

Міністерство освіти і науки України
Луцький національний технічний університет



ТЕОРІЯ ЕЛЕКТРИЧНОГО ЗВ'ЯЗКУ

Конспект лекцій

для здобувачів першого (бакалаврського) рівня вищої освіти
освітньої програми «Комп'ютеризовані телекомунікаційні мережі»
галузі знань 17 Електроніка, автоматизація та електронні комунікації
спеціальності 172 Електронні комунікації та радіотехніка
денної та заочної форм навчання

Луцьк 2025

УДК 621.391(07)

Т – 33

Електронна копія друкованого видання передана для внесення в репозитарій ЛНТУ

Директор бібліотеки _____ Наталія ПОЛЩУК

Рекомендовано до видання вченою радою факультету комп'ютерних та інформаційних технологій ЛНТУ, протокол № __ від « __ » _____ 2025 року.

Голова вченої ради ФКІТ _____ Інна КОНДІУС

Розглянуто і схвалено на засіданні кафедри електроніки та телекомунікацій ЛНТУ, протокол № __ від « __ » _____ 2025 року.

Завідувач _____ Валентин ЗАБЛОЦЬКИЙ к.т.н., доц. кафедри електроніки та телекомунікацій ЛНТУ

Укладач: _____ Наталія ЯКИМЧУК к.т.н., доц. кафедри електроніки та телекомунікацій ЛНТУ

Рецензент: _____ Віктор ЛИШУК к.т.н., доц. кафедри електроніки та телекомунікацій ЛНТУ

Відповідальний за випуск: _____ Валентин ЗАБЛОЦЬКИЙ к.т.н., доц., завідувач кафедри електроніки та телекомунікацій ЛНТУ

Т - 33 Теорія електричного зв'язку. Конспект лекцій для здобувачів першого (бакалаврського) рівня вищої освіти освітньої програми «Комп'ютеризовані телекомунікаційні мережі» галузі знань 17 Електроніка, автоматизація та електронні комунікації, спеціальності 172 Електронні комунікації та радіотехніка, всіх форм навчання / уклад. Н. М. Якимчук. Луцьк: ЛНТУ, 2025. 190 с.

Видання містить виклад основного теоретичного лекційного матеріалу з дисципліни «Теорія електричного зв'язку». Призначене для здобувачів першого (бакалаврського) рівня вищої освіти освітньої програми «Комп'ютеризовані телекомунікаційні мережі».

Н.М. Якимчук, 2025

ЗМІСТ

ВСТУП	4
Змістовий модуль 1. Основи теорії процесів в системах електрозв'язку	5
Тема 1. Загальні поняття про систему електрозв'язку та сигнали	5
Тема 2. Математичні моделі сигналів.....	19
Тема 3. Спектральні характеристики сигналів.....	31
Тема 4. Випадкові процеси в системах електрозв'язку.....	46
Тема 5. Методи формування та перетворення сигналів.....	58
Тема 6. Проходження сигналів через канали зв'язку.....	83
Змістовий модуль 2. Елементи теорії інформації в електрозв'язку.....	94
Тема 7. Інформаційні характеристики джерел повідомлень та каналів зв'язку.....	94
Тема 8. Теорія кодування повідомлень.....	102
Тема 9. Теорія оптимального прийому сигналів.....	120
Тема 10. Перешкодостійкість систем передачі повідомлень...	137
Тема 11. Оптимальна фільтрація випадкового сигналу.....	148
Тема 12. Основи адаптивного придушення перешкод.....	155
Тема 13. Багатоканальний зв'язок та розподіл інформації.....	167
Тема 14. Ефективність систем електрозв'язку.....	179
ІНФОРМАЦІЙНІ ДЖЕРЕЛА.....	189

ВСТУП

Дисципліна «Теорія електричного зв'язку» є однією з ключових складових навчальної програми для студентів спеціальності 172 «Електронні комунікації та радіотехніка». Вона забезпечує фундаментальні знання, необхідні для розуміння принципів функціонування сучасних телекомунікаційних систем і мереж, які є основою інформаційного суспільства.

Розвиток телекомунікаційних технологій за останні десятиліття відкрив широкі можливості для передачі інформації, одночасно висуваючи нові виклики, пов'язані з ефективністю, якістю та безпекою зв'язку. Глибоке розуміння теорії електричного зв'язку дозволяє майбутнім фахівцям адаптуватися до цих змін, розробляти і вдосконалювати системи зв'язку, а також вирішувати задачі, пов'язані із забезпеченням стійкості сигналів до шумів та перешкод.

Основною метою дисципліни є вивчення основ теорії сигналів, передачі інформації через різні канали зв'язку, аналізу шумів та їх впливу на якість зв'язку, а також методів оптимального прийому сигналів. Особлива увага приділяється математичним моделям, які описують процеси передачі інформації, та їхньому практичному застосуванню для проектування та аналізу телекомунікаційних систем.

