix} e_{21} & 0 \\ e_{22} & 0 \end{bmatrix}, \\ EC^T R^{-1}CE &= \begin{bmatrix} e_{11}^2 & e_{11}e_{12} \\ e_{21}e_{11} & e_{21}e_{12} \end{bmatrix} R^{-1}, & HQH^T &= Q \begin{bmatrix} 0 & 0 \\ 0 & k_{u2}^2 \end{bmatrix}. \end{aligned}$$

Просумувавши всі доданки (9.70), враховуючи, що матриця E симетрична і $e_{12} = e_{21}$, отримаємо наступну систему алгебраїчних рівнянь:

$$2e_{21} - R^{-1}e_{11}^2 = 0, \quad e_{22} - e_{11}e_{21}R^{-1} = 0, \quad Qk_{u2}^2 - R^{-1}e_{21}^2 = 0.$$

Звідси випливає, що

$$\begin{aligned} e_{21} &= \sqrt{Qk_{u2}^2 R} = \sqrt{k_{u2}^2 S_c(0) S_n(0)}, \\ e_{11} &= \sqrt{2e_{21} R} = \sqrt[4]{4k_{u2}^2 S_c(0) S_n^3(0)}. \end{aligned} \quad (9.108)$$

В установленому режимі комплексний коефіцієнт передачі розімкнутої слідкуючої системи у відповідності з рис. 9.7, б та рівностями (9.106) записується у вигляді:

$$K(j\omega) = \frac{K_a(1 + j\omega T_1)}{(j\omega)^2}, \quad (9.109)$$

де

$$K_a = R^{-1}e_{21} = k_{u2} \sqrt{S_c(0) / S_n(0)}, \quad T_1 = e_{11} / e_{21} = \sqrt[4]{4S_n(0) / k_{u2}^2 S_c(0)}.$$

Таким чином, в даній задачі, оптимальний фільтр в установленому режимі представляє собою слідкуючу систему з двома інтеграторами і демпфіруючим ланцюжком, параметри яких визначаються співвідношеннями (9.107).

9.7. Оптимальний фільтр для повідомлення, описуваного степеневим поліномом

Метод простору станів дозволяє достатньо просто знаходити структуру й параметри оптимального фільтра, який мінімізує помилку фільтрації в будь-який момент часу, й у випадку, коли повідомлення $\lambda(t)$ описується поліномом із випадковими коефіцієнтами

$$\lambda(t) = a_0 + a_1 t + \dots + a_r t^r, \quad (9.110)$$

де $a_0, a_1, \dots, a_r$ - випадкові величини з відомою кореляційною матрицею. Елементами цієї матриці являються значення $M\{a_i a_j\} = \sigma_{ij}^2$. Застосування методу простору станів в даному випадку засновано на можливості представити повідомлення $\lambda(t)$ як процес на виході формуючого фільтра, який складається із $r + 1$ послідовно включених інтеграторів (рис. 9.7, в). Дійсно, якщо внутрішня дія на вході ланцюжка інтеграторів відсутнє $u_{\text{вх}}(t) = 0$ і початкові значення $u_i(0)$ на входах інтеграторів встановлені рівними:

$$u_i(0) = \frac{a_i i!}{k_u^i},$$

то на виході ланцюжка утворюється процес (9.108). Формування повідомлення $\lambda(t)$ схемою, яка показана на рис. 10.7 описується диференціальним рівнянням $d^{r+1} \lambda / dt^{r+1} = 0$. Поклавши $\lambda(t) = x_1(t)$, його можна замінити системою $r + 1$ рівнянь першого порядку:

$$\left. \begin{aligned} dx_1 / dt &= x_2(t), \\ dx_2 / dt &= x_3(t), \\ &\dots\dots\dots \\ dx_{r+1} / dt &= 0. \end{aligned} \right\} \quad (9.111)$$

Записана у векторній формі ця система має вигляд:

$$dx / dt = Fx(t), \quad (9.112)$$

де вектор $x(t)$ і матриця F визначені наступним чином:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{r+1}(t) \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 1 \\ 0 & 0 & \dots & \dots & 0 \end{bmatrix}. \quad (9.113)$$

Значення вектора $\mathbf{x}(t)$ при $t = 0$ виражається через коефіцієнти $a_0, a_1, \dots, a_r$, як випливає із (9.109), (9.108), співвідношенням:

$$\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_{r+1}(0) \end{bmatrix} = \begin{bmatrix} a_0 \\ 1a_1 \\ \vdots \\ r!a_r \end{bmatrix}. \quad (9.114)$$

Співставлення рівнянь (9.63) і (9.101) показує, що випадок, коли повідомлення $\lambda(t)$ описується поліномом, повністю вкладається в рамки загальної постановки задачі при синтезі фільтра методом простору станів. Спрощення в даному випадку полягає в тому, що шум $\chi(t)$ відсутній, отже, матриця $\mathbf{Q} = \mathbf{0}$.

Оптимальний фільтр для виділення поліному можна знайти за загальними формулами (9.67), (9.68), (9.70), покладаючи в них $\mathbf{Q} = \mathbf{0}$. Рівняння (9.70), зокрема, при цьому приймає вигляд:

$$d\mathbf{E}/dt = \mathbf{F}\mathbf{E} + \mathbf{E}\mathbf{F}^T - \mathbf{E}\mathbf{C}^T \mathbf{R}^{-1} \mathbf{C}\mathbf{E}. \quad (9.115)$$

Необхідне нього розв'язку початкове значення $\mathbf{E}(0)$ визначається наступним чином.

Якщо напруга на виході фільтра при $t = 0$ рівна нулю, то значення $\mathbf{E}(0)$ рівне кореляційній матриці значень векторного процесу $\mathbf{x}(0)$ при $t = 0$:

$$\mathbf{E}(0) = \mathbf{R}_\lambda(0) \mathbf{M} \{ \mathbf{x}(0) \mathbf{x}^T(0) \}, \quad (9.116)$$

де $\mathbf{R}_\lambda(0)$ - матриця розміром $(r+1) \times (r+1)$. Елемент цієї матриці r_{ij} , який стоїть на перетині i -го рядка та j -го стовпця, рівний $r_{ij} = M \{ x_i(0) x_j(0) \}$. На основі цього (9.117) він визначається через відомий момент σ_{ij}^2 коефіцієнтів a_i, a_j співвідношенням:

$$r_{ij} = M \{ i! a_i j! a_j \} = i! j! \sigma_{ij}^2. \quad (9.118)$$

Таким чином, матриця $\mathbf{E}(0)$ повністю визначається за відомими значеннями σ_{ij}^2 .

Матрицю $\mathbf{E}(t)$ в даному випадку можна знайти, розв'язуючи рівняння шляхом переходу до системи лінійних рівнянь (10.72) та (10.73). Нагадаємо, що при цьому в процесі вирішення виникає необхідність визначити перехідну матрицю $\mathbf{T}(t, t_0)$ розміром $2n \times 2n$, де $n = r + 1$. Якщо для її пошуків використовується формула (9.81), то при обчисленні зворотної матриці $[\mathbf{sI} - \mathbf{A}]^{-1}$ зручно представити $\mathbf{sI} - \mathbf{A}$ у вигляді блочної матриці:

$$[\mathbf{sI} - \mathbf{A}] = \begin{bmatrix} \mathbf{m} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}.$$

Тоді відповідно [68]:

$$[s\mathbf{I} - \mathbf{A}] = \begin{bmatrix} \mathbf{M} & \mathbf{K} \\ \mathbf{B} & \mathbf{D} \end{bmatrix},$$

де

$$\begin{aligned} \mathbf{M} &= [\mathbf{m} - \mathbf{b}\mathbf{d}^{-1}\mathbf{c}]^{-1}, & \mathbf{B} &= -\mathbf{d}^{-1}\mathbf{c}\mathbf{M}, \\ \mathbf{D} &= [\mathbf{d} - \mathbf{c}\mathbf{m}^{-1}\mathbf{b}]^{-1}, & \mathbf{K} &= -\mathbf{m}^{-1}\mathbf{b}\mathbf{D}. \end{aligned} \quad (9.119)$$

Так як в задачі фільтрації повідомлення, описуваного поліномом $\mathbf{Q} = \mathbf{0}$, то в матриці $s\mathbf{I} - \mathbf{A}$, як видно із (9.80), блок $\mathbf{b} = \mathbf{0}$. Рівності (9.116) при цьому приймають дуже простий вигляд:

$$\mathbf{M} = -\mathbf{m}^{-1}, \quad \mathbf{B} = -\mathbf{d}^{-1}\mathbf{c}\mathbf{M}, \quad \mathbf{D} = \mathbf{d}^{-1}, \quad \mathbf{K} = \mathbf{0},$$

і процедура відшукування матриці $\mathbf{E}(t)$ спрощується.

Кореляційну матрицю $\mathbf{E}(t)$ помилок фільтрації при $\mathbf{Q} = \mathbf{0}$ можна знайти також за формулою:

$$\mathbf{E}(t) = \mathbf{\Omega}(t,0)\mathbf{P}^{-1}(t)\mathbf{\Omega}^T(t,0), \quad (9.120)$$

де

$$\mathbf{P}(t) = \mathbf{E}^{-1}(0) + \int_0^t \mathbf{\Omega}^T(\tau)\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\mathbf{\Omega}(\tau)d\tau;$$

$\mathbf{\Omega}(t,0)$ - перехідна матриця формуючого фільтра, рівна $\mathbf{\Omega}(t,0) = e^{\mathbf{F}t}$. Рівність (9.117) є приведеним в [69] розв'язком рівняння (9.113), вираженим через легко обчислювану матрицю $\mathbf{\Omega}(t,0)$. Розмір всіх матриць, які входять в (9.117), не перевищує $n \times n$.

Обсяг обчислень при відшукуванні матриці $E(t)$ обома способами приблизно однаковий. Розглянемо в якості прикладу побудову оптимального фільтра для повідомлення $\lambda(t)$, яке описується поліномом першого порядку.

Приклад 9.5. Нехай повідомлення

$$\lambda(t) = a_0 + a_1 t, \quad (9.121)$$

де a_0, a_1 - випадкові незалежні величини з нульовими математичними очікуваннями та дисперсіями σ_0^2 і σ_1^2 відповідно. На вхід фільтра надходить суміш:

$$\varphi(t) = \lambda(t) + n(t),$$

де $n(t)$ - білий шум зі спектральною щільністю $S_n(0)$ і функцією кореляції $R(\tau) = S_n(0)\delta(\tau)$.

Фільтр повинен забезпечити мінімум середнього квадрату помилки в будь-який момент часу t . Так як порядок поліному (9.18) $r = 1$, то вектор $\mathbf{X}(t)$ і матриця $\mathbf{F}$ в даному прикладі у відповідності з (9.11) приймають вигляд:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (9.122)$$

Суміш $\varphi(t)$, виражена через векторний процес $\mathbf{X}(t)$, записується при цьому у вигляді:

$$\varphi(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + n(t), \quad x_1(t) = \lambda(t).$$

Звідси після співставлення з (9.65) випливає, що

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathbf{R} = S_n(0). \quad (9.123)$$

Вектор $\mathbf{X}(t)$ і матриці $\mathbf{F}, \mathbf{C}, \mathbf{R}$, як впливає із (9.119), (9.120), (9.100), (9.102), такі ж, як і в прикладі 9.4, тому знайдені там диференціальні рівняння (9.105) й структурна схема (рис. 9.7) оптимального фільтра залишаються справедливими і для даного випадку. Однак коефіцієнти $e_{11}(t)$ і $e_{21}(t)$ ми отримуємо іншими. Для їх визначення необхідно розв'язати рівняння (9.113).

Застосуємо для обчислення матриці $\mathbf{E}(t)$ формулу (9.117). Знайдемо спочатку перехідну матрицю $\mathbf{\Omega}(t,0) = e^{\mathbf{F}t}$. У відповідності з (9.79), (9.81), (9.118):

$$\mathbf{\Omega}(t,0) = L^{-1} \{ [s\mathbf{I} - \mathbf{F}]^{-1} \},$$

де

$$[s\mathbf{I} - \mathbf{F}] = \begin{bmatrix} s & -1 \\ 0 & s \end{bmatrix},$$

а отже

$$[s\mathbf{I} - \mathbf{F}]^{-1} = \begin{bmatrix} 1/s & 1/s^2 \\ 0 & 1/s \end{bmatrix}. \quad (9.124)$$

Визначивши зворотнє перетворення Лапласа для кожного елемента матриці (9.121) отримаємо:

$$\mathbf{\Omega}(t,0) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}. \quad (9.125)$$

Перемноження матриць $\mathbf{\Omega}^T, \mathbf{C}^T, \mathbf{R}^{-1}, \mathbf{C}, \mathbf{\Omega}$ і наступне інтегрування дає наступний результат:

$$\int_0^t \mathbf{\Omega}^T \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \mathbf{\Omega} d\tau = \mathbf{R}^{-1} \begin{bmatrix} t & t^2/2 \\ t^2/2 & t^3/3 \end{bmatrix}. \quad (9.126)$$

Так як коефіцієнти a_0 і a_1 за умовою задачі незалежні величини з дисперсіями σ_0^2 і σ_1^2 , то матриця $\mathbf{E}(0)$ відповідно до (9.114) має вигляд:

$$\mathbf{E}(0) = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}. \quad (9.127)$$

Із (9.117), (9.123) і (9.124) випливає, що

$$\mathbf{P}^{-1}(t) = \frac{1}{\Delta} \begin{bmatrix} 1/\sigma_1^2 + t^3/3R & -t^2/2R \\ -t^2/2R & 1/\sigma_0^2 + t/R \end{bmatrix}, \quad (9.128)$$

$$\Delta = (1/\sigma_0^2 + t/R)(1/\sigma_1^2 + t^3/3R) - t^4/4R^2.$$

Підставляючи (9.122) і (9.125) в формулу (9.117) знаходимо матрицю $\mathbf{E}(t)$. Її елементи $e_{11}(t)$ і $e_{21}(t)$, що визначають змінні коефіцієнти передачі в структурній схемі оптимального фільтра (рис. 10.6), рівні:

$$e_{11}(t) = \frac{\sigma_0^2 + \sigma_1^2 t^2 + \frac{\sigma_0^2 \sigma_1^2 t^3}{3S_n(0)}}{1 + \frac{\sigma_0^2 t}{S_n(0)} + \frac{\sigma_1^2 t^3}{3S_n(0)} + \frac{\sigma_0^2 \sigma_1^2}{12S_n^2(0)} t^4},$$

$$e_{21}(t) = \frac{\sigma_1^2 t + \frac{\sigma_0^2 \sigma_1^2 t^2}{3S_n(0)}}{1 + \frac{\sigma_0^2 t}{S_n(0)} + \frac{\sigma_1^2 t^3}{3S_n(0)} + \frac{\sigma_0^2 \sigma_1^2}{12S_n^2(0)} t^4}.$$

Як видно з отриманих виразів, коефіцієнти $R^{-1}e_{11}(t)$, $R^{-1}e_{21}(t)$ і дисперсія помилки виділення повідомлення $\lambda(t)$, рівна $e_{11}(t)$, зі зростанням t прямує до нуля. Пояснюється це тим, що коефіцієнти a_0 і a_1 постійні в часі, й при

необмеженому зростанні часу спостереження можуть бути виміряні з як завгодно високою точністю.

9.8. Поняття про оптимальну нелінійну фільтрацію

У п.9 при розгляді оптимальної нелінійної фільтрації розглянуті схеми слідкуючого прийому (мал. 9.1, б, в), які забезпечують мінімум середньої квадратичної помилки.

Синтез системи, що здійснює нелінійне перетворення сигналу й забезпечує найбільшу точність слідкування за корисним параметром сигналу (повідомленням), який змінюється за час спостереження, може бути проведений методами оптимальної нелінійної фільтрації.

В теорії оптимальної нелінійної фільтрації поєдналися два основні підходи. Перший з них, запропонований Р.Л. Стратоновичем, заснований на описі параметра сигналу, який фільтрується, за допомогою Марковського процесу. Цей підхід дозволив вирішити ряд важливих і цікавих задач синтезу оптимальних нелінійних систем. Великий внесок у розвиток цього напрямку оптимальної нелінійної фільтрації та його застосування до синтезу оптимальних прийомних пристроїв при різних методах модуляції сигналу внесли роботи В. І. Тихонова.

Другий підхід, розроблений І. А. Більшаковим і В. Г. Репіним, заснований на припущеннях про нормальність процесу $\lambda(t)$ та високої точності його виділення, які дозволяють застосувати гаусову апроксимацію багатовимірного апостеріорного закону розподілу параметра $\lambda(t)$. В роботах такий підхід одержав назву квазілінійного методу. При проведенні синтезу системи оптимальної нелінійної фільтрації даним методом вхідна напруга записується у вигляді:

$$u_{\text{д\o}}(t) = u_c(t, \lambda, \lambda_1) + u_{\text{ш}}(t),$$

де $u_c(t, \lambda, \lambda_1)$ - напруга сигналу; $u_{\text{ш}}(t)$ - білий шум зі спектральною щільністю $S(\omega) = S(0)$; $\lambda(t)$ - корисний параметр сигналу, за яким ведеться слідкування; $\lambda_1(t)$ - вектор невідомих параметрів сигналу, які не несуть інформації (їх іноді називають заважаючими або супроводжуючими).

При спрощуючому припущенні, що час кореляції заважаючих параметрів сигналу, набагато менший від часу кореляції корисного параметра (повідомлення), оптимальна система нелінійної фільтрації представляє собою замкнуту слідкуючу систему, функціональна схема якої зображена на рис.9.8, а.

Параметр $\lambda(t)$, за яким ведеться спостереження (частота, фаза, тимчасове положення) нелінійно зв'язаний з радіосигналом $u_c(t, \lambda)$, що діє на вході радіотехнічної слідкуючої системи. Тому слідкуюча система повинна містити елементи, які здійснюють нелінійне перетворення радіосигналу. Задача оптимізації структури всієї системи за критерієм мінімуму помилки спостереження повинна вирішуватися при цьому на основі теорії оптимальної нелінійної фільтрації.

Нелінійну операцію виділення із сигналу інформації про величину неузгодженості виконує дискримінатор (рис.9.8, а). Якщо дискримінатор вже обраний або заданий, то задача синтезу звужується й зводиться до оптимізації структури фільтра. Під фільтром тут розуміється вся інерційна частина системи, що стежить за дискримінатором.

Припустимо, що помилка спостереження не виходить за межі лінійної ділянки характеристики дискримінатора, і представимо вихідну напругу дискримінатора у вигляді:

$$u_{\text{диск}}(t) = S_{\text{диск}} x(t) + \xi(t),$$

де $S_{\text{диск}}$ - крутизна дискримінаційної характеристики; x - помилка спостереження (неузгодженість); $\xi(t)$ - флюктуаційна напруга на виході дискримінатора, яку приймемо незалежною від помилки слідкування.

У загальному випадку оптимальний фільтр у контурі слідкуючої системи (рис.9.8, а) може бути нелінійним. Введемо ще одне обмеження і будемо шукати оптимальний фільтр у класі лінійних пристроїв. При цьому слідкуюча система, яка розглядається, прийме вигляд, показаний на рис.9.8, б де $g_{\delta}(t, u)$ - імпульсна перехідна функція оптимізуемого фільтра, який може мати змінні параметри.

З точки зору формування процесу $y(t)$ слідкуюча система, показана на рис.9.8, б, може бути замінена еквівалентним фільтром (мал. 9.8), в з імпульсною перехідною функцією $g(t, u)$, на вході якого діє суміш:

$$\varphi(t) = \lambda(t) + \xi(t)/S_{\hat{\mathbf{a}}} = \lambda(t) + n(t). \quad (9.129)$$

Завада $n(t)$ відповідає флюктуаційному коливанию (шуму), переліченому до входу дискримінатора. Імпульсна перехідна функція $g(t, u)$ еквівалентного фільтра зв'язана з аналогічною функцією $g_{\hat{\mathbf{d}}}(t, u)$ фільтра в контурі слідкуючої системи інтегральним рівнянням

$$S_{\hat{\mathbf{d}}} g_{\hat{\mathbf{d}}}(t, u) = g(t, u) + S_{\hat{\mathbf{d}}} \int_u^t g(\theta, u) g_{\hat{\mathbf{d}}}(t, \theta) d\theta. \quad (9.130)$$

Дискримінатор оптимальної системи виконує нелінійне перетворення вхідного сигналу, в результаті якого формується напруга $z(t)$, що описується виразом:

$$z(t) = \left. \frac{\partial Q_1(u_{\hat{\mathbf{a}}\hat{\mathbf{o}}}, \lambda, t)}{\partial \lambda} \right|_{\lambda=\lambda_0}, \quad (9.131)$$

де λ_0 - оцінка параметра $\lambda(t)$; $Q_1(u_{\hat{\mathbf{a}}\hat{\mathbf{o}}}, \lambda, t)$ - функція, що визначається співвідношенням:

$$\int_{t-\Delta t}^t Q_1(u_{\hat{\mathbf{a}}\hat{\mathbf{o}}}, \lambda, t) dt = \ln P_{\lambda}(u_{\hat{\mathbf{a}}\hat{\mathbf{o}}}), \quad (9.132)$$

$P_{\lambda}(u_{\hat{\mathbf{a}}\hat{\mathbf{o}}})$ - функція правдоподібності, тобто залежність від λ багатовимірної щільності імовірності реалізації вхідного процесу $u_{\hat{\mathbf{a}}\hat{\mathbf{o}}}(t)$, розглянутої на інтервалі

часу Δt . Інтервал Δt вибирається значно більшим від часу кореляції заважаючих параметрів, але значно меншим, ніж час кореляції корисного параметра сигналу.

Функція правдоподібності $P_{\lambda}(u_{\hat{\alpha}})$ обчислюється шляхом усереднення функції правдоподібності, записаної для сигналу, всі параметри якого, крім λ , відомі точно. Усереднення проводиться за всіма можливими значеннями заважаючих параметрів, з врахуванням їх апіорної щільності ймовірності.

Оскільки відповідно до (9.128) похідна функція Q_1 обчислюється в точці $\lambda = \lambda_0$, до дискримінатора (рис.9.8, а) підводиться також оцінка $\lambda_0(t)$. Фільтр, слідкуючий в оптимальній системі за дискримінатором, є лінійним, і в загальному випадку нестационарним. Його імпульсна перехідна функція $g_{\hat{\delta}}(t, u)$ збігається з імпульсною перехідною функцією оптимального фільтра в контурі лінійної слідкуючої системи, зображеної на рис. 9.8, б. При цьому в схемі на рис. 9.8, б варто прийняти:

$$S_{\hat{\alpha}} = K, \quad S_{\xi}(\omega) = S_{\xi}(0) = K, \quad (9.133)$$

де $S_{\xi}(0)$ - спектральна щільність процесу $\xi(t)$.

$$K = -\overline{\partial^2 Q_1} / \partial \lambda^2. \quad (9.134)$$

Методика визначення імпульсної перехідної функції $g_{\hat{\delta}}(t, u)$ оптимального фільтра в контурі лінійної слідкуючої системи докладно викладена в [1]. Як впливає з її змісту, для відшукування оптимальної функції $g_{\hat{\delta}}(t, u)$ необхідно спочатку, вирішивши інтегральне рівняння, знайти функцію $g(t, u)$ еквівалентного лінійного фільтра. Потім за знайденою $g(t, u)$ функція $g_{\hat{\delta}}(t, u)$ визначається в результаті рішення інтегрального рівняння (9.127). Якщо повідомлення $\lambda(t)$ і завада $u_{\theta}(t)$ стаціонарні, і потрібно мінімізувати помилку фільтрації тільки в установленому режимі, то оптимальний лінійний фільтр має постійні параметри, і його комплексний коефіцієнт передачі $K(j\omega)$ записується у вигляді

$$K(j\omega) = \frac{1}{k} \cdot \frac{\Phi(j\omega)}{1 - \Phi(j\omega)}, \quad (9.135)$$

де $\Phi(j\omega)$ - коефіцієнт передачі оптимального лінійного фільтра, який виділяє повідомлення $\lambda(t)$ із суміші з білим шумом, спектральна щільність якого дорівнює $1/k$.

Достоїнством методу запропонованого І.А. Більшаковим і В.Г. Репіним, є можливість роздільного синтезу дискримінатора й лінійного фільтра. При зміні спектра корисного параметра $\lambda(t)$ дискримінатор оптимальної системи залишається незмінним, і заново доводиться шукати тільки оптимальний лінійний фільтр. Марковський підхід до теорії оптимальної нелінійної фільтрації носить більш загальний характер, оскільки принципово його можна застосувати у випадку негаусової апостеріорної щільності імовірності виділюваного параметра. Більшість практичних результатів при марківському підході отримані поки в рамках гаусового наближення, при якому апостеріорна щільність імовірності процесу $\lambda(t)$ також покладається гаусовою. Однак і в цьому випадку можна вказати коло задач, при рішенні яких марківський підхід має певні переваги. Трудомісткість синтезу на основі марківського підходу і квазілінійним методом приблизно однакова, якщо виконуються наступні умови:

- 1) повідомлення та завада є стаціонарними випадковими процесами, заданими своїми кореляційними функціями чи енергетичними спектрами;
- 2) потрібно мінімізувати помилку фільтрації в установленому режимі;
- 3) всі параметри сигналу, крім корисного, несучого інформацію про повідомлення, відомі.

Деяку перевагу при виконанні цих умов, очевидно можна віддати квазілінійному методу. Якщо умови 1 і 2 не виконуються, то оптимальні згладжуючі ланцюги слідкуючої системи повинні бути нестаціонарними. Синтез їх на основі марківського підходу виявляється більш простим. Відмічене положення має перевагу синтезу оптимальних лінійних нестаціонарних фільтрів методом простору станів у порівнянні з безпосереднім рішенням інтегральних рівнянь Віннера - Хопфа.

Якщо заважаючі параметри сигналу змінюються в часі, і період їх зміни чи час кореляції порівняні з часом кореляції корисного параметра сигналу, то переважним є проведення синтезу на основі марківської теорії.

Метод синтезу, заснований на марківській теорії, зручний також у тих випадках, коли потрібно з мінімальною помилкою відфільтрувати декілька параметрів (або кілька складових параметра) сигналу.

Контрольні питання

1. Якій умові відповідає оптимальний приймач неперервних сигналів?
2. В якому випадку кореляційна обробка прийнятих сигналів є оптимальною?
3. В чому полягає принцип слеящего приема?
4. В якому випадку теорію оптимального прийому можна розглядати як теорію оптимальної нелінійної фільтрації?
5. Поясніть постановку задачі синтезу оптимальних пристроїв (систем зв'язку).
6. Поясніть фізичний зміст оптимальної смуги перепуску пристрою зв'язку.
7. Яким чином обчислюються значення оптимальних параметрів (систем) зв'язку із умови мінімуму середньої квадратичної помилки?
8. Поясніть постановку задачі оптимальної фільтрації Колмогорова-Віннера та умови фізичної реалізованості оптимального рішення.
9. Які недоліки теорії лінійної фільтрації Колмогорова-Віннера для стаціонарних процесів?
10. В чому полягають особливості методу просторів станів при синтезі фільтрів Колмогорова?
11. Запишіть рівняння оптимального фільтра.
12. В чому полягають особливості оптимального фільтра для повідомлення, описуваного степеневим поліномом?
13. Які два підходи в теорії нелінійної оптимальної фільтрації Ви знаєте?
14. Які достоїнства методу І.А. Більшакова та В.Г. Рєпіна?
15. Які переваги методу синтезу, основанийого на марківській теорії?

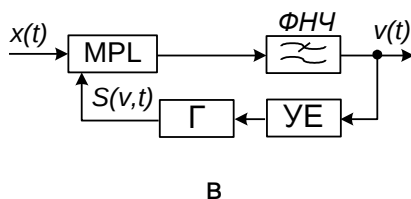
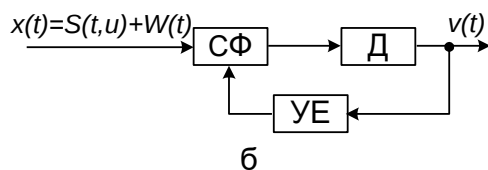
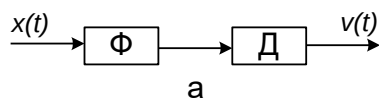
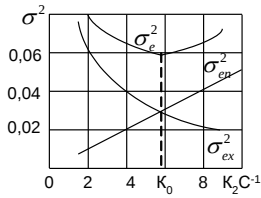
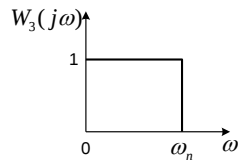


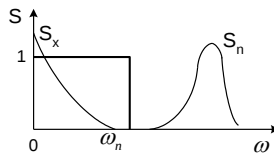
Рис.9.1. Структурні схеми: а - оптимального приймача неперервних сигналів; б - приймача зі слідкуючим фільтром; в - слідкуючого кореляційного приймача



а

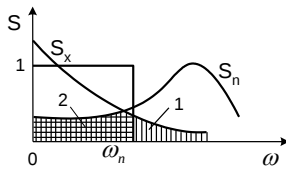


б

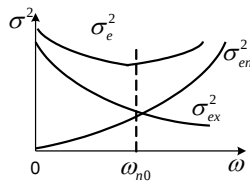


в

Рис.9.2. До визначення оптимального коефіцієнта підсилення пристрою (системи) зв'язку (а) та оптимальної смуги перепуску при спектрах сигналу й завади, які не перехрещуються, (в) та ідеальна ЛЧХ (б)

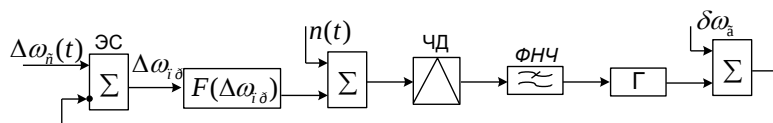


а

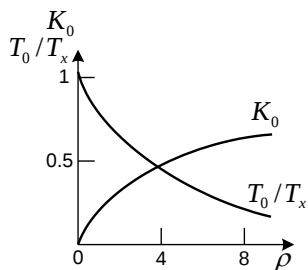


б

Рис.9.3. До пояснення оптимальної смуги перепуску пристрою (системи) зв'язку (а) та залежність середньої квадратичної помилки пристрою зв'язку від частоти (б)



а



б

Рис.9.4. Функціональна схема системи автоматичного підстроювання частоти (а) та залежність оптимальних параметрів системи від відношення сигнал/шум (б)

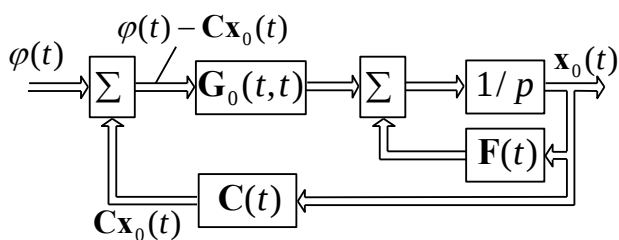


Рис.9.5. Структурна схема фільтру, побудованого за рівнянням (9.47)

РОЗДІЛ 10. КОДУВАННЯ ДЖЕРЕЛ ІНФОРМАЦІЇ ТА КАНАЛІВ ЗВ'ЯЗКУ

10.1. Основні поняття та визначення

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 кодовою 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), що виникають через перешкоди, - коригувальні 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 взаємодозначне 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-Фанно*.

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 якісь 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 коригувальних кодів. *Зниження швидкості передавання інформації* при роботі з каналами 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 робота на великій відстані, а також робота по каналах зі звуженою смугою пропускання (300-2700 Гц).

Застосування коригувальних кодів є ефективнішим методом підвищення надійності, ніж робота зі зниженою швидкістю. Зазначимо, що обидва методи не слід протиставляти один одному. Використання коригувальних кодів є методом більш загальним, а зниження швидкості передавання можна вважати окремим випадком застосування коригувального коду. Розглянемо деякі поняття, пов'язані з коригувальними кодами.

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. *Коригувальні (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 використанні коригувальних кодів помилка в одному розряді спричинює заміну дозволеної кодової комбінації недозвільною, 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 дією помилок більшої кратності, що є результатом найбільш інтенсивних перешкод.

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) відбулася одинична помилка, внаслідок чого спотворився перший розряд, і прийнята комбінація 1011. Ця комбінація є недозвільною, що свідчить про наявність у ній 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, прийнята комбінація 1011. Ця комбінація відрізняється від дозволеної комбінації 0011 одним розрядом, а від іншої дозволеної 1100 - трьома. Отже, прийнята комбінація "ближче" 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 досить трьох розрядів, а не чотирьох. Отже, підвищення завадостійкості потребує введення додаткового розряду. Другий код володіє ще більшою завадостійкістю, і це потребує ще більшого надлишку - трьох додаткових розрядів.

Коригувальні 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 коригувальні коди. Останні поділяються на два класи - блокові і неперервні (рис. 10.1). 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 і нероздільні.

Роздільними називаються коди, у 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, у 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 Хемінга і циклічним.

Кількість розрядів, якими різняться дві кодові комбінації, називається *ковою 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 *хемінгою відстанню*. Так, у трирозрядному коді з дозволеними комбінаціями 101, 110, 011, 000 мінімальна кодова відстань дорівнює двом, для простих кодів - одиниці.

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

$$d_0 \geq \sigma + 1 \quad (10.1)$$

$$d_0 \geq 2t + 1 \quad (10.2)$$

де σ - кількість помилок, що виявляються; t - кількість помилок, що виправляються; d_0 - мінімальна кодова відстань.

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

$$d_0 \geq \sigma + t + 1 \quad (10.3)$$

Мінімальна кодова відстань лише частково характеризує коригувальні властивості коду, оскільки звичайно забезпечується виправлення і виявлення помилок і, більш високої кратності, ніж обумовлене співвідношеннями (10.1) - (10.3).

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. *Надлишком 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. Найбільш близьке наближення звичайно забезпечує критерій Варшавова

$$r \geq \log_2 \left(1 + \sum_{i=1}^{d_0-2} C_{n-1}^i \right) \quad (10.4)$$

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

$$r \geq \log_2(n+1) \quad (10.5)$$

причому r - найменше ціле число, при якому задовольняється нерівність (10.5).