У рамках курсу студенти вивчають спектральні характеристики сигналів, основи теорії кодування, принципи формування та перетворення сигналів, а також методи боротьби з перешкодами. Додатково розглядаються сучасні методи адаптивного фільтрування, принципи багатоканального зв'язку та оцінювання ефективності роботи систем зв'язку. Це створює міцний фундамент для подальшого вивчення дисциплін, що спеціалізуються на проектуванні та експлуатації телекомунікаційних мереж.

Конспект лекцій побудовано таким чином, щоб забезпечити студентів структурованим та доступним матеріалом, який можна використовувати не лише для освоєння курсу, але й для самостійного поглиблення знань. Матеріал викладено в логічній послідовності, кожна тема супроводжується прикладами, графічними ілюстраціями та завданнями для самостійної роботи.

Опановуючи дисципліну «Теорія електричного зв'язку», студенти отримають змогу не лише засвоїти теоретичні основи передачі інформації, але й навчатися практично застосовувати ці знання у вирішенні реальних інженерних задач. Це забезпечить їх готовність до роботи в динамічній сфері електронних комунікацій, яка є невід'ємною частиною сучасного світу.

Змістовий модуль 1. Основи теорії процесів в системах електрозв'язку

Тема 1. Загальні поняття про систему електрозв'язку та сигнали

Інформація та форма її подання. Класифікація сигналів. Основні характеристики повідомлень і сигналів. Завади і спотворення в системах зв'язку. Структура системи зв'язку. Загальні відомості про канали передачі інформації.

1.1 Інформація, повідомлення, сигнали та перешкоди

Інформація – сукупність знань щодо деякого процесу, подій, предмету, які зменшують невизначеність, що була в отримувача до одержання ним цих знань. Для передавання або зберігання інформації використовуються різні знаки (або символи), які дозволяють представляти її в деякій формі.

Системи зв'язку призначені передачі інформації. Інформація надсилається за допомогою повідомлень. Отже, **повідомлення** – матеріальна форма подання інформації.

Прикладами повідомлень можуть бути текст телеграми, фраза у телефонній розмові, послідовність цифр передачі даних, зображення у системі фототелеграфії, послідовність зображень (кадрів) у системі телебачення тощо. Повідомлення є сукупністю знаків (символів).

Повідомлення можуть бути дискретними (що складаються із символів, що належать кінцевій множині – алфавіту) або неперервними (континуальними, аналоговими), що описуються функціями неперервного часу.

Інформація передається **каналом зв'язку** шляхом відправлення деяких сигналів, отже **сигнал** - це матеріальний носій повідомлення. Сигнали формуються шляхом зміни тих або інших параметрів фізичного носія відповідно до переданого повідомлення. Цей процес (зміни параметрів носія) прийнято називати **модуляцією**.

У сучасних системах зв'язку використовують в основному три види сигналів: електричні, електромагнітні та оптичні, які завдяки простоті їх генерування та перетворення найкраще пристосовані для передачі великих об'ємів даних на великі відстані. У системах електрозв'язку як дані (інформація) так і сигнали можуть бути представлені у двох основних формах: **цифровій** та **аналоговій**. Електричні сигнали, як правило, представлені у вигляді дискретних імпульсів, тоді як електромагнітні та оптичні сигнали можна представити як суми синусоїдних хвиль.

Зокрема, сигнали керування передаються виключно у цифровій формі у вигляді послідовності прямокутних імпульсів. Для безспотворної передачі цих сигналів необхідна смуга частот, що охоплює діапазон від нуля до нескінченності. Реальні канали мають кінцеву смугу частот, з якою необхідно узгодити сигнали, що передаються. Узгодження забезпечується, по-перше, шляхом модуляції – перенесення сигналу на задану смугу частот і, по-друге, шляхом кодування – перетворення даних у вигляд, що дозволяє виявляти та виправляти помилки, що виникають через перешкоди в каналі зв'язку. Водночас, при використанні високочастотних провідних та кабельних ліній, таких як екранована та неекранована вита пара, коаксіальні кабелі, сигнали можуть передаватися у їхньому природному вигляді без модуляції. Канали, що працюють без модуляції, називаються телеграфними та забезпечують передачу даних зі швидкістю, як правило, від 50 до 200 біт/с.

Таким чином, розуміння різних типів сигналів та середовища їх передачі є важливим для ефективного функціонування систем електрозв'язку.

Природною формою подання сигналу вважається його опис деякою функцією часу (залежною змінною найчастіше є напруга чи струм). Математичний опис та подання сигналів дозволяє створити математичну модель сигналу.

По виду математичного опису сигнали можна поділити:

Якщо математична модель дозволяє точно описати сигнал, такий сигнал називається **детермінованим**. У разі неможливості точного опису сигналу в будь-які моменти часу сигнал називається **випадковим**.