10.2. Коригувальні 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. У систематичних кодах перевірні розряди є лінійною комбінацією інформаційних. Це означає, що значення будь-якого перевірного розряду $b_i = c_{i1} a_1 \oplus c_{i2} a_2 \oplus \dots \oplus c_{ik} a_k$, $c_{i1}, c_{i2}, \dots, c_{ik}$ - 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:

$$a_1 a_2 \dots a_k b_1 b_2 \dots b_r \oplus a'_1 a'_2 \dots a'_k b'_1 b'_2 \dots b'_r =$$

$$(a_1 \oplus a'_1)(a_2 \oplus a'_2) \dots (a_k \oplus a'_k)(b_1 \oplus b'_1)(b_2 \oplus b'_2) \dots (b_r \oplus b'_r).$$

де

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

Неважко бачити, що перевірни 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. Ця обставина дає можливість визначити всі дозвалені кодові комбінації, користуючись лише обмеженою кількістю дозволених комбінацій.

Усі дозвалені 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 комбінації. Кожне додавання дає нам нову дозволена комбінацію.

Процес побудови множини кодових комбінацій можна виразити як $X_i = c_1 X_1 \oplus c_2 X_2 \oplus \dots \oplus c_g X_g$ де $x_1, \dots, x_g$ - початкові кодові комбінації; $c_1, \dots, c_g$ - коефіцієнти, що набувають значення 1 або 0. Надаючи різним коефіцієнтам c_i , значення 1 або 0, можна скласти вихідні комбінації у 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

$$c_g^0 + c_g^1 + c_g^2 + \dots + c_g^g = 2^k \quad (10.6)$$

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

Співвідношення (10.6) є слухним тоді, коли комбінації, що маємо 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 противному 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 випадку 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), що складає k рядків і n стовпців і називається *породжувальною матрицею*:

$$G = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_{11} & b_{12} & \dots & b_{1r} \\ a_{21} & a_{22} & \dots & a_{2k} & b_{21} & b_{22} & \dots & b_{2r} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_{k1} & b_{k2} & \dots & b_{kr} \end{pmatrix}$$

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:

$$G = \left| E_k, C_{r,k} \right| = \begin{vmatrix} 10...b_{11}b_{12}...b_{1r} \\ 01...b_{21}b_{22}...b_{2r} \\ \\ 00...1b_{k1}b_{k2}...b_{k3} \end{vmatrix} \quad (10.7)$$

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 так званій канонічній формі.

Розглянемо тепер приклад побудови систематичного коду. Нехай треба побудувати код $n=7$, що забезпечує виправлення одиничної помилки.

Відповідно до формули (10.2) $d_0=3$. Користуючись нерівністю (10.5), знайдемо $r=3$. Відповідно $k=4$. Побудуємо твірну матрицю коду (7,4):

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

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 менше 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:

$1 \oplus 2$	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

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 так званого перевірного співвідношення, а визначення всіх r перевірних розрядів вимагає r перевірних співвідношень. Перевірні співвідношення прийнято записувати одне під одні 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_{kr} , що містить k стовпців і r рядків, причому кожен її рядок відповідає стовпцю перевірних розрядів породжувальної матриці в канонічній формі:

$$H = |D_{k:r}, E_r| = \begin{vmatrix} b_{11} & b_{21} & \dots & b_{k1} & 10 & \dots & 0 \\ b_{12} & b_{22} & \dots & b_{k2} & 01 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{1r} & b_{2r} & \dots & b_{kr} & 00 & \dots & 1 \end{vmatrix} \quad (10.8)$$

Матриця (10.8) є перевіркою. За її допомогою операція кодування здійснюється дуже просто. Позиції, займані одиницями в i -му рядку приписаної частини перевіркої матриці, визначають ті інформаційні розряди, що мають брати участь у формуванні i -го перевірного розряду. Наприклад, якщо перший рядок перевіркої матриці має вигляд 1011100, то перевірний 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 (10.8). 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 основі 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 комбінацією синдрому і його 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 $2^k - 1$ комбінацій 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 комбінація помилок, що виправляється, 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

$$H = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{vmatrix} \quad (10.9)$$

Тому

$$a_5 = a_2 \oplus a_3 \oplus a_4;$$

$$a_6 = a_1 \oplus a_3 \oplus a_4;$$

$$a_7 = a_1 \oplus a_2 \oplus a_4.$$

де $a_1, \dots, a_4$ - інформаційні розряди; a_5, a_6, a_7 - перевірни.

Декодування відбувається обчисленням перевірних співвідношень:

$$\begin{aligned} & a_2 \oplus a_3 \oplus a_4 \oplus a_5 \\ & a_1 \oplus a_3 \oplus a_4 \oplus a_6 \\ & a_1 \oplus a_2 \oplus a_4 \oplus a_7 \end{aligned} \quad (10.10)$$

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 (10.10):

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

Отже, при наявності помилки в 1-му розряді синдром дорівнює 011. Аналогічно можна одержати види синдромів і при наявності всіх інших можливих однократних помилок (див. нижче).

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

Нехай, наприклад, у кодувальний пристрій надійшла інформаційна послідовність 1101. Згідно з (10.9) ця послідовність кодується так: 1101001. Тепер припустимо, що при передаванні відбулася помилка в 2-му розряді і комбінація 1101001 прийнята як 1001001. При декодуванні відповідно до (10.10) будуть отримані такі результати:

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

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, що має два різновиди. В одному з них є S -кратне повторення комбінації простого коду $a_1, a_2, \dots, a_k$:

$$\underbrace{a_1 a_1 \dots a_k}_1 \underbrace{a_1 a_2 \dots a_k}_2 \dots \underbrace{a_1 a_2 \dots a_k}_S$$

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

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

Код з повторенням має довжину $n = S \cdot k$, число перевірних розрядів $r = (S-1) \cdot k$, 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 відстоять 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 виявляються при дії пакетів помилок. Так, при $S = 2$ завжди виявляються пакети помилок завдовжки до $n/2$. Виявляються також усі помилки непарної кратності. При $S = 2$ перевірна матриця має вигляд:

$$H = \begin{vmatrix} E_k & E_k \end{vmatrix} \quad (10.11)$$

Недоліком кодів з повторенням є дуже високий 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:

$$H = \begin{vmatrix} \bar{E}_k & E_k \end{vmatrix} \quad (10.11)$$

де $\bar{E}_k$ - 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 більш складні 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

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

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

$$H = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \end{bmatrix} \quad (10.13)$$

Як і код із 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. При 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

$$\begin{aligned} b_1 &= a_1 \oplus a_2 \oplus a_3 \oplus a_4 \oplus a_5, \\ b_2 &= a_1 \oplus a_3 \oplus a_5 \end{aligned}$$

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:

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

Коди Хемінга

Кодами Хемінга 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 виявляють усі подвійні помилки. Коди Хемінга з мінімальною кодовою відстанню $d_0 = 3$ мають довжину $n \leq 2^r - 1$. Відповідно

$$2^k \leq \frac{2^n}{n+1} \quad (10.14)$$

Характерною рисою перевірної матриці цього коду з $d_0 = 3$ є те, що її стовпці - це будь-які різні ненульові комбінації завдовжки r . Наприклад, для коду Хемінга (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 перевірної матриці можна довільно переставляти, при цьому коригувальні властивості коду не зміняться.

Коди Хемінга з мінімальною кодовою відстанню $d_0 = 4$ утворюються на основі розглянутих кодів з $d_0 = 3$ введенням ще одного перевірного розряду, що доповнює кодову комбінацію до парного числа одиниць. Перевірна матриця даного коду також утвориться з перевірної матриці коду Хемінга з $d_0 = 3$ приписуванням праворуч нульового стовпця і приписуванням додаткового рядка, що складається тільки з одиниць. Перевірна матриця, наприклад коду (10,4), 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 Хемінга з $d_0 = 4$, можливі три випадки:

- помилок немає, тоді всі перевірки означають відсутність помилок;
- одинична помилка, тоді всі перевірки означають наявність помилки;
- подвійна помилка, тоді 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 споживачу спотвореного повідомлення, тобто знайти подвійні помилки.

Іноді під кодами Хемінга 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 вид 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.

10.3. 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 коригувальні властивості, що і забезпечило їм широке поширення на практиці.

При побудові циклічних кодів кодові комбінації подають у вигляді поліномів

$$G(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0 \quad (10.15)$$

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 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 [8]. Дійсно,

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

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

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

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

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

$$G = \begin{vmatrix} P(x) \\ xP(x) \\ x^2P(x) \\ \vdots \\ x^{k-1}P(x) \end{vmatrix} \quad (10.16)$$

Безпосередньо з твірної матриці (10.16) випливає, що всі дозволені кодові комбінації циклічного коду без остачі діляться на вироблюваний 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:

$$\begin{array}{r} \oplus 1101001 \mid 111 \\ \underline{111} \mid 10101 \\ \oplus 110 \mid \\ \underline{111} \mid \\ \oplus 101 \mid \\ \underline{111} \mid \\ 10 \mid \end{array}$$

Двоцлен 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 дозоволених.

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 $R(x)$:

$$\frac{x^r G(x)}{P(x)} = Q(x) \oplus \frac{R(x)}{P(x)} \quad (10.17)$$

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

Помноживши обидві частини рівності (10.17) на $P(x)$, маємо

$$F(x) = Q(x)P(x) = x^r G(x) \oplus R(x) \quad (10.18)$$

Отже, кодову комбінацію циклічного коду можна отримати двома 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 змогу одержати твірну матрицю відразу в канонічній формі:

$$G = \begin{bmatrix} E_k & C_{r;k} \end{bmatrix}$$

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:

$$h(x) = \frac{x^n + 1}{P^{-1}(x)} \quad (10.19)$$

де $P^{-1}(x)$ - поліном, об'єднаний з твірним поліномом $P(x)$. Нагадаємо, що в об'єднаних поліномах члени розташовані в зворотному порядку. Так,

наприклад, поліноми 100111 і 111001 є об'єднаними. Перший рядок перевіркої матриці циклічного коду є перевірним поліномом $h(x)$, помноженим на x^{r-1} (тобто доповнений праворуч $(r-1)$ -нулями). Наступні рядки перевіркої матриці є циклічним зсувом першої.

Перевірні матриця може бути також побудована звичайним способом, виходячи з твірної матриці. Побудована таким способом перевірна матриця зовні може відрізнятися від побудованої за допомогою перевірного полінома, однак обидві матриці завжди можуть бути зведені до одного виду.

Наведемо приклад побудови семирозрядного циклічного коду з $d_0 = 3$. Для цього потрібні три перевірних розряди ($d_0 = 3$) і утворювальний 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 .

Коригувальна 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.

Циклічний код з $d_0 = 3$ є різновидом кодів Хемінга. 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 на стор. 393. Відомості про незвідні поліноми більш високих степенів є в літературі [1,3].

Циклічний код з $d_0 = 4$ є також різновидом кодів Хемінга 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 $r = m+1$. Так, наприклад, при $n = 15$ твірний поліном має вигляд $(x+1)(x^3 + x + 1) = x^4 + x^3 + x^2 + 1$.

Степінь полінома Вигляд полінома

1	$1+x$
2	x^2+x+1
3	x^3+x+1
	x^3+x^2+1
4	x^4+x+1
	x^4+x^3+1
	$x^4+x^3+x^2+x+1$
5	x^5+x^2+1
	x^5+x^3+1
	$x^5+x^3+x^2+x+1$
	$x^5+x^4+x^2+x+1$
	$x^5+x^4+x^3+x+1$
	$x^5+x^4+x^3+x^2+1$

10.4. Коды Боуза – Чоудхурі – Хоквінгема

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

$$P(x) = HCK \{a_1(x)a_3(x)...a_i(x)...a_{d_0-2}(x)\} \quad (10.20)$$

Мінімальні поліноми для $m \leq 7$ наведені в табл. 8.12, причому всі вони є незвідними.

Та

блиця 8.12

i	Мінімальні поліноми при m , яке дорівнює					
	2	3	4	5	6	7

1	x^2+x+	x^3+x+	x^4+x+1	x^5+x^2+1	x^6+x+1	x^7+x^3+1
3	1	1	$x^4+x^3+x^2+x+$	$x^5+x^4+x^3+x^2$	$x^6+x^4+x^2+x+$	$x^7+x^3+x^2+x+$
5	-	-	1	+1	1	1
7	-	-	-	$x^5+x^4+x^2+x+$	$x^6+x^5+x^2+x+$	$x^5+x^4+x^3+x^2$
	-	-	-	1	1	+1
				-	x^6+x^3+1	$x^7+x^6+x^5+x^4$
						+ x^2+x+1

Нехай необхідно побудувати код Боуза-Чоудхурі-Хоквінгема з $n=31$ і $d_0=5$. Маємо

$$P(x) = HCK \{a_1(x)a_3(x)\} = a_1(x)a_3(x)$$

Оскільки в даному випадку $m=5$, то значення мінімальних поліномів випишемо з 5-го стовпця табл. 8.12:

$$a_1(x) = x^5 + x^2 + 1, \quad a_3(x) = x^5 + x^4 + x^3 + x^2 + 1,$$

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

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 коді 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 вводиться додаткова перевірка на парність, що охоплює всі інформаційні розряди аналогічно тому, як це було зроблено в циклічних і нециклічних кодах Хемінгаз $d_0=4$.

Циклічні коди поширені в системах передавання даних і використовуються у різних за призначенням апаратурі передавання даних (АПД). Незважаючи на те, що умови застосування кожної конкретної АПД різні, а використовувані канали можуть сильно відрізнятися за своїми характеристика-

ми, є тенденція стандартизації методів підвищення надійності приймання за допомогою коригувальних 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 виду $x^{16} + x^{12} + x^5 + 1$. 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. *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. З матричної точки зору укорочення на 1 розряд полягає у викреслюванні з відповідної матриці одного верхнього рядка і одного лівого стовпця. 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	k	r
3	1	2
7	4	3
15	11	4
31	26	5
63	57	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, наприклад, x^4+x+1 (див. стор. 393). Твірна матриця циклічного (15,11)-коду має вигляд

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

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

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

Коригувальна здатність укороченого циклічного коду не нижче коригувальної здатності 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.

10.5. 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 3 одиниці і 4 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 одна або кілька одиниць у кодовій комбінації переходять у нулі, а така сама кількість нулів - в одиниці. У цілком асиметричних каналах (тобто каналах, у яких переходять 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, яка працює на короткохвильових радіоканалах, які є асиметричними.

Істотним недоліком кодів з постійною вагою є те, що вони належать до числа нерозділних кодів, у комбінаціях яких неможливо виділити інформаційні і перевірні розряди. Це значно затрудняє 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: інформаційній частині кодової комбінації підраховується число одиниць, після чого формується перевірні розряди, що являють собою запис цього числа в двійковій формі. Так само формується 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

найважливішим достоїнством 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. Коригувальна здатність кодів з інвертованими перевірками цілком збігається з коригувальною 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 сприяє підвищенню стійкості 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.

10.6. 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 як нм-небудь кодом, а потім (не обов'язково тим самим кодом) - кожен стовпець. Застосуємо як перший код код з парним числом одиниць, а як 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, у яких кожен інформаційний розряд входить у комбінації трьох, чотирьох і т. д. ітеративних кодів, однак багатовимірні ітеративні коди менш поширені.

Властивості ітеративного коду визначаються параметрами ітеративних кодів, залежно від яких ітеративний 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:

$$\left. \begin{aligned} n &= \prod_{i=1}^S n_i \\ k &= \prod_{i=1}^S k_i \\ d_0 &= \prod_{i=1}^S d_{0i} \end{aligned} \right\} \quad (10.21)$$

де n_i, k_i, d_{0i} – параметри ітеративних кодів; S - кратність ітерування.

Отже, найпростіший ітеративний код, утворений перевіркою на парність (ненарність) 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. 10.2).

Найпростіший ітеративний код має досить високу виявляючу здібність - 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 Хемінга. 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 Хемінга - порядку 10^{-8} - 10^{-10} .

Істотним недоліком ітеративних кодів, які використовують для перевірок за рядками і стовпцями 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).

10.7. Ефективність систематичних коригувальних кодів.

Розглянемо ефективність використання систематичних коригувальних кодів, що виявляють помилки. Найважливішою характеристикою таких кодів є коефіцієнт підвищення вірності

$$K_{п.в.} = \frac{P_{ПОМ}}{P_{н.ПОМ.}} \quad (10.22)$$

де $P_{ПОМ}$ - частота появи помилкових кодових комбінацій у дискретному каналі (тобто перед пристроєм захисту від помилок); $P_{н.ПОМ.}$ - частота появи кодових комбінацій з невиявленими помилками в каналі передавання даних (тобто після пристрою захисту від помилок).

Характерною рисою використання коригувальних 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 прийманим 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 поводження коду в різних каналах - в різних закономірностях розподілу помилок у них.

Відомо, що коригувальний код, що має мінімальну кодову відстань d_0 , виявляє всі сполучення помилок кратністю до $d_0 - 1$ включно. Тому ймовірність виявлення помилки кодом можна визначити як

$$P_{H.пом} = \frac{B(d_0)}{C_n^{d_0}} P(d_0, n) + \frac{B(d_0 + 1)}{C_n^{d_0+1}} P(d_0 + 1, n) + \dots = \sum_{i=d_0}^n \frac{B(i)}{C_n^i} P(i, n) \quad (10.23)$$

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/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 створенні систем передавання даних завжди відбирають "сильні" коди, які щонайкраще відповідають розподілу помилок у дискретних каналах, і найбільші значення коефіцієнта, що забезпечують підвищення надійності. 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. Трудність задачі визначається дуже складним 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.