Якби передане повідомлення було детермінованим, тобто заздалегідь відомим з повною імовірністю, то передача його не мала б сенсу. Таке детерміноване повідомлення не містить інформації. Його можна використати лише для випробувань системи зв'язку або окремих її елементів. Тому повідомлення варто розглядати як випадкові події (або випадкові величини, випадкові функції). Інакше кажучи, має існувати деяка множина варіантів повідомлення, з яких реалізується з певною ймовірністю одиниць. Тому й сигнал є випадковою функцією.

Високочастотне модульоване коливання називається **радіосигналом**.

Сигнал без високочастотного заповнення є **відеосигналом**.

Якщо сигнал може бути описаний функцією $s(t) = s(t + T)$, де T – період, він називається **періодичним**.

При неможливості такого представлення сигнал є **неперіодичним**.

Сигнал, що описує в часі процес, що безперервно змінюється, називається **аналоговим**. Сигнал кінцевої тривалості – **імпульсний**.

Іноді зручно передавати лише значення неперервного сигналу (відліки чи вибірки), взяті окремі моменти часу. Такий квантований за часом сигнал називається **дискретним**. Якщо ж передавати не самі вибірки в вигляді коротких імпульсів, а їх числові значення, то спочатку необхідно ці значення отримати. Ця процедура у техніці зв'язку називається **квантуванням за рівнем**. Таким чином, сигнал, квантований за часом та рівнем, називається **цифровим**.

Цікаво відзначити, що детерміновані сигнали не несуть у собі жодної інформації. Однак з їх допомогою можна передавати інформацію, якщо випадковим буде розташування цих сигналів на часовій осі. Наприклад, телеграфний сигнал складається із семи імпульсів прямокутної форми із заданими параметрами. Перший (стартовий) та останній (стоповий) імпульси позначають початок та кінець посилки. Інформаційний зміст посилки залежить від літери алфавіту, що передається в даний момент, і представляє відповідну цій літері комбінацію струмових і безструмових посилок.

По виду повідомлень, що передаються, сигнали можна розділити на радіомовні, телевізійні, телеграфні і т. д.

По смузі частот сигнали зазвичай поділяються на вузькосмугові та широкосмугові.

Для **широкосмугових** сигналів $\Delta F/F_{\text{ср}} \gg 1$,
де $\Delta F = F_{\text{max}} - F_{\text{min}}$ – абсолютна ширина спектра сигналу,
 $F_{\text{ср}} = (F_{\text{max}} + F_{\text{min}})/2$ – середня частота спектра сигналу,
 F_{max} – максимальна частота в спектрі сигналу,
 F_{min} – мінімальна частота у діапазоні сигналу.

Для **вузькосмугових** сигналів $F/F_{\text{ср}} < 1$.

Сигнали так само поділяються на **складні** та **прості** в залежності від величини бази сигналу B (добуток тривалості сигналу на ширину смуги його спектру).

Для складних сигналів $B > 1$,
де $\Delta F \cdot \Delta T$ – база сигналу, ΔF – абсолютна ширина спектру сигналу, ΔT – тривалість сигналу. Для простих сигналів $B = 1$.

По виду модуляції сигнали розрізняються за ознакою того параметра, який змінюється за законом повідомлення, що передається. Так як будь-яке гармонійне коливання характеризується амплітудою, частотою та миттєвою фазою, то і радіосигнали бувають з амплітудною модуляцією (АМ), частотною

(ЧМ) і фазовою модуляцією (ФМ). В даний час в системах зв'язку використовується велика різноманітність сигналів зі складними видами модуляції, наприклад з амплітудно-імпульсною модуляцією (АІМ), кодово-імпульсною модуляцією (КІМ), широтно-імпульсною модуляцією (ШІМ). На даний час розроблено не один десяток складних видів модуляції і, природно, велика кількість відповідних сигналів з різними характеристиками.

На рис. 1.1 наведено осцилограми різних сигналів, що широко застосовуються в системах зв'язку.