10.8. Мажоритарне декодування

В основі мажоритарного способу виправлення помилок - визначення, який розряд кодової комбінації, то 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 перевірної матриці будь-якого систематичного коду

$$H = \begin{vmatrix} h_{11} & h_{12} & h_{13} & \dots & h_{1n} \\ h_{21} & h_{22} & h_{23} & \dots & h_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ h_{r1} & h_{r2} & h_{r3} & \dots & h_{rn} \end{vmatrix}$$

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 перевірної матриці також утворюють перевірні співвідношення. Виконавши μ лінійних операцій над рядками перевірної матриці, можна (для деяких кодів) одержати нову матрицю

$$L = \begin{vmatrix} l_{11} & l_{12} & l_{13} & \dots & l_{1n} \\ l_{21} & l_{22} & l_{23} & \dots & l_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ l_{\mu 1} & l_{\mu 2} & l_{\mu 3} & \dots & l_{\mu n} \end{vmatrix}$$

що характеризується двома важливими властивостями:

- один зі 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 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 , вибирають за більшістю (тобто мажоритарне) 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 колова 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 $P(x) = x^8 + x^7 + x^6 + x + 1 \approx 111010001$. Мінімальна кодова відстань $d_0 = 5$ і код може виправити дворазову помилку. Обчислимо перевірний поліном:

$$h(x) = \frac{x^n + 1}{p^{-1}(x)} = x^7 + x^3 + x + 1 \approx 10001011.$$

$$\frac{B(i)}{C_n^i} \approx 2^{-r} \quad (10.24)$$

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, що 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,

$$P_{H.ПOM.} \approx 2^{-r} \sum_{i=d_0}^n P(i, n) = 2^{-r} P(\geq d_0, n) \quad (10.25)$$

де $P(\geq d_0, n)$ - імовірність того, що в кодовій комбінації завдовжки n буде помилок більше ніж d_0 . Нагадаємо, що ймовірність появи в кодовій комбінації $P_{ПOM}$ помилок записується як $P_{ПOM} = P(\geq 1, n)$ Тоді коефіцієнт підвищення вірності коду можна визначати як

$$K_{П.Н.} = \frac{P_{П.ОМ}}{P_{Н.П.ОМ}} = \frac{P(\geq 1, n)}{2^{-r} P(\geq d_0, n)} = 2^r \frac{P(\geq 1, n)}{P(\geq d_0, n)} \quad (10.26)$$

Відношення $P(\geq 1, n)/P(\geq d_0, n)$ завжди більше 1. Тому коефіцієнт підвищення вірності коду $K_{П.Н.}$ майже завжди перевищує 2^r . Для більшості систематичних кодів величина 2^r є досить надійною нижньою оцінкою коефіцієнта підвищення вірності $K_{П.В.}$.

Співвідношення (10.26) 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, що 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 менше 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 перешкоди 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

$$P_{Н.П.ОМ} = \frac{1}{2^{r-1}} P(r+1, n) + \frac{1}{2^r} P(\geq r+2, n) \approx \frac{1}{2^r} P(> r, n) \quad (10.27)$$

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 $P(> 1, n)$ дорівнює, і в цьому легко переконатися, ймовірності появи в кодовій комбінації помилки. Тому коефіцієнт підвищення вірності при використанні циклічного коду

$$K_{П.В.} \approx 2^r \frac{P(\geq 1, n)}{P(> r, n)} \quad (10.28)$$

Величина 2^r для більшості випадків є досить надійною нижньою оцінкою коефіцієнта підвищення вірності циклічних кодів.

Отже, при використанні систематичних кодів можна вважати, що ймовірність невиявлення кодом помилкового блоку в 2^r рази менше, 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 надійності, яка значно перевищує 2^r . Є також "слабкі" коди, коефіцієнт підвищення надійності яких менше ніж 2^r .

Побудуємо перевірну матрицю. При цьому як перший рядок використаємо перевірний поліном, помножений на x^{r-1} , а інші рядки одержимий циклічним зсувом першого:

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

Перетворимо перевірну матрицю так. Додамо 1, 5, 7, і 8-й рядки матриці:

$$1 \oplus 5 \oplus 7 \oplus 8 = 1000000010 \ 00101$$

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 рядки перевірної матриці H :

$$L = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

Як бачимо, у цій матриці один зі 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 перевірками:

$$a_0 = a_1 \oplus a_3 \oplus a_7;$$

$$a_0 = a_2 \oplus a_6 \oplus a_{14};$$

$$a_0 = a_4 \oplus a_{12} \oplus a_{13};$$

$$a_0 = a_8 \oplus a_9 \oplus a_{11}.$$

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.

Мі 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:

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

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.

Схема мажоритарного декодера циклічного $(15, 7)$ -коду показана на рис. 10.3. Декодер складається з n -розрядного регістра зсуву, набору дво-вхідних 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 відповідно до системи поділених 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 тактів: протягом перших тактів заповнюється регістр зсуву, а протягом наступних визначається кожний розряд.

10.9. 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 -ійних 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 -ійних кодів можуть бути визначені всі поняття і доведені усі властивості, отримані раніше для двійкових кодів, а саме - породжуюча і перевірюча матриці, систематичні коди, дуальні коди, границі мінімальних відстаней, стандартні розташування,

синдроми, вагові спектри і границі для $p_{\text{но}}$ і $p_{\text{од}}$, циклічні і БЧХ-коди, алгебраїчні і мажоритарні алгоритми декодування. Найбільш важливим класом m -їйних кодів є коди Ріда-Соломона (скорочено РС-коди). Вони можуть бути побудовані як систематичні циклічні (n, k) -коди при $n = q - 1, n - k = 2t, t$ - число помилок, що виправляються. Коди РС є частиною стандарту цифрового запису на компакт-диски.

По визначенню вектор x є словом m -їйного (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. Породжуючий многочлен коду РС має вигляд

$$g(D) = (D - \alpha)(D - \alpha^2) \dots (D - \alpha^{n-k}) \quad (10.29)$$

Як бачимо з визначення РС-кода, він є частковим випадком m -їйних БЧХ-кодів, і відповідно до доведеного раніше, мінімальна кодова відстань таких кодів буде в точності дорівнювати $d = n - k + 1$.

Легко показати, що ніякий лінійний систематичний m -їйний ($m \geq 2$) код не може мати $d > n - k + 1$. Дійсно, якщо вибрати значення k мінус одного інформаційного символу рівного нулю, то це дасть ненульове кодове слово вагою не більш ніж $n - k + 1$, що по властивості лінійного коду і визначає верхню границю для d як $n - k + 1$. Оскільки РС-код реалізує верхню границю для мінімальної кодової відстані, то він виявляється оптимальним серед усіх m -їйних (n, k) -кодів, в розумінні виправлення і виявлення помилок гарантованої кратності.

Можна дати простий опис РС-кода в несистематичному представленні. Тоді кодовий вектор визначається як

$$x = [F(1), F(\alpha), F(\alpha^2), \dots, F(\alpha^{q-2})] \quad (10.30)$$

де $F(D) = b_0 + b_1 D + \dots + b_{k-1} D^{k-1}, (b_0, b_1, \dots, b_{k-1})$ - значення інформаційних m -їйних символів.

Вибір довжини коду $n = q - 1$ є досить сильним обмеженням, особливо при великому порядку поля. Тому можна будувати так звані *скорочені* коди РС, що мають довільну довжину $n \leq q - 1$. Їх можна одержати з повних РС кодів, що мають довжину $n = q - 1$, якщо покласти частину інформаційних символів рівними нулю і викинути їх з кодових блоків. Легко бачити, що укорочення коду не може зменшити мінімальної кодової відстані, і тому (n, k) -код при $n \leq q - 1$ буде як і раніше мати $d = n - k + 1$.

Коди РС, будучи частковим випадком БЧХ кодів, мають алгебраїчний алгоритм виправлення помилок з поліноміальною складністю. Коди m можуть бути використані разом із двійковими кодами для побудови так званих *каскадних кодів*.

10.9.1. 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 буде, наприклад, ітеративний код із двох лінійних систематичних кодів (n_1, k_1) і (n_2, k_2) - див. табл. 7.4. Спочатку повідомлення кодується кодом 1-го ступеню (n_1, k_1) . Нехай k_2 блоків коду 1-го ступеню записані у вигляді рядків матриці. Її стовпці містять по k_2 символів, що будемо вважати інформаційними для коду 2-го ступеню (n_2, k_2) і підпишемо до них $n_2 - k_2$ перевірочних 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 і т.д.

При декодуванні кожного блоку 1-го ступеню виявляють і виправляють помилки. Після того як прийнятий весь двовимірний блок, знову виправляють помилки і витирання, але вже по стовпцях, кодом 2-го ступеню, причому доводиться виправляти тільки ті помилки, що не були виправлені (чи були виправлені невірно) кодом 1-го ступеню. Легко переконатися, 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. 7.6) будується таким способом: 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 $n_2 - k_2$ перевірочних символів m -йного коду (кожен перевірочний 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 $n_1 - k_1$ перевірочних 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 -йного коду (n_2, k_2) , причому виправляються помилки і стирання, що залишилися після декодування внутрішнього коду. Внутрішній код звичайно розрахований на виправлення одиночних помилок, зовнішній на виправлення груп помилок (які є одиночними помилками в збільшених m -йних символах, (рис. 10.4). Як зовнішній код звичайно

використовується m -ийний код Ріда-Соломона (див. вище), що забезпечує найбільше можливе d_2 при заданих значеннях n_2 і k_2 , якщо $n_2 < m$.

Побудований каскадний код еквівалентний лінійному двійковому коду з мінімальною відстанню $d \geq d_1 d_2$. Фактично розглянутий вище алгоритм декодування каскадного коду виявляється більш ефективним і простим, ніж алгоритм декодування двійкового коду з еквівалентною мінімальною відстанню d .

Використання каскадних кодів у 2СК без пам'яті дозволяє забезпечити експонентне спадання імовірності помилкового декодування при збільшенні повної

довжини блоку $n = n_1 n_2$, при цьому швидкість передачі $R = \frac{k_1 k_2}{n_1 n_2}$ може бути

зроблена як завгодно близькою до пропускну здатності каналу, а складність декодування поліноміально залежить від повної довжини блоку n . Ще більш ефективним виявляється застосування каскадних кодів у каналах з пам'яттю. Процес нарощування ступеней каскадного коду можна продовжити. Каскадні коди в багатьох випадках найбільш перспективні серед відомих блокових завадостійких кодів.

10.9.2. Кодування в каналах з пам'яттю

Таблиця 10.2

	Інформаційні символи 1-го ступеню	Перевірочні символи 1-го ступеню
Інформаційні символи 2-го ступеню	$b_{1,1}, b_{1,2}, \dots, b_{1,k_1}$ $b_{2,1}, b_{2,2}, \dots, b_{2,k_2}$ $b_{k_2,1}, b_{k_2,2}, \dots, b_{k_2,k_1}$	$b_{1,k_1+1}, b_{1,k_1+2}, \dots, b_{1,n_1}$ $b_{2,k_1+1}, b_{2,k_1+2}, \dots, b_{2,n_1}$ $b_{k_2,k_1+1}, b_{k_2,k_1+2}, \dots, b_{k_2,n_1}$
Перевірочні символи 2-го ступеню	$b_{k_2+1,1}, b_{k_2+1,2}, \dots, b_{k_2+1,k_1}$ $b_{n_2,1}, b_{n_2,2}, \dots, b_{n_2,k_1}$	$b_{k_2+1}, b_{k_2+1,k_1+1}, \dots, b_{k_2+1,n_1}$ $b_{n_2,k_1+1}, b_{n_2,k_1+2}, \dots, b_{n_2,n_1}$

Як було відзначено в §5.7, пам'ять 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 імовірність помилкового декодування для каналів з пам'ятю. Більш того, навіть найкращі коди для каналів з пам'ятю можуть виявитися значно гіршими, ніж посередні коди в каналах без пам'яті.

Реальні фізичні канали зв'язку володіють, як правило, пам'ятю, що може пояснюватися кореляцією випадкових перешкод і параметрів каналів зв'язку. Типовими прикладами каналів з пам'ятю є канали з завмираннями. Універсальним методом зведення каналів з пам'ятю до каналів без пам'яті є, як відзначено в §5.7, перемежінню символів на передачі і їхньому деперемежінню на прийомі. Після перемежіння символів можна використовувати алгоритми виправлення помилок для каналів без пам'яті. Платою за дану перевагу є затримки декодування, додаткова пам'ять, а також згадане раніше зменшення пропускну здатності каналу зв'язку. Остання обставина може бути "пом'якшена" переходом до алгоритму кодування-декодування з перемежінням і оцінкою імовірності помилок. У цьому випадку пам'ять каналу використовується для оцінки імовірностей помилок символів у кодових блоках, що потім враховуються при виправленні помилок.

Якщо пам'ять каналу виявляється в появі чітко виражених *пачок* помилок, тобто групи підряд слідуєчих одне за одним помилкових символів, то можна використовувати спеціальні коди, орієнтовані на виправлення помилок саме такої конфігурації, а не незалежних помилок (див. більш докладно коди Файра). Ефективними виявляються в каналах з пам'ятю і 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'ятю методу адаптивного кодування (декодування), коли при відсутності перемежіння символів на передачі декодер у режимі виправлення пачок помилок запускається лише тоді, коли зафіксоване групування помилок у каналі. В інший час декодер працює в режимі виправлення одиночних помилок. Метод адаптивного кодування (декодування) істотно зменшує затримку в ухваленні рішення в порівнянні з методом перемежіння символів на передачі і їх деперемежіння на прийомі.

10.10. Системи зі зворотним зв'язком

Дотепер розглядалися системи, у яких передача інформації здійснювалася в одному напрямку: від передавача до приймача. 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 передачі не тільки звичайної інформації, але також і спеціальних повідомлень, призначених для підвищення завадостійкості прямого каналу. Системи зв'язку, у 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 прийомному кінці. У випадку виявлення помилок по зворотному каналу передається сигнал запиту для повторення перекрученого повідомлення. Якщо на передавальному кінці -сигнал запиту прийнятий правильно, то поточна передача переривається й автоматично повторюється помилково прийняте повідомлення. Така система зветься *системою з керуючим зворотнім зв'язком* або *системою з автозапитом* (перезапитом).

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 перекрученого повідомлення. Ця система називається *системою з інформаційним зворотнім зв'язком* чи *системою з порівнянням*. Зупинимося спочатку більш докладно на системах з автозапитом. Одним з методів виявлення .перекручувань у прийнятому повідомленні є застосування коригувальних кодів, що виявляють помилки. Визначимо імовірність помилкових комбінацій у системі з автозапитом і виявляючим кодом. При прийомі кодової комбінації можливі три несумісні події, сумарна імовірність яких дорівнює одиниці:

$$P_{PP} + P_{HO} + P_{OO} = 1$$

де P_{PP} — імовірність правильного прийому, P_{HO} — імовірність помилок, що не виявляються даним кодом, P_{OO} — імовірність помилок, що виявляються. При відсутності чи невиявленні помилок комбінація надходить на вихід системи. Якщо наявні виявлені помилки, то прийнята комбінація стирається і посилається запит для її повторення. Тому при правильному прийомі сигналу запиту імовірність повторення комбінації дорівнює P_{OO} . У результаті вторинної передачі комбінації знову виникає первісне положення з трьома можливими результатами і т.д. Процес передачі окремої комбінації в такій системі можна представити графічно за допомогою векторної діаграми. На рис. 10.5 векторами показані ймовірні переходи системи з одного стану в інший при прийомі деякої комбінації.

Надалі будемо вважати, що в системі відсутнє яке-небудь обмеження кількості повторних передач однієї і тієї ж комбінації, а переходи системи в різні стани незалежні. При цих умовах неважко визначити імовірність помилкового переходу, що складається з l повторень комбінації з наступною подачею її на вихід при наявності невиявлених помилок. Як видно з рис. 10.5, ця імовірність дорівнює $P_{\infty}^l P_{HO}$. Повну

імовірність помилки, ймовірно, представляє сумарна імовірність усіх помилкових переходів

$$P_{OШ} = P_{HO} + P_{OO}P_{HO} + P_{OO}^2P_{HO} + \dots = P_{HO} \sum_{l=0}^{\infty} P_{OO}^l$$

Обчислюючи по відомій формулі суму геометричної прогресії, одержимо

$$P_{OШ} = \frac{P_{HO}}{1 - P_{OO}} \quad (10.29)$$

При виведенні цього виразу не враховувалася імовірність помилки в сигналі запиту, що в багатьох випадках допустимо, через те, що за такі сигнали вибираються спеціальні кодові послідовності, що володіють високою завадостійкістю.

Якщо $P_{OO} \ll 1$, то $P_{OШ} \approx P_{HO}$ і застосування системи з автозапитом

дасть підвищення завадостійкості, обумовленою величиною P_{HO} , що для різних типів кодів була приведена в попередньому розділі.

У розглянутих системах передача повідомлень відбувається з деякою надмірністю, обумовленою надмірністю самого коду і повторною передачею комбінацій, що еквівалентно збільшенню їхньої тривалості. Визначимо величину цієї надмірності. Звичайно при виявленні помилки приймач переходить у режим очікування запитуваної комбінації, тривалість якого $\tau_{ож}$ не менше сумарної тривалості проходження сигналів в обидва кінці. Тоді при l повтореннях час, затрачуваний на прийом однієї комбінації, дорівнює $\tau_k + l\tau_{ож}$, де τ_k – номінальна тривалість комбінації. Імовірність того, що комбінація надійшла на вихід після l (повторень див. рис. 10.5), визначається добутком

$$P_l = P_{OO}^l (P_{ПР} + P_{HO}) = P_{OO}^l (1 - P_{OO}).$$

Звідси еквівалентна середня тривалість комбінації з урахуванням втрат на повторну передачу може бути представлена сумою

$$\begin{aligned}\bar{\tau}_k &= \sum_{l=0}^{\infty} (\tau_k + l\tau_{ож}) P_l = \sum_{l=0}^{\infty} (\tau_k + l\tau_{ож}) P_{oo}^l (1 - P_{oo}) = \\ &= \tau_k [1 + N(P_{oo} + P_{oo}^2 + P_{oo}^3 + \dots)],\end{aligned}$$

де $N = \frac{\tau_{ож}}{\tau_k}$ - число додатково переданих комбінацій.

Використовуючи формулу суми геометричної прогресії, одержимо

$$\bar{\tau}_k = \tau_k [1 + N(\frac{1}{1 - P_{oo}} - 1)] = \tau_k \frac{1 + (N - 1)P_{oo}}{1 - P_{oo}}$$

Таким чином, повна збитковість у системі з автозапитом при використанні n -значного виявляючого коду, буде дорівнювати (7.6):

$$\chi = 1 - \frac{k}{n\bar{\tau}_k / \tau_k} = 1 - \frac{k(1 - P_{oo})}{n[1 + (N - 1)P_{oo}]}$$

де k - значність первинного коду та $n\frac{\bar{\tau}_k}{\tau_k}$ — еквівалентна значність виявляючого коду,

при наявності втрат за рахунок повторних передач.

Розглянемо тепер характеристики систем з інформаційним зворотнім зв'язком. У найпростішому випадку прийняті приймачем повідомлення повністю передаються по зворотному каналу на передавач, де відбувається порівняння з переданими повідомленнями. При наявності розбіжностей перекручені повідомлення повторюються. Цей різновид інформаційного зворотного зв'язку іноді називають ретрансляційним зворотнім зв'язком. У подібній системі помилка не буде виявлена, якщо вона відбудеться в одному і тому ж символі двічі: при передачі по прямому і зворотному каналах. Якщо імовірності помилок у цих каналах відповідно рівні P_0 і P_0^* , то можна показати, що імовірність невиявленої помилки в одному символі при $P_0 \ll 1$ і $P_0^* \ll 1$ приблизно дорівнює їхньому добутку $P_0 P_0^*$. Тому що P_0^* зазвичай величина досить мала, вииграш у завадостійкості може бути значним.

Перевага ретрансляційного зворотного зв'язку - порівняльна конструктивна простота. Однак для ретрансляції необхідно повністю займати зворотній 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 зворотньому каналу передаються відповідні їм контрольні символи. На передавальному кінці відбувається зіставлення переданих інформаційних і прийнятих контрольних символів. Ті комбінації, у яких порушується відповідність між цими двома групами символів, вважаються помилковими і вони знову повторюються. Підбираючи коди з різною надмірністю і коригувальною здатністю, тут легше, ніж у ретрансляційній системі, забезпечити найкращі умови передачі повідомлень. Аналіз показує, що застосування кодів з надмірністю, рівною 0,5, тобто кодів, контрольні елементи яких також повністю займають зворотній канал, дає можливість одержати більш високу завадостійкість, ніж у ретрансляційній системі.

Якщо імовірності помилок символів у прямому і зворотному каналах однакові, повна імовірність помилки кодової комбінації при інформаційному зворотному зв'язку виражається тією ж формулою, що й у системі з автозапитом (8.20). У відношенні надмірності (з урахуванням прямої і зворотної передач) ці системи також рівноцінні. Однак у тому випадку, коли імовірність помилки в зворотному каналі значно менше, ніж у прямому, системи з інформаційним зворотним зв'язком володіють більш високою завадостійкістю. У межі при нескінченно великому відношенні сигналу

до перешкоди в зворотному каналі ($P_{ОШ}^* = 0$) і використанні кодів з достатньою надмірністю інформаційний зв'язок може забезпечити безпомилкову передачу в прямому каналі. У цьому легко упевнитися на прикладі ретрансляційного зв'язку.

Дійсно, при $P_{ОШ}^* = 0$ всі помилки в прямому каналі будуть виявлені, а отже, і виправлені. Близькі до цього умови мають місце на практиці при здійсненні зв'язку потужної наземної станції зі станцією якого-небудь рухливого об'єкта: літака, ракети і т.д. Використовуючи наземну станцію для передачі сигналів зворотного зв'язку, можна забезпечити високу завадостійкість і в каналі рухливого об'єкта - Земля, де потужність сигналу, як правило, невелика.

На закінчення необхідно відзначити, що доцільність застосування систем зі зворотним зв'язком залежить від рівня перешкод у прямому і зворотному каналах. При дуже сильних перешкодах зростає кількість повторень і пропускна здатність системи різко падає. У цих умовах більш раціональною може виявитися система без зворотного зв'язку.

10.11. Сполучення процедур демодуляції і декодування

Вище були розглянуті двухетапні процедури визначення на прийомі переданого кодового слова. На першому етапі в схемі, яку можна назвати першою вирішальною ІРС, визначався вид переданого елемента (символу) кодової комбінації.

На другому етапі по другій вирішальній схемі 2PC по послідовності $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ визначалося кодове слово. Друга вирішальна схема при двухетапній процедурі винесення рішення називається в літературі також твердим (дискретним) декодером, а найчастіше просто декодером. Двухетапне декодування називають *поелементним* (*посимвольным*).

Дискретний декодер виносить рішення про передане кодове слово тільки по послідовності «0» і «1», не використовуючи інформацію про якість прийому кожного з елементів. Це, як неважко здогадатися, не сприяє одержанню максимальної вірності рішення і, крім того, обмежує можливості декодера. Для пояснення даної тези розглянемо декодування для коду (7, 6), що при поелементному прийомі здатний тільки виявляти помилки. Якщо декодер буде мати інформацію про якість прийому кожного з елементів, то можна і виправляти одиночні помилки. Якість (надійність) прийому i -го елемента визначається умовною імовірністю його неправильного прийому $p_i(H/Y)$. Чим більше $p_i(H/Y)$, тим гірше 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 місце розташування помилки відповідає місцеві розташування найменш надійного символу. Змінюючи його на протилежний, зробимо виправлення одиночної помилки. Вище був розглянутий один з варіантів так званого м'якого, чи аналогового, декодування, при якому використовується інформація про надійність прийнятих символів. Чим повніше використовується декодером інформація про надійність символів, тим менше імовірність неправильного декодування. Найбільш повне використання інформації про надійність прийнятих символів відповідає прийому в цілому. Свою назву такий метод прийому одержав унаслідок розгляду тракту прийому як єдиної вирішальної схеми, що здійснює винесення рішення про передане кодове слово на основі 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 в цілому в умовах гауссових перешкод і дискретного каналу без пам'яті можна, використовуючи схему, аналогічну схемі оптимального приймача. Ця схема дозволяє здійснити поділ двох сигналів, що відповідають символам «0» і «1». У нашому випадку число сигналів, що потрібно розділити, дорівнює 2^k , де k - число інформаційних елементів у кодовому слові. Якщо взяти $k=20$, то буде потрібно розділити 1048576 сигналів і, отже, таке ж число гілок прийому. Реалізувати такий приймач, навіть на сучасній елементній базі, неможливо.

Застосування прийому в цілому в порівнянні з поелементним прийомом по гауссовому каналу без пам'яті дає енергетичний виграш приблизно в 3 дБ при

фіксованій імовірності помилкового декодування. Для каналів з змінними параметрами і пам'яттю цей вигравш буде значно більше і складає вже 10...20 дб.

Складність прийому в цілому змушує шукати такі двухетапні процедури винесення рішення про передане кодове слово, при яких буде забезпечена більш проста реалізація тракту прийому, ніж при прийомі в цілому, і в той же час буде використана інформація про надійність прийнятих елементів, тобто вирішена тією чи іншою мірою задача сполучення процедур демодуляції і декодування. У залежності від ступеня використання інформації про надійність елементів буде забезпечене більше чи менше наближення по завадостійкості до прийому в цілому.

Розрізняють два підходи до проблеми м'якого декодування. При першому мінімізується імовірність неправильного декодування кодового слова, при другому мінімізується середня імовірність 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 імовірності помилки на елемент веде до мінімізації імовірності помилки в послідовності елементів, і навпаки.

Найбільш часто на практиці вирішується задача мінімізації імовірності неправильного декодування. При цьому враховуються обмеження на складність реалізації прийомного тракту. В усіх цих випадках робиться спроба породити відносно невелике число кодових слів, серед яких з великою імовірністю міститься передане кодове слово, що знаходиться на найменшій відстані від прийнятої послідовності елементів (символів). Різні методи мінімізації імовірності неправильного декодування розрізняються способом породження цих кодових слів. Нижче розглядаються деякі з них.

Прийом по найбільш надійним символам. Ідея методу декодування по найбільш надійним символам (елементам) запропонована Л. Ф. Бородіним і може бути пояснена в такий спосіб. Визначимо вид двійкових символів $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ і їхню надійність $\gamma_i, i = \overline{1, N}$. Нехай найменш надійні символи розташовані на перевіірочних позиціях. Зітремо ці символи. Тоді, якщо символи, що залишилися прийняті без помилок, 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 по який може бути винесене рішення про передане кодове слово, визначається нерівністю $k \leq n_1 \leq n - (d_0 - 1)$. На пристрій, що кодує, надходить послідовність

одиниць і нулів, а також інформація про надійність кожного із символів прийнятої $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ послідовності.

Декодування починається з виділенням k найбільш надійних символів і перевірки можливості однозначного декодування по цих символах. Якщо однозначне декодування неможливе, то додається ще один найбільш надійний із залишившихся $(n-k)$ символів і всі операції повторюються. При прийомі по методу 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 гауссовських каналів метод Бородіна забезпечує в порівнянні з поелементним прийомом виграш приблизно в 1,4 дБ, а для коду з $k = n - 1$ забезпечує таку ж 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. У результаті цього одержуємо вектор $\hat{A} = \{\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n\}$, що складається з 0 і 1.

2. На прийняту кодову комбінацію A накладаються сформовані певним чином тестові послідовності T . Отримані послідовності $\hat{A}^* = \hat{A} \oplus T$ декодуються твердим (дискретним) декодером по раніше розглянутих у попередніх параграфах правилам (див., наприклад, § 6.1). У результаті одержуємо кодові слова $\hat{A}'_1, \hat{A}'_2, \dots, \hat{A}'_j, \dots, \hat{A}'_N$, де N - число тестових послідовностей.

3. З урахуванням вірності (надійності) кожного з елементів визначаємо відстань послідовності $\hat{A}$ до кожного з кодових слів $\hat{A}'_i$, де $i=1, 2, \dots, N$. Прийнятим вважається те кодове слово, відстань до якого буде мінімальним. Відстань L_j до j -го кодового слова можна обчислити по формулі

$$L_j = \sum_{i=1}^n \gamma_i (\hat{a}_i \oplus \hat{a}'_{ij}); j = \overline{1, N} \quad (10.30)$$

де γ_i - позитивне число, що характеризує ступінь надійності прийнятого i -го елемента. Наприклад, при вісьмирівневому 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 .

Схема, що реалізує декодування по Чейзу, представлена на рис. 7.10, де подвійними лініями зображені векторні зв'язки. На вхід схеми з виходу безупинного каналу подається сигнал $V(t)$, що демодулюється; рівень сигналу на виході демодулятора визначає надійність прийнятих елементів. Вид елемента визначається в першій вирішальній схемі ІРС.

Три запропонованих Чейзом алгоритми відрізняються методом формування тестових послідовностей **Т**. Алгоритм 1 вимагає формування найбільшого числа послідовностей і забезпечує найкращу завадостійкість. В алгоритмах 2 і 3 формується менше число тестових послідовностей і відповідно забезпечується гірша завадостійкість.

Для алгоритму 1 потрібно взяти як тестові послідовності нульовий вектор і всі комбінації ваги $[d_0/2]$, де знак $[.]$ означає найбільше ціле число, що не перевершує задане. Число таких тестових комбінацій $C_n^{[d_0/2]}$. Для алгоритму 2 використовується $2^{[d_0/2]}$ тестових послідовностей, 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 ($i=0, 2, 4, \dots, d_0-1$ для непарного d_0 ; $i=0, 1, 3, \dots, d_0-1$ для парного d_0). Загальне число можливих тестових послідовностей $[d_0/2+1]$.

Приклад 10.1. Розрахувати 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 $N = [d_0/2+1] = 2+1 = 3$. Таким чином, найменше число тестових послідовностей потрібно при використанні алгоритму 3.

Розглянемо на прикладі процедуру декодування по Чейзу для алгоритму 3.

Приклад 10.2. Код (6,3) заданий виробляючою матрицею (7.3). Передано кодову комбінацію 000 000. У результаті дії перешкод прийнята кодова комбінація з помилками на позиціях 2 і 3 $A=011000$. Тому що $d_0=3$, те необхідно сформулювати дві тестові послідовності, перша з яких 000 000, а друга містить $d_0 - 2 = 2$ одиниці на позиціях найменш надійних символів. Нехай надійність символів визначається числами 732656, тобто найменш надійними є символи, що коштують на позиціях 2 і 3. Тоді друга тестова послідовність має вид 011000, і в результаті додавання прийнятої послідовності з тестовими маємо

$$\begin{array}{r} 011000 \\ \oplus \\ 000000 \\ \hline 011000 = \hat{A}_1^* \end{array} \qquad \begin{array}{r} 011000 \\ \oplus \\ 011\ 00 \\ \hline 000000 = \hat{A}_2^* \end{array}$$

Після декодування твердим декодером одержимо кодові слова $A'_1=011100$ і $A'_2=000000$. Обчислюючи L - по формулі (7.32) 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 000 000. У розглянутій вище ситуації удалося виправити дворазову помилку.

При розробці систем ПДС доводиться вирішувати не тільки задачі сполучення процедур демодуляції і декодування, але і модуляції і кодування, тому що характеристики дискретного каналу залежать від виду модуляції. Найбільш загальний підхід до рішення 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

$$\frac{B(i)}{C_n^i} \approx 2^{-r} \quad (10.24)$$

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, що 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,

$$P_{H.ПOM.} \approx 2^{-r} \sum_{i=d_0}^n P(i, n) = 2^{-r} P(\geq d_0, n) \quad (10.25)$$

де $P(\geq d_0, n)$ - імовірність того, що в кодовій комбінації завдовжки n буде помилок більше ніж d_0 . Нагадаємо, що ймовірність появи в кодовій комбінації $P_{ПOM}$ помилок записується як $P_{ПOM} = P(\geq 1, n)$. Тоді коефіцієнт підвищення вірності коду можна визначати як

$$K_{П.Н.} = \frac{P_{ПOM}}{P_{H.ПOM}} = \frac{P(\geq 1, n)}{2^{-r} P(\geq d_0, n)} = 2^r \frac{P(\geq 1, n)}{P(\geq d_0, n)} \quad (10.26)$$

Відношення $P(\geq 1, n)/P(\geq d_0, n)$ завжди більше 1. Тому коефіцієнт підвищення вірності коду $K_{П.Н.}$ майже завжди перевищує 2^r . Для більшості систематичних кодів величина 2^r є досить надійною нижньою оцінкою коефіцієнта підвищення вірності $K_{П.Н.}$.

Співвідношення (10.26) 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, що 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 менше 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 перешкоди 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

$$P_{н.пом} = \frac{1}{2^{r-1}} P(r+1, n) + \frac{1}{2^r} P(\geq r+2, n) \approx \frac{1}{2^r} P(> r, n) \quad (10.27)$$

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 $P(> 1, n)$ дорівнює, і в цьому легко переконатися, ймовірності появи в кодовій комбінації помилки. Тому коефіцієнт підвищення вірності при використанні циклічного коду

$$K_{п.в.} \approx 2^r \frac{P(\geq 1, n)}{P(> r, n)} \quad (10.28)$$

Величина 2^r для більшості випадків є досить надійною нижньою оцінкою коефіцієнта підвищення вірності циклічних кодів.

Отже, при використанні систематичних кодів можна вважати, що ймовірність невиявлення кодом помилкового блоку в 2^r рази менше, 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 надійності, яка значно перевищує 2^r . Є також "слабкі" коди, коефіцієнт підвищення надійності яких менше ніж 2^r .