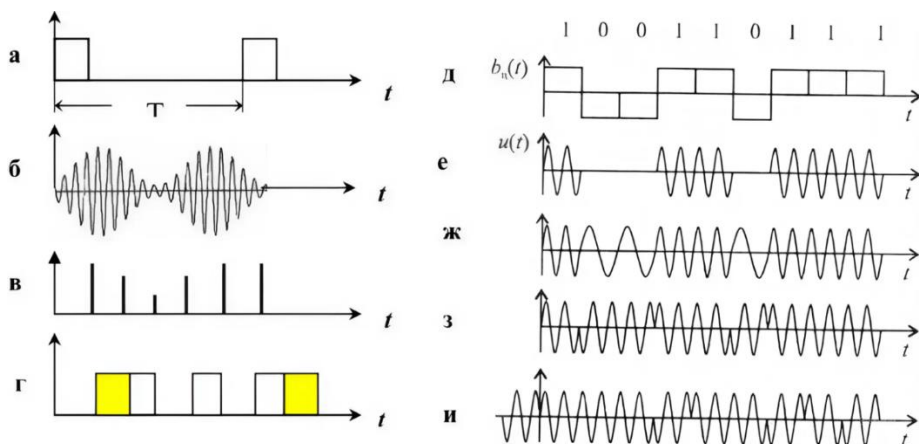


Рисунок 1.1 – Типові приклади сигналів електровз'язку

На цьому рисунку зображені наступні сигнали: а – періодичний імпульсний, б – неперервний (аналоговий) радіосигнал з АМ, в – дискретний, г – випадковий, д – цифровий кодований, е – цифровий з АМ, ж – цифровий із ЧС, з – цифровий з ФМ, та – цифровий з фазовою маніпуляцією.

Слід зазначити, що жорсткої класифікації до реальних сигналів застосувати неможливо. Наприклад, сигнал (рис. 1.1 а) можна класифікувати як детермінований періодичний імпульсний відеосигнал, а сигнал (рис. 1.1 з) як випадковий цифровий радіосигнал з ЧС.

В теорії електричного зв'язку прийнято розглядати сигнал як «об'єкт транспортування». З цієї точки зору сигнал можна описати трьома «габаритними характеристиками», подібними до довжини, ширини і висоти вантажу, що перевозиться, скажімо, залізницею. Перша з таких характеристик – **тривалість сигналу** T_c , яка вимірюється в секундах (с).

Будь-який сигнал можна уявити сумою (суперпозицією) гармонійних коливань з певними частотами, тому друга «габаритна характеристика» – **ширина спектру**, або смуга частот сигналу ΔFc , що дорівнює різниці найвищої та нижчої частот його гармонійних складових і вимірюється в герцах (Гц). У техніці зв'язку спектр сигналу часто свідомо скорочують. Це обумовлено тим, що апаратура й лінія зв'язку мають обмежену смугу частот. Скорочення спектру здійснюється виходячи із припустимих перекручувань сигналу. Наприклад, при телефонному зв'язку потрібно, щоб мова була розбірлива й щоб кореспонденти могли розпізнати один одного за голосом. Для виконання цих умов досить передати мовний сигнал у смузі від 300 до 3400 Гц. Передача більш широкого спектра мови в цьому випадку недоцільна, тому що веде до технічних ускладнень і збільшення витрат. Аналогічно необхідна ширина спектра телевізійного сигналу визначається необхідною чіткістю зображення. При стандарті в 625 рядків верхня частота сигналу досягає 6 МГц.

Спектр сигналу зображення багато ширше за спектр сигналу звукового супроводу. Це істотно ускладнює побудову систем телевізійного віщання в порівнянні із системами звукового мовлення. Ширина спектра телеграфного сигналу залежить від швидкості передачі й звичайно приймається рівною $\Delta Fc = 1,5\nu$, де ν - швидкість передачі (телеграфування) у Бодах (Бод), тобто число символів, переданих у секунду. Так, при телетайпній передачі $\nu = 50$ Бод й тоді $\Delta Fc = 75$ Гц.

Третьою «габаритною» характеристикою є **динамічний діапазон Dc** – це відношення найбільшої миттєвої потужності сигналу до тієї найменшої потужності, яку необхідно відрізнити від нуля при заданій якості передачі. Він виражається в децибелах (дБ), а також в неперах (Нп), причому $1\text{Нп} = 8,68$ дБ, а $1\text{дБ} = 0,115$ Нп. Так, динамічний діапазон мови диктора, наприклад, дорівнює 25...30 дБ, невеликого вокального ансамблю 45...65 дБ, симфонічного оркестру 70...95 дБ. Щоб уникнути перевантажень каналу в радіомовленні, динамічний діапазон часто скорочують до 35...45 дБ.

Добуток цих трьох величин називається **об'ємом сигналу**:

$$Vc = Tc\Delta FcDc. \quad (1.1)$$

Чим більший об'єм сигналу, тим більше інформації можна «вкласти» у цей обсяг і тим складніше такий сигнал передавати по каналу зв'язку з необхідною якістю.

1.2 Завади і спотворення в системах зв'язку

Корисні сигнали в системах зв'язку відрізняються від інших тим, що корисні сигнали служать для передачі повідомлень, в той час як заважаючі сигнали є причиною їх спотворення (втрати інформації).

Корисні сигнали рідко присутні в електричних колах у чистому вигляді. Під час передачі вони зазнають впливу завад, що призводить до їх спотворення. Це, у свою чергу, викликає помилки у відтворенні повідомлень.

Завада (перешкода) визначається як будь-який випадковий вплив на корисний сигнал, що погіршує вірність прийому та відтворення інформації, переданої через лінію зв'язку, може мати чітко визначене джерело.