Побудуємо перевірну матрицю. При цьому як перший рядок використаємо перевірний поліном, помножений на x^{r-1} , а інші рядки одержимий циклічним зсувом першого:

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

Перетворимо перевірну матрицю так. Додамо 1, 5, 7, і 8-й рядки матриці:

$$1 \oplus 5 \oplus 7 \oplus 8 = 1000000010 \ 00101$$

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 рядки перевірної матриці H :

$$L = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

Як бачимо, у цій матриці один зі 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 перевірками:

$$a_0 = a_1 \oplus a_3 \oplus a_7;$$

$$a_0 = a_2 \oplus a_6 \oplus a_{14};$$

$$a_0 = a_4 \oplus a_{12} \oplus a_{13};$$

$$a_0 = a_8 \oplus a_9 \oplus a_{11}.$$

Додавши до цих співвідношень тривіальну перевірку $a_0=a_0$, одержимо систему поділених відносно a_0 перевірок:

$$\begin{aligned}a_0 &= a_0; \\a_0 &= a_2 \oplus a_6 \oplus a_{14}; \\a_0 &= a_4 \oplus a_{12} \oplus a_{13}; \\a_0 &= a_8 \oplus a_9 \oplus a_{11}; \\a_0 &= a_1 \oplus a_3 \oplus a_7.\end{aligned}$$

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.

Мі 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:

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

звідки випливає

$$\begin{aligned}a_{14} &= a_0 \oplus a_2 \oplus a_6; \\a_{14} &= a_1 \oplus a_5 \oplus a_{13}; \\a_{14} &= a_3 \oplus a_{11} \oplus a_{12}; \\a_{14} &= a_7 \oplus a_8 \oplus a_{10}.\end{aligned}$$

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.

Схема мажоритарного декодера циклічного (15, 7)-коду показана на рис. 10.3. Декодер складається з n -розрядного регістра зсуву, набору дво-вхідних 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 відповідно до системи поділених 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 тактів: протягом перших тактів заповнюється регістр зсуву, а протягом наступних визначається кожний розряд.

10.12. Згорткові (гратчасті) коди

Вище розглядалися блокові коди, коли значення елементів, що входять у різні блоки, виявлялися незалежними одне від одного.

Для систематичних двійкових блокових (n, k) -кодів послідовність інформаційних символів джерела розбивається на блоки довжиною k біт і потім у кодері до кожного такого блоку додається $r = n - k$ перевірочних символів, після чого блоки довжиною n символів передаються в канал зв'язку. Декодування блоків також здійснюється незалежно одне від одного.

Однак можливий і інший, *безупинний* принцип кодування і декодування, коли на вхід кодера надходить безупинна послідовність інформаційних символів джерела, а з виходу кодера також знімається безупинна послідовність символів, що є функцією вхідних символів і структури кодера. У декодері такого типу на вхід надходить безупинна послідовність символів з каналу зв'язку (можливо, спотворена помилками), а на виході відновлюється (можливо, з помилками, але, як правило, меншими за каналні) послідовність інформаційних символів. Найбільш розповсюдженим класом безупинних кодів є так звані *згорткові* коди, для яких операція формування вихідної послідовності по заданій вхідній послідовності є лінійною. Згорткові коди були уперше відкриті Л. Фінком і П. Елайесом, незабаром Возенкрафт розробив метод послідовного декодування згорткового коду. У 1967 р. з'явився алгоритм Вітербі для оптимального декодування згорткових кодів.

Структура двійкового згорткового коду зі швидкістю $R = \frac{k}{n}$ показана на

рис.10.7,а. Згортковий кодер складається зі зсувного регістра, що включає V комірок пам'яті, блоку суматорів по mod12, входи кожного з яких зв'язані з деякими виходами комірок пам'яті регістра, обумовленими коефіцієнтами $h_{i,j} = (0, 1)$. Виходи суматорів зчитуються за допомогою комутатора K і подаються в канал зв'язку. Таким чином, на кожному такті в регістр зрушення послідовно надходить черговий блок з k

інформаційних символів джерела й одночасно він звільняється від k символів, що містяться в його крайніх правих комітках пам'яті. На цьому ж такті формуються n вихідних символів, що послідовно зчитуються в канал зв'язку. Таким чином, якщо V_i швидкість надходження символів у кодер, то для відсутності зростаючих затримок у часі швидкість передачі символів по каналу зв'язку повинна бути не менше ніж $V_k = \frac{n}{k} V_H$, звідки випливає, що відношення k/n дійсно визначає швидкість згорткового коду. Величина V (чи довжина регістра) звичайно називається довжиною *кодowego обмеження*. У деяких роботах довжина кодowego обмеження визначається числом осередків зсувного регістру мінус 1.

(Можливо й інше, більш загальне представлення згорткового кодера у вигляді схеми з k регістрами зрушення з кодowymi обмеженнями $v_i (i = 1, 2, \dots, k)$. На вхід кожного з них подається один інформаційний символ за час одного такту [13].)

На рис. 7.10 показаний окремий випадок згорткового коду зі швидкістю $1/2$ і довжиною кодowego обмеження $V = 3$. При нульовій інформаційній послідовності вихідна кодова послідовність також дорівнює нулю. У табл. 10.3 приведений приклад формування вихідної послідовності для кодера, показаного на рис. 10.7,б. Вихідна послідовність кодера може бути представлена як цифрова згортка вхідної інформаційної послідовності й імпульсного відгуку кодера (звідси назва кодів – *згорткові*).

Таблиця 10.3

Вхід	0	0	1	0	0	0	1	1	0	0	0
Вихід			1	1	1	0	1	0	0	1	0
д			1	0	1	0	1	1	1	1	0

Згортковий код характеризується наступними параметрами: *відносною швидкістю* коду $R=k/n$ і *надмірністю*, $x = 1-R$ число інформаційних і кодowych символів, що відповідають одному такту роботи кодера (для кодера рис. 10.7,б $R=1/2$); *довжиною кодowego обмеження* V (довжина регістра кодера); *поліномом коду*, що породжує, коефіцієнти яких описують зв'язки суматорів з осередками регістра кодера (для

верхнього $g^{(1)} = 1 + D + D^2$ суматора, для нижнього $g^{(2)} = 1 + D^2$ суматора). Поліноми звичайно записують скорочено, позначаючи кожні три відводи (двійкових коефіцієнтів) як одну вісімкову цифру (код рис. 10.7, б позначається (7,5)).

Крім перерахованих вище параметрів згортковий код характеризується *вільною відстанню* d_{cg} , під якою розуміють відстань по Хемінгу між двома напівнескінченими кодовими послідовностями. Якщо дві однакові інформаційні послідовності кодувати за допомогою кодера, зображеного на рис. 10.7, б, то відповідні їм кодові послідовності будуть збігатися один з одним. Якщо в деякий момент в одній інформаційній послідовності виявиться символ 0, а в інший 1, то з цього моменту кодові послідовності будуть відрізнятися одне від одного незалежно від подальшого змісту інформаційних послідовностей. Мінімальна відстань по Хемінгу між будь-якими двома напівнескінченими кодовими послідовностями з того моменту, як відповідні їм інформаційні послідовності починають розрізнятися, називається *вільною відстанню згорткового коду* d_{cg} .

Вільна відстань d_{cg} характеризує завадозахисні властивості згорткового коду (аналогічно тому як мінімальна відстань d характеризує завадозахисні властивості блокових кодів). Воно показує, яке найменше число помилок повинне відбутися в каналі для того, щоб одна кодова послідовність перейшла в іншу і помилки не були виявлені. Для коду, приведеного в нашому прикладі, вільна відстань $d_{cg} = 5$.

Пошук гарних згорткових кодів (з найбільшим d_{cg} при заданих R і V) звичайно здійснюється методом перебору всіх поліномів, що породжують, на ЕОМ. Таблиці гарних кодів приведені в [13].

Згорткові коди є частковим випадком (лінійною реалізацією) так званих *гратчастих* кодів. Можна також думати, що ґратки є просто іншим (іноді більш зручним) способом представлення і звичайних згорткових кодів.

ґратками називається орієнтований граф з періодично повторюваною структурою "комірок". Кожна комірка містить стовпчики з однакового числа вершин (вузлів), з'єднаних ребрами. Між процедурою кодування згортковим кодом і ґратками мається взаємно однозначна відповідність, що задається наступними правилами [6]:

- 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 виходить два ребра - верхнє для 0 і нижнє для 1);
- над кожним ребром відзначені значення символів, переданих у канал зв'язку, якщо кодер знаходився в стані, що 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 комірок пам'яті (2, 3) у реєстрі зрушення на рис. 10.7, б, ґратка з чотирма станами, що відповідає даному кодеру, буде мати вид, показаний на рис. 10.7, в (ґратка може відображати і нелінійний кодер, коли вихідні символи не є лінійною функцією вхідних). Так само як і блокові коди, згорткові допускають представлення у виді напівнескінченних породжуючих чи перевірючих матриць, однак представлення у виді ґратки виявляється більш зручним для опису алгоритмів декодування.

Згорткові коди мають наступні основні переваги над блоковими при їхньому використанні для виправлення помилок.

1. Вони не вимагають синхронізації по блоках, а лише синхронізації комутаторів *ДО* (на передачі і прийомі).
2. Якщо кодове обмеження V вибрати рівним довжині блокового коду, то виправляюча здатність згорткового коду виявляється більше, ніж виправляюча 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 виграш.
4. Згорткові коди допускають просте об'єднання кодування і модуляції (так називана *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* (AB). Розглянемо його на прикладі м'якого декодування в постійному каналі з адитивним білим гаусовським шумом.

Оскільки прийнятий сигнал на k -му тактовому інтервалі нам відомий, то можна обчислити евклідови (чи гільбертові) відстані між прийнятим сигналом і всіма можливими сигналами:

$$\Delta_{ki} = \int_0^T (z_k(t) - s_k^{(i)}(t))^2 dt, i = 0, 1, \dots, m-1, k = 0, 1, \dots,$$

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 буде тоді відповідати вибору такого шляху на ґратках (тобто послідовності безупинно триваючих ребер), що $\sum_k \Delta_{ki}$ виявиться мінімальною. Здавалося б, для ґратки довжини n (тобто для послідовності переданих символів довжиною n) потрібно для цього перебрати 2^n можливих варіантів, але в дійсності це не так. Ключовий момент АВ полягає в тому, що для кожної вершини на даному кроці (такті) мається безліч метрик, що відповідають з'єднанню 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 простому прикладі. Нехай ґратки мають усього два стани і структуру, показану на рис. 10.8, а, де над ребрами підписані відповідні метрики. Думаємо, що перший інформаційний символ 0. Тоді шляхи, залишені ("виживші") на різних кроках, показані на рис. 10.8, б. Видно, що на 4-му кроці одержуємо виживший шлях, що в умовах наших позначень (орієнтація ребра вниз - 1, вгору - 0) відповідає інформаційній послідовності 0100.

Складність АВ визначається на кожному кроці числом порівнянь метрик, що з'єднують усі вершини, і воно обмежено величиною M^2 , де M – число станів ґрати. Оскільки зі схеми згорткового кодера одержуємо, що $M = 2^{V-1}$, де V – число комірок пам'яті регістра зрушення кодера, то бачимо, що складність АВ експоненціально залежить від довжини кодових обмежень, але лінійно залежить від довжини переданої послідовності. Тому довжина кодових обмежень V при використанні АВ як алгоритм декодування звичайно вибирається не більш 10...15, що, утім, виявляється цілком достатнім для одержання великого енергетичного виграшу. АВ вимагає обробки всієї послідовності сигналів для оптимального декодування навіть першого інформаційного символу. Така процедура вимагає значної пам'яті на прийомі і затримки для декодування елементів повідомлення. Для виключення цих недоліків використовується модифікація АВ у виді *усіченого алгоритму*, коли рішення про інформаційний символ на i -м такті приймається за результатами обробки по АВ

послідовності символів на даному i -м і L наступних тактових інтервалах. Теорія й експеримент показують, що якщо L вибрати порядку декількох довжин кодових обмежень, то енергетичні втрати при використанні такої модифікації виявляться невеликими.

Помітимо, що в дійсності АВ є ефективним методом вирішення значно більш загальної оптимізаційної задачі. Нехай потрібно знайти такий дискретний вектор $X_N = (x_1, \dots, x_N), x_i \in X, |X| = r$, що максимізує (мінімізує) функцію $\Lambda(X_N)$. Тоді прямий метод вимагає перебору r^N різних векторів. Однак, якщо функція $\Lambda(X_N)$ допускає представлення

$$\Lambda(X_N) = \sum_{k=1}^N \lambda_k(\sigma_k)$$

де $\lambda_k(\sigma_k)$ - довільні функції векторних аргументів виду

$$\sigma_k = \begin{cases} (x_{k-v}, x_{k-v+1}, \dots, x_k), & k > v, \\ (x_1, \dots, x_k), & k \leq v, \end{cases}$$

то можна довести, що максимум такої функції знаходиться за допомогою наступного узагальненого АВ (ОАВ):

Крок 1. Знайти $\hat{x}_1(x_2, \dots, x_{v+1}) = \underset{x_1}{\text{Arg max}} \Lambda_{v+1}$ (при x_{v+1}). (Оскільки

$|X| = r$, то для кожного значення аргументу $(x_2, \dots, x_{v+1})$ це потребує не більш ніж r обчислень.)

Крок 2. Знайти $\hat{x}_2(x_3, \dots, x_{v+2}) = \underset{x_2}{\text{Arg max}} \Lambda_{v+2}$ (при $\hat{x}_{v+2}$), де 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 $\hat{x}_S(x_{S+1}, \dots, x_{v+S}) = \underset{x_S}{\text{Arg max}} \Lambda_{v+S}$, 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 $(x_{N-v}, \dots, \hat{x}_N) = \underset{(x_{N-v}, \dots, x_N)}{\text{Arg max}} \Lambda_N$ (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, ОАВ має поліноміальну складність від довжини послідовності N на відміну від експонентної складності при прямому методі перебору всіх альтернативних гіпотез.

Вибираючи різні форми функцій $\lambda_k(\sigma_k)$, можна одержати досить прості методи вирішення оптимізаційних задач, а отже, і оптимальні алгоритми обробки сигналів не тільки для свёрточних кодів, але і для інших моделей каналів, наприклад для каналу з міжсимвольною інтерференцією, описану в § 5.8.

Універсальність АВ полягає в тому, що він може бути використаний для різних розподілів сигналів і перешкод, неоднорідних каналів, для сполучення декодування і демодуляції і не тільки для незалежних розподілів, але і для випадків залежності, описуваних марковськими послідовностями.

У каналах із МСІ алгоритм Кловського-Ніколаєва (АКН) зі зворотним зв'язком за рішенням при фіксованій затримці ухвалення рішення $D=Q$ (Q - пам'ять каналу) практично не поступається по завадостійкості АВ з тією ж затримкою рішення, але реалізується простіше. АКН, отриманий спочатку для субоптимального поліелементного прийому 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 ухваленні рішення L , повинна враховувати не тільки пам'ять каналу (Q), але і довжину кодового обмеження (Кловський Д.Д., Белоус С.А., Карташевський В.Г. Прийом сигналів зі згортковим кодуванням у каналі з міжсимвольною інтерференцією // Проблеми передачі інформації. - Т.7.- Вип. 2.- 1991, с. 97-100).

Згорткові коди можуть декодуватися й іншими алгоритмами (наприклад, послідовного декодування і синдромного декодування), що не є, взагалі кажучи, оптимальними.

Послідовне декодування було уперше уведене Возенкрафтом, однак найбільш широко використовуваний у даний час алгоритм належить Фано [13].

У той час, коли відповідно до алгоритму Вітербі виробляється просування і відновлення метрики для всіх шляхів, що можуть здаватися нам кращими, послідовний декодер істотно обмежує число шляхів, що фактично оновлюються. Основна ідея послідовного декодування полягає в тому, що використовуватися повинний лише той шлях, що має вид найбільш ймовірного. 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 граничне декодування - застосовуються при декодуванні як згорткових, так і блокових кодів. Про табличний метод синдромного декодування ми вже говорили. Граничне декодування - це розроблений Мессі метод, в основі якого лежить спеціальний вибір породжувачих многочленів коду, що допускають рішення перевірочних рівнянь за допомогою мажоритарного вибору символу помилки за деякими оцінками. Останні отримуються підсумовуванням одного чи декількох символів синдрому.

Для оцінки якості м'якого декодування за допомогою АВ скористаємося визначенням: мінімальною евклідовою відстанню d_e згорткового коду при обраному методі модуляції називається мінімальна сума евклідових метрик помилкових шляхів на ґратках, що починаються і закінчуються на правильному шляху.

Тоді для імовірності помилки p_e в першому символі в детермінованому каналі без пам'яті з БГШ при декодуванні по АВ буде справедлива наступна 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 , що починаються і закінчуються на правильному шляху; N_0 спектральна щільність БГШ.

Використовуючи більш повно структуру ґратки згорткового коду, можна оцінити також і середню імовірність 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'ятю, доцільно використовувати каскадні коди з внутрішніми згортковими кодами і м'яким декодуванням по АВ (чи АКН) і зовнішніми РС-кодами з використанням алгебраїчних методів декодування. Така конструкція дозволяє одержати енергетичний вииграш, що досягає 5 дБ при еквівалентній імовірності помилки $P_g = 10^{-5}$ і прийнятної складності декодування. Помітимо, що АВ може бути використаний 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, як уже згадувалося раніше, *кодована модуляція* чи- в іншій термінології - використовуються деякі *сигнально-кодові конструкції* (СКК). Загальну ідею такого методу ілюструє рис. 10.9, а, де індексом i позначається номер такту. У цьому випадку кодер згорткового коду зі швидкістю k/n зручніше представляти у виді рівнобіжного набору k регістрів зрушення з різними довжинами кодових обмежень V_i . На виході такого кодеру, у кожен момент часу з'являється n -мірний двійковий вектор, що відображається в один з 2^n безупинних сигналів, переданих у канал зв'язку. Для каналів з обмеженою смугою типовими є використання сигналів з багаторазовою фазовою чи амплітудно-фазовою модуляцією, причому ключовим моментом тут є *розбивка безлічі* сигналів на *сузір'я*.