Необхідно відзначити, що те саме коливання може бути корисним сигналом по відношенню, наприклад, до однієї системи зв'язку або радіолокації і завадою – по відношенню до іншої. Варто також відзначити, що всі завади, як і всі інформаційні сигнали, є випадковими (якщо перешкода детермінована, то її можна виключити з коливання, що спостерігається, і таким чином позбутися її шкідливого впливу на повідомлення).

На рис. 1.2 наведено приклади випадкового сигналу та випадкової (шумової) перешкоди.

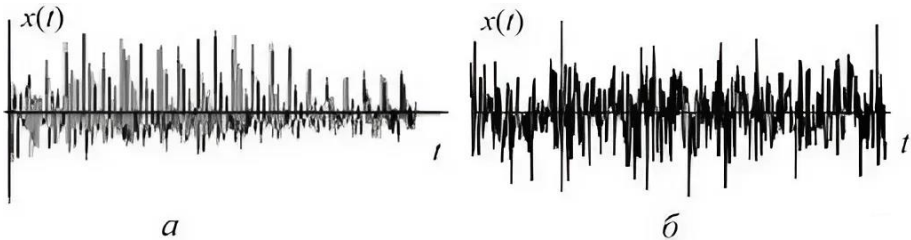


Рисунок 1.2 – Випадковий (мовний) сигнал (а) та випадкова перешкода (шум) (б)

Для аналогових сигналів внаслідок впливу завад характерна деградація якості, що проявляється у зміні форми та потужності сигналу. Проходячи через середовище передачі, аналогові сигнали зазнають деградації: зменшується амплітуда сигналу зі збільшенням відстані. В результаті, форма отриманого сигналу може бути подібною до первісної, але вона завжди буде спотвореною.

Цифрові сигнали, з іншого боку, зазнають бітових помилок (bit errors), коли на виході отримується інша послідовність нулів і одиниць, ніж на вході.

Завади поділяються на зовнішні та внутрішні, а також на **адитивні**, **мультиплікативні** та **змішані**.

Аддитивні завади, що впливають на сигнал, класифікуються за основними властивостями на три класи (рис.1.3):

– **Зосередженими по спектру** (вужькосмуговими) називають завади, основна частина потужності яких припадає на окремі ділянки діапазону частот, менших за смугу пропускання системи зв'язку (це можуть бути електромагнітні завади, викликані пристроями, такими як двигуни, генератори або інші електричні машини, які створюють магнітні поля і заважають передачі сигналу, чи інтерференція від сусідніх каналів (**Перехресні завади Crosstalk**), коли в багатопровідній кабельній системі сигнал одного проводу індукується в сусідній провід через електромагнітне поле. Це може траплятися в телефонних лініях, кабельних мережах або навіть у бездротових мережах, де перехресні завади можуть призводити до витоку інформації або зниження якості сигналу.).

Для зниження таких завад використовують різні методи, такі як екранування кабелів, збільшення відстані між каналами, використання спеціальних типів проводів (наприклад, витої пари), а також застосування фільтрів і пристроїв корекції сигналу.

– **Імпульсними** (зосередженими в часі) називають регулярну або хаотичну послідовність імпульсних сигналів, однорідних з корисним сигналом. Джерелами таких перешкод є цифрові і комутуючі елементи кіл або працюючі поруч з ними пристрої.

– **Флуктуаційна завада** (флуктуаційний шум) представляє випадковий процес з нормальним розподілом – гауссівський процес. Ці перешкоди мають місце практично у всіх реальних каналах зв'язку, і їх називають шумами. Отже, **шум** – це випадкові, непередбачувані флуктуації сигналу, які можуть заважати сприйняттю корисної інформації. Шум може виникати з різних джерел, таких як термічні коливання, електромагнітні завади або зміни в середовищі передачі і присутній у всіх компонентах системи. Шуми зазвичай не обмежені ні в часі, ні по спектру і впливає на якість сигналу, знижуючи його амплітуду та розпізнаваність, що робить їх особливо складними для корекції.

Розглянемо основні види шумів, які впливають на електричний зв'язок: адитивний шум, тепловий шум, шум мерехтіння та квантовий шум.

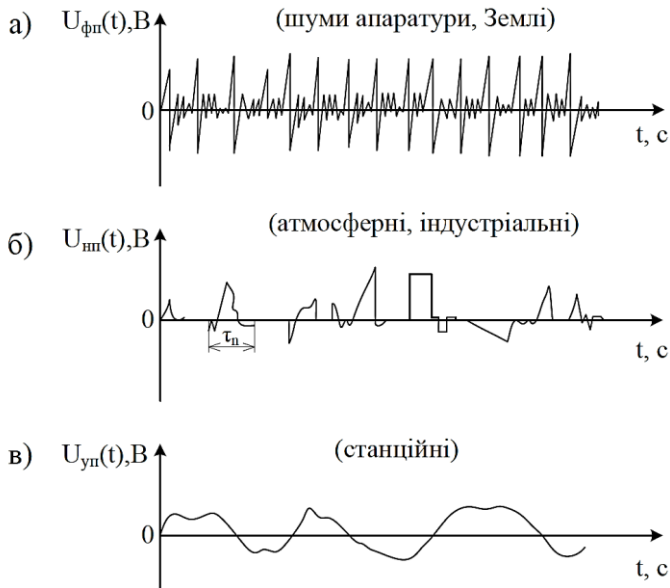


Рисунок 1.3 – Графіки завад а) флуктуаційна, б) імпульсна, в) вузькосмугова

1) **Тепловий шум** (шум Джонсона) – виникає через тепловий рух електронів у провіднику. Він існує в будь-якому каналі передачі та в кожному передаючому пристрої. Природа цього шуму робить його присутнім у будь-якому середовищі, де передається електричний сигнал. Його величина несуттєва і може бути визначена за формулою:

$$N = kTB \quad (1.2)$$

де N – потужність шуму, K – стала Больцмана $1,38 \cdot 10^{-23}$ Дж/К, T – температура середовища, K , B – ширина спектру, Гц.

Для систем зв'язку оцінюють $N = -174$ dBm/1Hz.

2) **Шум мерехтіння** (рожевий шум) – це вид шуму, потужність якого зменшується з підвищенням частоти. Цей шум характерний для низькочастотних середовищ і має складнішу природу, але може виникати через нерегулярні коливання в електронних компонентах.

3) **Квантовий шум** – пов'язаний із принципами квантової механіки і зазвичай проявляється в системах передачі світлових сигналів. Він виникає через випадковий характер випромінювання та **поглинання** фотонів у середовищі передачі.

4) **Інтермодуляційний шум** (інтерференція) – з'являється, коли сигнали на однакових частотах передаються через одне і те ж середовище і взаємно впливають один на одного. Це можна уявити як одночасну розмову багатьох людей в аудиторії, де різні голоси змішуються, створюючи перешкоди. Підвищення потужності сигналу може збільшувати інтермодуляційний шум, але цього можна уникнути, розносячи різні передавачі за частотами або за часом, щоб зменшити взаємний вплив. Це дозволяє забезпечити якість сигналу та знизити вплив шуму.

5) **Імпульсний шум** – це короткий пік шуму, наприклад, викликаний блискавкою якимись електричними пошкодженнями або перешкодами в системі. (трансформатор, мікрохвильова)

Основні види спотворень сигналів, що виникають внаслідок дії завод:

1) **Згасання сигналу в системах зв'язку** (attenuation) – це зменшення амплітуди сигналу під час його передачі через канал зв'язку. Це явище викликане фізичними властивостями середовища, через яке проходить сигнал. Наприклад, навіть у високоякісних провідниках, таких як мідь, частина електромагнітної енергії втрачається і перетворюється на теплову енергію. Як наслідок, амплітуда сигналу зменшується, що може призвести до спотворення інформації — цього явища називають спотворенням згасання (attenuation distortion).

Коли сигнал проходить через канал зв'язку, важливо, щоб його амплітуда залишалася достатньою для коректної інтерпретації на прийомі. Якщо сигнал зменшується до критично низького рівня, його може бути важко розпізнати, і в таких випадках необхідно використовувати повторювачі сигналу для підсилення.

Особливості згасання сигналу включають кілька важливих аспектів:

– Затухання сигналу є функцією від відстані: чим далі сигнал проходить, тим більше його амплітуда зменшується.

– Отриманий сигнал повинен мати достатній рівень потужності для розпізнавання його приймачем, перевищуючи чутливість останнього.

– Сигнал, що передається, повинен мати більшу потужність, ніж шум у каналі, щоб його можна було коректно сприймати.

– Згасання залежить від частоти, на якій передається сигнал: для різних частот існує різний рівень затухання. Зазвичай, чим вища частота сигналу, тим більше його затухання в каналі. (У 2G мережі в діапазонах 800 МГц з потужністю 20-40 Вт сигнал загасає повільніше ніж на більш високих

діапазонах частот, радіус стільника тут близько десяти кілометрів; передавачі 3G та 4G працюючи в діапазонах наприклад 2 і 6 ГГц вже давали вам радіус порядку одного кілометра, в мережах 5G – діапазон міліметрових хвиль, вище 25 ГГц – до 100 метрів радіус соти).

– Для того щоб компенсувати відмінності в згасанні на різних частотах ми використовуємо методику еквалайзера, для чого змінюємо амплітуду певної частоти.

2) **Спотворення сигналу через затримку (Delay Distortion)** – виникає, коли сигнал, проходячи через середовище передачі, зазнає запізнення через відбиття від об'єктів або інших факторів. Такі спотворення часто зустрічаються у радіозв'язку, коли сигнал передається антеною і поширюється в просторі.