Приклад такої розбивки для 8-кової ФМ показаний на рис. 10.9, б. Сузір'я знаходяться в нижньому ряді малюнка. На рис. 10.10 представлена схема кодування Унґербоека, що використовує представлення сигналів у виді сузір'їв.

Тут двійкова послідовність символів джерела розбивається на блоки по k біт, і перші $\tilde{k}$ біт цих блоків подаються на вхід згорткового кодера, а ті, що залишилися надходять на модулятор у некодованому виді. Загальний принцип полягає в тому, що некодовані біти вибирають сигнал у сузір'ї, а кодовані біти визначають вибір сузір'я. (Так, наприклад, для схеми розбивки, показаної на рис. 10.9, б, кодер Унґербоека, зображений на рис. 10.10, буде мати параметри: $k = 2, \tilde{k} = 1, R = 1/2$).

Оскільки, як видно з рис. 10.9, б сигнали в кожному із сузір'їв вилучені 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 і спектральної ефективності (Biglieri E, Divsalar D, Simon M.N. Introduction to trellis – coded Modulation with Applications.- Publ. Mc. Millan.1.- 1991.- 548p.).

При практичному використанні завадостійких кодів головним обмеженням є складність пристрою декодування, що може бути виражена або числом логічних схем у декодері, або числом обчислювальних операцій необхідних для декодування. Тому серед кодів, що забезпечують заданий виграш, варто вибирати ті, котрі допускають менш складну реалізацію, або навпаки, при заданій складності декодування варто вибирати коди, що забезпечують найбільший виграш.

Контрольні питання

1. Дайте визначення коду.
2. Дайте визначення 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 Хемінга?
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 зв'язку зі зворотним зв'язком Ви знаєте?
15. Які переваги і недоліки систем зі зворотним зв'язком?
16. Якими особливостями володіють згорткові коди?
17. Що таке сигнально кодова конструкція?
18. У чому складаються особливості алгоритму Вітербі?
19. Що називається ґратками?

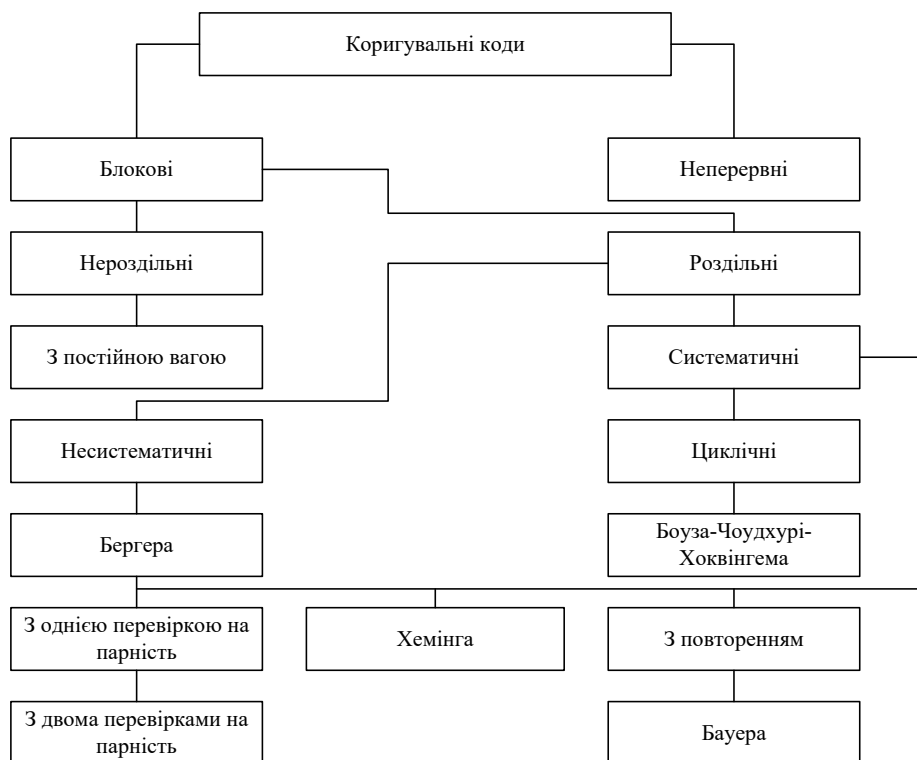


Рис. 10.1. Класифікація коригувальних кодів

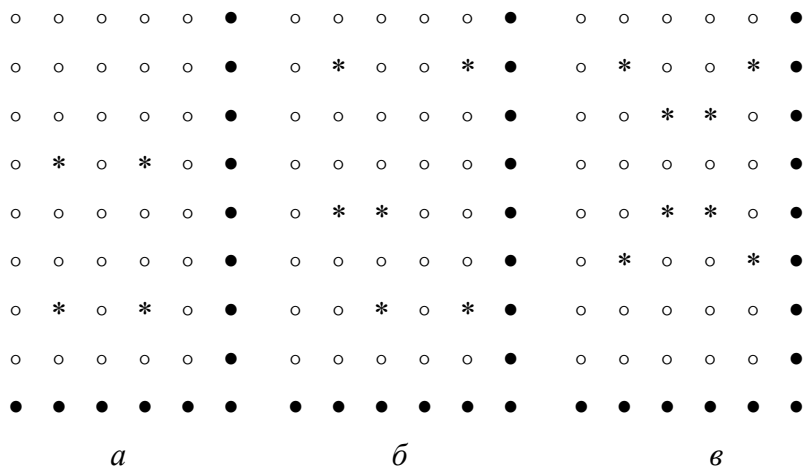


Рис.10.2 Деякі помилки, що не виявляються найпростішим ітераційним кодом:

- - чотириразові помилки; * - шести-, восьмиразові і т.д. помилки

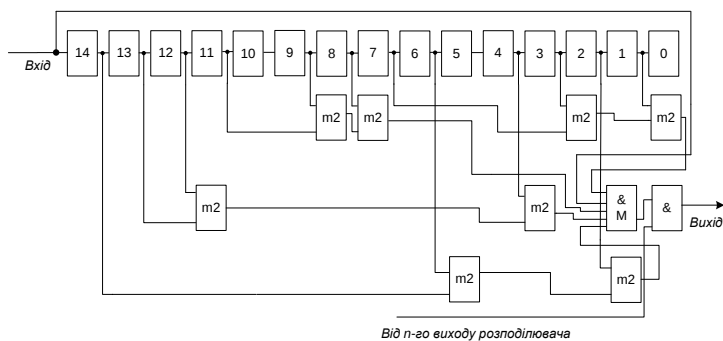


Рис. 10.3. Побудова мажоритарного декодера циклічного (15,7)- коду з подільними перевірками

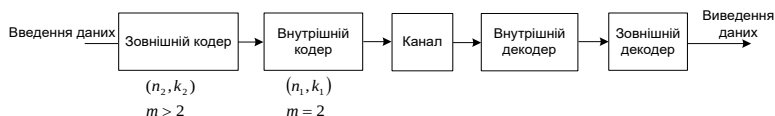


Рис. 10.4. Схема каскадного кодування і декодування

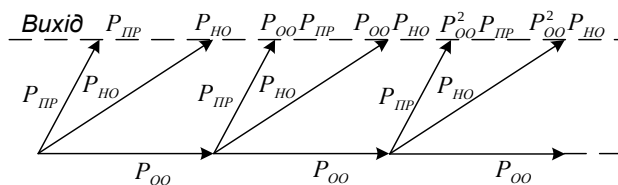


Рис. 10.5. Векторна діаграма станів системи з автозапитом

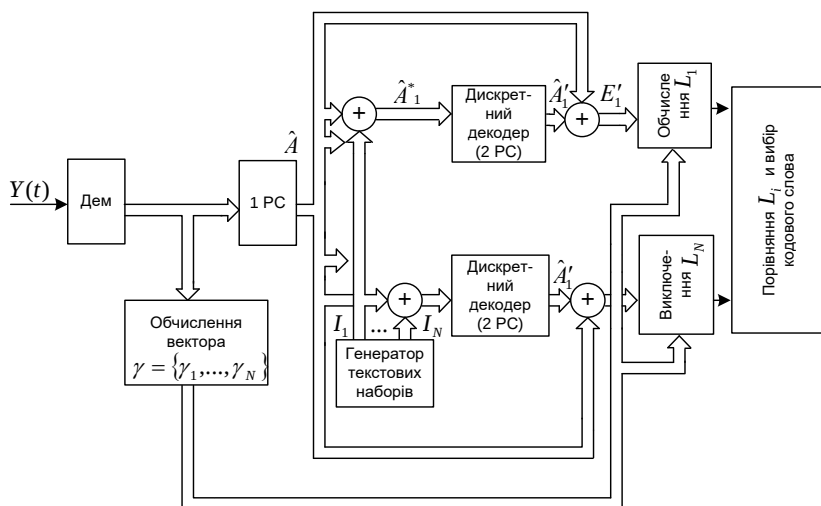


Рис. 10.6. Декодер Чейза

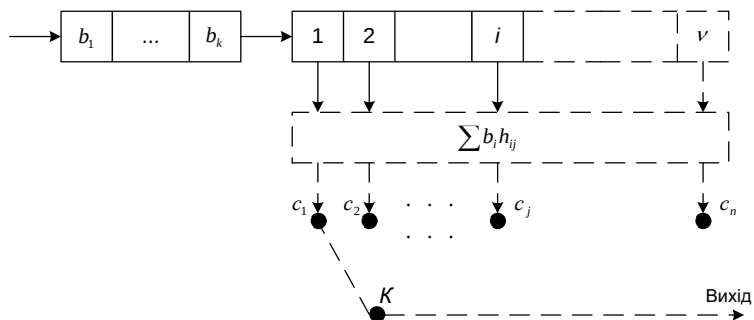


Рис. 10.7,а Структура свёрточного кодера зі швидкістю k/n

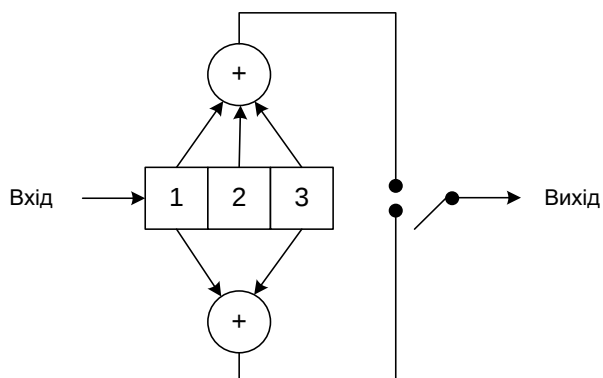


Рис. 10.7, б. Свёрточный кодер зі швидкістю $R=1/2$

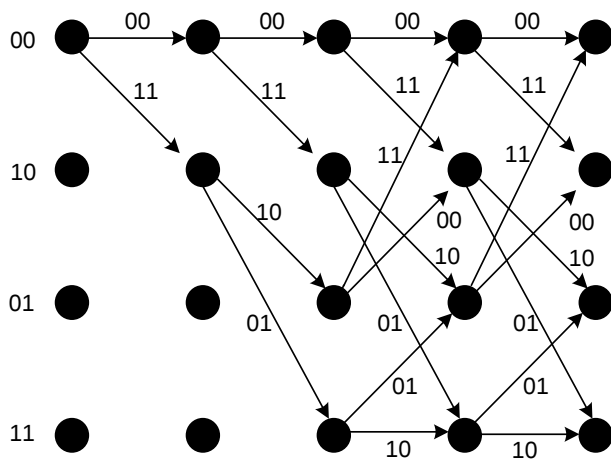


Рис. 10.7, в. Ґратки кодера, показаного на рис. 10.7, б

Рис.10.7. Структура сверточних кодерів зі швидкістю k/n (а) і швидкістю $R=1/2$ (б) і ґраток кодера (в), показаного на рис.10.7, б

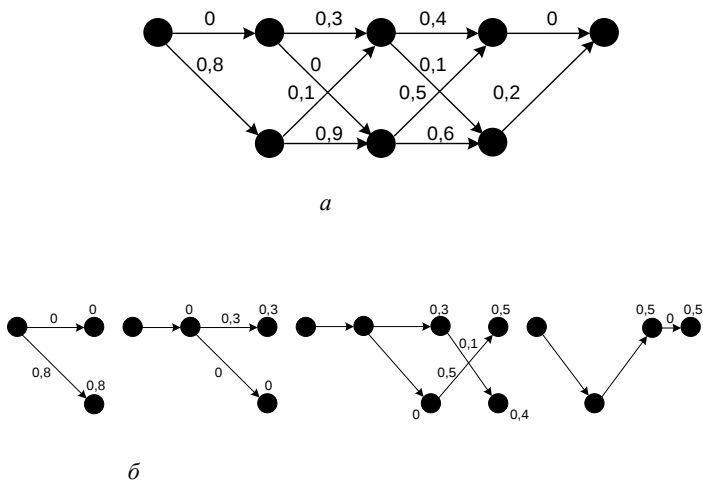
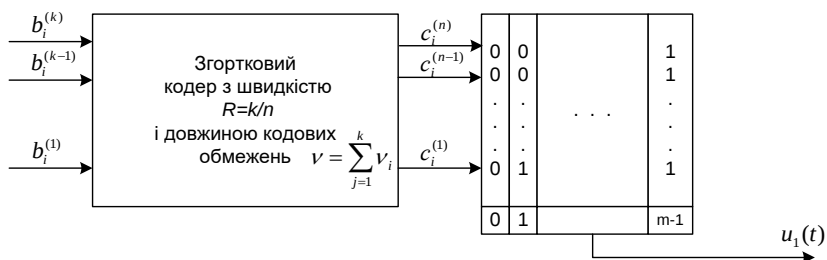
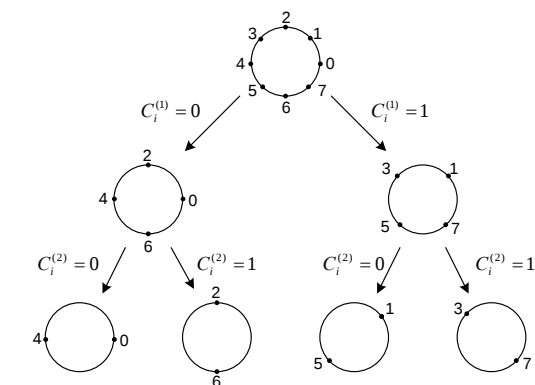


Рис. 10.8. Ґратки з метриками (а) і побудова “вижившого” шляху по АВ.



Відображення двійкових n -послідовностей на безліч сигналів $(u_0, u_1, \dots, u_{m-1})$



б

Рис. 10.9. Система з кодовою модуляцією (а) і розбивка 8-кових ФМ сигналів на сузір'я (б)

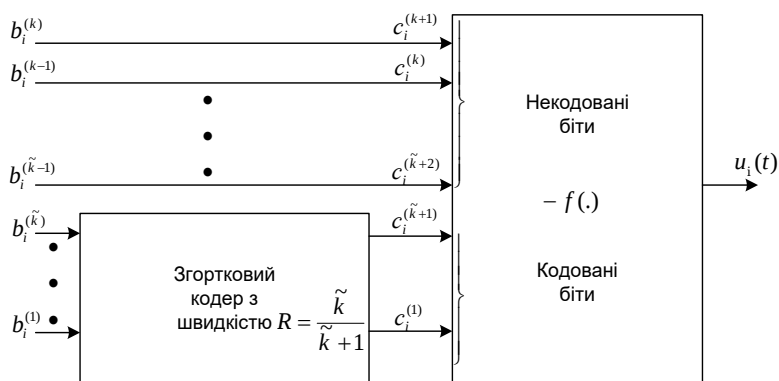


Рис. 10.10. Кодер Унгербоєка

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. **Л.Н. Беркман, Л.О. Комарова, Н.С. Чумак.** Системи електрозв'язку та сигнали. Навчальний посібник. — Київ: ДУІКТ, 2015.
2. **Л.Н. Беркман, Л.О. Комарова, О.І. Чумак.** *Основні поняття та теореми теорії інформації.* Навчальний посібник. — Київ: ДУІКТ, 2015.
3. **Л.О. Комарова, О.І. Чумак.** Основні поняття та теореми теорії інформації. Навчальний посібник. — Київ: ДУІКТ, 2015.
4. **В.П. Куляс.** Теорія інформації та кодування. — Київ: Видавництво "Наукова думка", 2004.
5. **С.І. Харченко.** Основи теорії сигналів та систем. — Харків: ХНУРЕ, 2012.
6. **К. Шеннон.** Математична теорія зв'язку. — Київ: Видавництво "Наукова думка", 1963.
7. **А.І. Зобов.** Цифрові системи передачі інформації. — Київ: 2010.
8. **В.М. Глушков.** Теорія перетворення інформації. — Київ: 1986.
9. **В.В. Андрєєв.** Основи теорії кодування в телекомунікаційних системах. — Київ: ЛНУ, 2011.
10. **М.М. Соколов.** Телекомунікаційні системи: загальні принципи та підходи. — Львів: ЛНУ, 2013.
11. **Л.М. Кравчук.** Основи теорії кодування та захисту інформації. — Київ: ДУІКТ, 2017.