Розглянемо приклад: антена передає сигнал, який може досягти приймача по найкоротшому прямому шляху. Однак сигнал також може відбиватися від будівель або інших об'єктів. У результаті ми отримуємо кілька копій одного й того ж сигналу, які доходять до приймача з різними затримками. Таким чином, на приймач надходить основний сигнал разом із ще кількома запізненими копіями. Ці запізнені копії сигналу починають взаємодіяти з оригінальним сигналом, що призводить до викривлень і появи напруги, якої спочатку не було. Це явище відоме як **міжсимвольна інтерференція**. Запізнення сигналу впливає на якість передачі, оскільки зменшується точність розрізнення окремих символів у переданому повідомленні.

Таким чином, розуміння природи шумів та завад є важливим для забезпечення надійності систем електрозв'язку та підвищення якості передачі даних. Для зменшення впливу спотворень через затримку в системах зв'язку використовують спеціальні методи обробки сигналу, що допомагають компенсувати затримки і відновити оригінальну інформацію.

1.3 Загальні засади побудови систем електрозв'язку

Система електричного зв'язку забезпечує передавання повідомлення з певною якістю від джерела повідомлень до отримувача повідомлень. Система електричного зв'язку може забезпечувати одностороннє передавання повідомлень (**мовлення**) або двостороннє передавання повідомлень (**зв'язок**). У першому випадку має місце **симплексне** (тільки одностороннє) передавання, у другому випадку має місце двостороннє передавання: повний **дуплекс** – система забезпечує одночасно приймання і передавання

повідомлень; **напівдуплекс** – система забезпечує приймання і передавання повідомлень по черзі.

Сучасна система зв'язку є складною сукупністю пристроїв, що виконують перетворення повідомлень і сигналів з метою найбільш ефективної передачі інформації. По виду повідомлень, що передаються, розрізняють:

- 1) телеграфію (передача тексту),
- 2) телефонію (передача мови),
- 3) фототелеграфію (передача нерухомих зображень),
- 4) телебачення (передача рухомих зображень),
- 5) телеметрію (передача результатів вимірів),
- 6) телеуправління (передача керуючих команд),
- 7) передачу даних (у обчислювальних системах та АСУ).

За діапазоном частот – відповідно до декадного поділу діапазонів електромагнітних хвиль від міріаметрових ($3 \div 30$) кГц до дециміліметрових ($300 \div 3000$) ГГц.

За призначенням – мовні (високоякісна передача мови, музики, відео від малої кількості джерел повідомлень велику кількість їх одержувачів) та професійні (зв'язкові), у яких кількість джерел та одержувачів повідомлень одного порядку.

До показників ефективності відносяться достовірність та швидкість передачі інформації, а також деякі інші величини.

Узагальнена структурна схема системи зв'язку (рис. 1.4) відбиває найбільш типові перетворення, яким піддається повідомлення у системі зв'язку, вона справедлива будь-яких видів повідомлень.

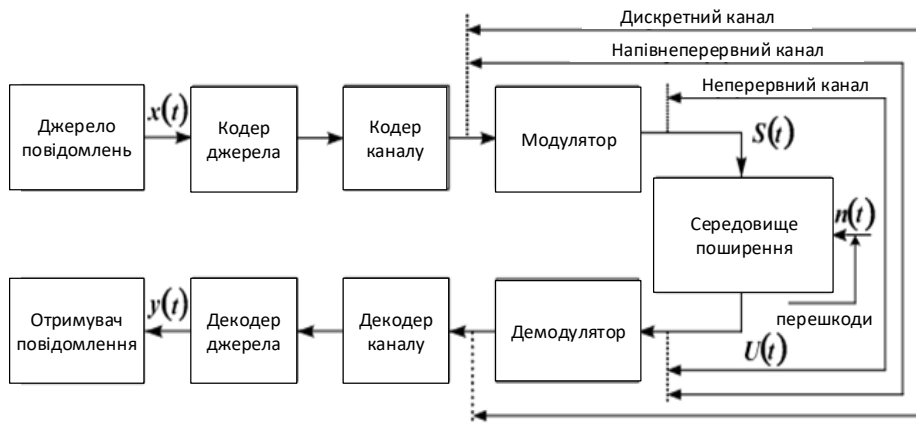


Рисунок 1.4 – Структурна схема системи електричного зв'язку

Розглянемо призначення основних блоків системи зв'язку.

Джерело повідомлень – джерело інформації, що підлягає передачі (людина, докільця тощо.), де повідомлення – це мова, музика, текст, зображення, параметри деяких об'єктів тощо.

Кодер джерела та кодер каналу:

- а) перетворює неелектричне повідомлення в електричний сигнал;
- б) перетворює аналоговий сигнал на цифровий;
- в) здійснює ефективне кодування з метою зменшення необхідної швидкості передачі інформації при заданій якості (усунення надмірності повідомлення);
- г) здійснює завадостійке кодування, що дозволяє поліпшити якість повідомлення, що приймається.

Модулятор – змінює амплітуду, частоту або фазу опорного (несучого) сигналу у відповідності до модулюючого сигналу, що надходить від кодера.