12. Wang, J., Li, X., & Chen, Y. "Advances in 6G Wireless Communication Systems: Theory and Application". IEEE Communications Magazine, 2023.
13. Zhang, H., & Shi, Q. "Deep Learning for Signal Processing: A Comprehensive Overview". IEEE Signal Processing Magazine, 2023.
14. Liu, Y., & He, B. "Quantum Communications: Challenges and Future Directions". Journal of Quantum Information Science, 2023.
15. Tang, F., & Li, H. "Error Correction in Deep Neural Networks for Signal Processing". IEEE Transactions on Communications, 2023.
16. Shen, Y., & Wei, L. "Machine Learning Approaches in Wireless Signal Processing". Wireless Communications and Mobile Computing, 2023.
17. Miller, R., & Kim, J. "Optimization Algorithms for MIMO Systems in 5G and Beyond". IEEE Access, 2023.
18. Xu, W., & Ma, Y. "AI-Driven Modulation Techniques in Communication Systems". IEEE Communications Letters, 2023.
19. Zhang, X., & Zhou, Y. "Adaptive Beamforming for 6G Networks: Algorithms and Architectures". IEEE Transactions on Wireless Communications, 2023.
20. Li, Z., & Huang, L. "Entropy Coding and Its Applications in Modern Data Compression". IEEE Journal on Selected Areas in Information Theory, 2023.
21. Park, J., & Kim, S. "Nonlinear Signal Processing Techniques in OFDM Systems". IEEE Wireless Communications, 2023.
22. Gupta, R., & Singh, A. "Data Security in Quantum Communication Networks". IEEE Transactions on Information Forensics and Security, 2023.
23. Wang, Y., & Liu, T. "High-Efficiency Demodulation Techniques for 6G Applications". Journal of Communications and Networks, 2023.
24. Roy, S., & Bose, P. "Performance Evaluation of LDPC Codes in 5G and 6G Networks". IEEE Transactions on Communications, 2023.
25. Yang, L., & Xiong, Y. "Channel Estimation in 6G Terahertz Communications". IEEE Transactions on Wireless Communications, 2023.
26. Sun, Z., & Du, Y. "Next-Generation Signal Processing for Massive MIMO Systems". IEEE Access, 2024.
27. Chen, X., & Zhou, Z. "Multi-User Detection and Decoding in 6G Communication Systems". IEEE Transactions on Signal Processing, 2023.
28. Fang, H., & Yang, Z. "Wideband Spectrum Sensing in Cognitive Radio Networks for 6G". IEEE Transactions on Wireless Communications, 2023.
29. Johnson, E., & Brown, D. "Shannon Capacity Limits in 6G Communication Systems". Journal of Information and Communication Engineering, 2023.
30. Zhao, Y., & Yang, X. "The Role of AI in Enhancing the Efficiency of Wireless Communication Systems". IEEE Wireless Communications, 2024.
31. Tang, J., & Wang, K. "Signal Detection in Multi-Antenna Communication Systems Using Deep Learning". IEEE Transactions on Communications, 2023.
32. Zhou, F., & Liu, J. "Network Coding for Secure Communications in 6G". IEEE Transactions on Information Theory, 2023.
33. Yang, F., & Zhang, L. "Coherent and Noncoherent Detection Techniques in 6G Wireless Systems". IEEE Transactions on Wireless Communications, 2023.
34. Chen, H., & Lin, M. "Sparse Signal Recovery for Compressive Sensing in Wireless Communication". IEEE Transactions on Signal Processing, 2023.
35. Li, J., & Liu, G. "The Role of Polar Codes in Next-Generation Communication Systems". IEEE Transactions on Communications, 2023.

36. Wei, Y., & Zhang, B. "Channel Estimation and Equalization for High-Density MIMO Systems". IEEE Transactions on Wireless Communications, 2023.
37. Gupta, V., & Sharma, P. "Optimization of Error-Correcting Codes for 6G Communication Systems". IEEE Transactions on Information Theory, 2023.
38. Wang, X., & Zhao, L. "AI-Based Resource Allocation in Next-Generation Wireless Networks". IEEE Access, 2023.
39. Roy, R., & Singh, M. "Reinforcement Learning for Dynamic Spectrum Access in Cognitive Radio Networks". IEEE Transactions on Wireless Communications, 2023.
40. Chen, W., & Yu, F. "Signal Modulation Techniques for 6G Satellite Communication Systems". IEEE Transactions on Communications, 2024.
41. Park, M., & Lee, K. "Quantum Error Correction in Next-Generation Communication Networks". IEEE Journal on Selected Areas in Quantum Communications, 2023.
42. Wang, D., & Li, X. "Deep Learning for Channel Estimation in Massive MIMO Systems". IEEE Communications Magazine, 2023.
43. Zhao, F., & Chen, S. "Secure Communications in 6G Networks: Challenges and Opportunities". IEEE Transactions on Communications, 2023.
44. Zhang, X., & Zhao, Y. "Frequency-Selective Channel Estimation for Millimeter-Wave Systems". IEEE Transactions on Wireless Communications, 2023.
45. Liu, W., & Sun, H. "AI-Driven Beamforming Techniques for 6G Wireless Systems". IEEE Wireless Communications, 2024.
46. Yang, J., & Zhao, Q. "Efficient Resource Management for 6G Communication Networks". IEEE Access, 2023.
47. Sun, L., & Yang, W. "Performance of LDPC Codes in Ultra-Reliable Low-Latency Communications". IEEE Transactions on Wireless Communications, 2023.
48. Tang, R., & Liu, P. "Machine Learning for Error Detection in Communication Systems". IEEE Transactions on Signal Processing, 2023.
49. Zhao, J., & Zhang, T. "Massive MIMO Signal Processing for 6G Networks". IEEE Transactions on Communications, 2024.
50. Wang, K., & Yu, X. "Advanced Signal Processing for Terahertz Communication Systems". IEEE Transactions on Wireless Communications, 2023.
51. Xie, Y., & Zhou, P. "AI-Powered Error Correction for Next-Generation Communication Systems". IEEE Access, 2023.
52. Zhang, L., & Zhao, Y. "Cognitive Radio and Dynamic Spectrum Management in 6G". IEEE Communications Letters, 2023.
53. Liu, Y., & Zhao, X. "Data Transmission in Quantum Communication Networks". IEEE Journal on Selected Areas in Communications, 2023.
54. Li, Z., & Wang, H. "Signal Detection Techniques in 6G Wireless Communication". IEEE Transactions on Signal Processing, 2023.
55. Zhang, T., & Sun, Y. "High-Efficiency Encoding and Decoding Techniques for 6G Applications". IEEE Transactions on Information Theory, 2023.
56. Zhao, H., & Li, Y. "Efficient Signal Modulation for Multi-User Communication in 6G Networks". IEEE Wireless Communications, 2023.
57. Yang, Z., & Liu, X. "AI-Assisted Signal Processing for Low-Latency Communication Systems". IEEE Transactions on Wireless Communications, 2024.
58. Wang, R., & Zhang, P. "Spectrum Sensing for Ultra-Reliable Low-Latency Communications in 6G". IEEE Transactions on Communications, 2023.
59. Li, C., & Liu, G. "Security in Wireless Communication Systems Using Blockchain Technology". IEEE Transactions on Information Forensics and Security, 2023.

60. **Yang, K., & Zhao, J.** "Optimizing Channel Capacity in 6G Networks Using Deep Learning". IEEE Transactions on Communications, 2023.
61. **Zhang, F., & Liu, H.** "Error-Correcting Codes for Quantum Communication Networks". IEEE Transactions on Communications, 2023.

ЗМІСТ

ВСТУП.....	3
Загальні відомості про теорію інформації і передачі сигналів	3
Структура системи зв'язку.....	6
Основні характеристики сигналів	7
Демодуляція і декодування.....	17
Цифрове кодування неперервних повідомлень	18
Завади та спотворення в каналі	19
Основні характеристики системи зв'язку	21
Питання, задачі і вправи	23
РОЗДІЛ 1. МАТЕМАТИЧНІ МОДЕЛІ ДЕТЕРМІНОВАНИХ СИГНАЛІВ	27
1.1 Детерміновані сигнали.....	28
1.1.1 Періодичні сигнали	29
1.1.2.Неперіодичні сигнали.....	31

1.2	Часове представлення сигналів. Дискретизація інформації	33
1.3	Квантування за рівнем	46
1.4	Геометричні представлення сигналів.....	51
1.5	Розклад сигналів в узагальнений ряд Фур'є.....	62
	Контрольні питання.....	68
	Запитання, задачі і вправи	69
РОЗДІЛ 2. ВИКОРИСТАННЯ ШУМОПОДІБНИХ СИГНАЛІВ У СИСТЕМАХ ЗВ'ЯЗКУ		78
2.1	Основні поняття визначення.....	78
2.2	Кодовий поділ абонентів.....	85
2.3	Основні структурні схеми широкосмугових систем зв'язку ШСС.....	94
2.4	Шумоподібні сигнали та їх спектри.....	101
2.5	Оптимальна обробка сигналів	106
2.6	Кореляційні функції сигналів	109
2.7	Основні типи ШПС	117
	Контрольні питання.....	125
РОЗДІЛ 3. МЕТОДИ ОБРОБКИ СИГНАЛІВ		125
3.1	Загальні відомості про прийом сигналів.....	126
3.2	Прийом сигналів із накопиченням	129
3.3	Кореляційний прийом	136
3.4	Автокореляційний прийом.....	138
3.5	Когерентний та некогерентний прийоми.....	141
3.6	Прийом на узгоджений фільтр	146
	Контрольні питання.....	153
РОЗДІЛ 4. РОЗРОБКА АЛГОРИТМІВ ОПТИМАЛЬНОГО ПРИЙОМУ СИГНАЛІВ ІЗ ФАЗОРІЗНИЦЕВОЮ МОДУЛЯЦІЄЮ ВИСОКОЇ КРАТНОСТІ		154
	Постановка задачі.....	154
4.1	Фазорізнцева модуляція першого та другого порядків та властивості інваріантності.....	155
4.2	Багатопозиційні фазомодульовані сигнали	164
4.3	Фазорізнцева модуляція високого порядку	179
4.4	Система з перемінними параметрами, інваріантна до адитивної перешкоди	187

4.5. Система з постійними параметрами, варіантна до неадитивної завади.....	191
4.6. Адаптивна система, інваріантна до адитивної перешкоди	198
Контрольні питання.....	200
РОЗДІЛ 5. ПОРІВНЯННЯ МЕТОДІВ ОПТИМАЛЬНОЇ ОБРОБКИ	
БАГАТОПОЗИЦІЙНИХ СИГНАЛІВ	201
Постановка задачі.....	201
5.1. Оптимальний некогерентний прийом сигналів із ФРМ першого порядку	204
5.2. Оптимальний некогерентний прийом сигналів із ФРМ другого порядку	218
5.3. Автокореляційна обробка фазомодульованих сигналів	224
5.3.1. Застосування автокореляційної обробки сигналів.....	224
5.3.2. Оптимальні автокореляційні (енергетичні) демодулятори сигналів із ФРМ першого порядку.....	230
5.4.Розробка і дослідження алгоритму обробки сигналу в багатоканальних модемах з ортогональним розподілом каналів.....	241
Постановка задачі.....	241
5.4.1 Особливості і шляхи рішення задачі когерентного прийому багаточастотних групових сигналів.....	243
5.4.2. Синтез оптимального алгоритму когерентної обробки багаточастотних групових сигналів.....	247
5.4.3. Моделювання когерентного багаточастотного модему.	256
Контрольні питання.....	265
РОЗДІЛ 6. РОЗРАХУНОК ОБСЯГУ КЕРУЮЧОЇ ІНФОРМАЦІЇ ДЛЯ	
ЗАБЕЗПЕЧЕННЯ ТОЧНОСТІ ПОКАЗНИКІВ ЯКОСТІ СИСТЕМ І МЕРЕЖ ЗВ'ЯЗКУ	
.....	265
6.1.Кількісна міра інформації.....	265
6.2.Ентропія джерела дискретних повідомлень	269
6.3.Ентропія неперервних повідомлень.....	270
6.4.Ефективність систем передавання інформації	274
6.5.Ентропія розподілу ймовірностей керованого процесу	275
6.6. Ентропія в замкнених системах управління.....	284
6.7.Ентропія та інформація в розімкнених системах управління.....	293
6.8.Розрахунок обсягу управляючої інформації для забезпечення необхідної точності параметрів мережі	296
Контрольні питання.....	304

РОЗДІЛ 7. ОПТИМАЛЬНИЙ ПРИЙОМ ДИСКРЕТНИХ СИГНАЛІВ. ОСНОВИ ТЕОРІЇ ЗАВАДОСТІЙКОСТІ	305
7.1. ЗАГАЛЬНІ ВІДОМОСТІ ПРО ЗАВАДОСТІЙКОСТІ.....	305
7.2. Критерії оптимального прийому сигналів	306
7.3. Оптимальний прийом дискретних сигналів	310
7.4. Імовірність помилки при когерентному прийому двійкових сигналів.....	314
7.5. Імовірність помилки при когерентному прийому багатопозиційних сигналів	320
7.6. Некогерентний приймання дискретних сигналів	323
7.7. Передача дискретних сигналів по каналам з змінними параметрами	327
Контрольні питання.....	333
РОЗДІЛ 8. ОСНОВНІ ПОНЯТТЯ ТЕОРІЇ ІНФОРМАЦІЇ	334
8.1. Кількісна міра інформації	334
8.2. Ентропія джерела дискретних повідомлень	346
8.3. Взаємна інформація.....	358
8.4. Швидкість і пропускна спроможність дискретного каналу без перешкод	362
8.5. Оптимальні статистичні кодування повідомлень	364
8.6. Дискретні канали з завадами	368
8.7. Інформація в неперервних повідомленнях. Епсілон-ентропія.....	372
8.8. Швидкість та пропускна спроможність неперервного каналу. Формула Шеннона.....	376
8.9. Визначення щільності розподілу ймовірностей і втрат інформації на виході нелінійного елемента.	382
8.10. Ефективність систем передачі інформації.....	396
Контрольні питання.....	399
РОЗДІЛ 9. ОПТИМАЛЬНИЙ ПРИЙОМ НЕПЕРЕРВНИХ ПОВІДОМЛЕНЬ. ОПТИМАЛЬНА ФІЛЬТРАЦІЯ	405
9.1. Оптимальний прийом неперервних повідомлень	405
9.2. Синтез пристроїв зв'язку при заданій структурній схемі	410
9.3. Фізичний зміст оптимальної смуги перепуску.....	411
9.4. Синтез оптимальних пристроїв (систем) зв'язку	413
9.5. Синтез оптимальних фільтрів Калмана методом простору станів	420
9.6. Приклади синтезу оптимальних фільтрів	434

9.7. Оптимальний фільтр для повідомлення, описуваного степеневим поліномом	445
9.8. Поняття про оптимальну нелінійну фільтрацію	453
Контрольні питання.....	458
РОЗДІЛ 10. КОДУВАННЯ ДЖЕРЕЛ ІНФОРМАЦІЇ ТА КАНАЛІВ ЗВ'ЯЗКУ	462
10.1 Основні поняття та визначення	462
10.2 Кориговальні коди	466
10.3 Циклічні коди.....	476
10.4.Коди Боуза - Чоудхурі – Хоквінгема	481
10.5 Несистематичні коди.....	485
10.6 Ітеративні коди	486
10.7 Ефективність систематичних коригувальних кодів.....	488
10.8 Мажоритарне декодування	489
10.9 Узагальнення теорії кодування на недвійкові коди	495
10.9.1 Ітеративні і каскадні коди.....	497
10.9.2 Кодування в каналах з пам'яттю	498
10.10 Системи зі зворотним зв'язком	499
10.11 Сполучення процедур демодуляції і декодування.....	503
10.12 Згорткові (гратчасті) коди	513
Контрольні питання.....	522
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:	529

This image shows a full page of blank, lined paper. It features approximately 20 horizontal blue or grey lines spaced evenly apart, typical of notebook paper. The lines extend across the entire width of the page, leaving small margins at the top and bottom. There are no vertical lines, text, or other markings on the page.

[illegible]

[illegible]

**В.П. Шульга, Л.Н. Беркман, Н.В. Галаган,
К.С. Нестеренко, О.Г. Варфоломєєва, Н.С. Хаб'юк**

СУЧАСНІ МЕТОДИ ПЕРЕДАЧІ ТА ОБРОБКИ ІНФОРМАЦІЇ В ІНФОКОМУНІКАЦІЙНИХ МЕРЕЖАХ

Підручник



Підписано до друку 04.04.2025 .
Формат 60х84 1/16. Папір офсетний.
Друк цифровий. Ум. друк. арк. 33,75
Віддруковано ФО-П Білецький Р.Г.,
Типографія Аграр Медіа Прінт.
Адреса потужностей: м. Київ, вул. В.Некрасова, 3
e-mail: info@agrarmedia.com
тел.: (067) 407 87 23

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців ДК №81116 від 21.06.2024р.
www.agrarmedia.com