Середовище поширення (лінія зв'язку) – сукупність технічних пристроїв (кабель, двопровідна лінія, оптична лінія зв'язку) або ефір, якими сигнал надходить від передавача до приймача.

Демодулятор – перетворює височастотний модульований сигнал в низькочастотний (сигнал на виході демодулятора, приблизно, відповідає тому, що було на вході модулятора).

Декодер: а) приймає рішення щодо кожної послідовності (1 або 0);

б) декодує кодові комбінації, виправляє частину помилок;

г) перетворює кодові комбінації на повідомлення зручні для одержувача.

Отримувач повідомлення – людина, комп'ютер чи інші технічні пристрої.

З кодуванням не слід плутати шифрування повідомлень. Мета шифрування полягає у запобіганні несанкціонованого вилучення або навмисного зміни інформації. Під час шифрування відбувається заміна відкритого повідомлення шифрограмою (шифр-текстом), а під час розшифрування відбувається зворотнє перетворення. Шифрування виконується до перетворення повідомлення в первинний сигнал або кодову послідовність.

Основною частиною практично всіх систем електров'язку є **канал передачі даних** – це шлях, за допомогою якого дані передаються від одного пристрою до іншого, що складається з передавального і приймального пристрою і лінії зв'язку.

У теорії передачі сигналів канали класифікують за характером сигналів на вході та виході та розрізняють:

- неперервні;
- дискретні;
- дискретно-неперервні канали.

У неперервних каналах сигнали на вході та виході аналогові, неперервні за рівнями; у дискретних каналах – вони відповідно дискретні; а у дискретно-неперервних – сигнали на вході дискретні, але на виході неперервні і навпаки.

Сукупність технічних засобів, включених між модулятором та демодулятором (рис. 1.4), тобто вихідні каскади передавача, передаюча антена, середовище поширення, приймальна антена і лінійна частина приймача, утворюють неперервний канал, тоді як вхідні та вихідні радіосигнали неперервні за своєю природою.

Розглядаючи частину системи між виходом кодера та входом декодера, отримуємо дискретний канал.

Нарешті, частина системи між виходом кодера та входом демодулятора утворює дискретно-неперервний канал.

Канал зв'язку бути фізичним (таким як оптоволокну, мідь або бездротові з'єднання) або логічним (віртуальним шляхом через мережу). В свою чергу фізичні канали поділяються на проводові та безпроводові.

Канали зв'язку класифікують за різними ознаками:

- за призначенням;
- за характером ліній зв'язку;
- по діапазону частот;
- за характером сигналів на вході та виході каналів тощо.

До проводових каналів належать зокрема волоконно-оптичний кабель – для передачі даних у вигляді світлових сигналів та мідний кабель (включаючи коаксіальний кабель, телефонний кабель тощо), що використовується для передачі широкосмугових сигналів у кабельних телевізійних мережах та Інтернет-послуг, сюди ж можна віднести кабелі електроживлення.

Безпроводовий канал: це тип передачі даних, при якому використовуються радіохвилі або інші безпроводні технології для передачі сигналів між вузлами мережі, можуть бути організовані по радіорелейних чи супутникових лініях.

Канали зв'язку можна характеризувати за аналогією з сигналами наступними трьома параметрами:

- часом доступу T_k ,

- шириною смуги пропускання ΔF_k ,
- динамічним діапазоном $D_k = 10 \cdot \lg \frac{P_{k.доп.}}{P_{ш}}$ [дБ],

де $P_{k.доп.}$ – максимально допустима потужність сигналу в каналі, $P_{ш}$ – потужність власних шумів каналу.

Узагальненим параметром каналу є його ємність:

$$V_k = T_k \cdot \Delta F_k \cdot D_k. \quad (1.3)$$

Очевидною необхідною умовою узгодження сигналу та каналу є виконання нерівностей:

$$T_k \geq T_c; \Delta F_k \geq \Delta F_c; D_k \geq D_c; \quad (1.4)$$

$$V_k \geq V_c. \quad (1.5)$$

Однак нерівність (1.5) може виконуватися й тоді, коли одна або дві з нерівностей (1.4) не виконані. Це означає, що можна компенсувати зміну тривалості сигналу шириною спектру і навпаки. Записаний на плівку телефонний сигнал, який має ширину спектру 3 кГц, необхідно передати через канал, смуга пропускання якого 300 Гц.

Це можливо здійснити, відтворюючи сигнал зі швидкістю, до 10 разів меншою за ту, з якою він був записаний. При цьому всі частоти вихідного сигналу зменшаться в 10 разів і в стільки ж раз збільшиться час передачі. Прийнятий сигнал також записується на плівку, а потім, відтворивши його зі швидкістю, в 10 разів більшою, можна відновити вихідний сигнал. Аналогічно можна передати сигнал швидше, якщо смуга пропускання каналу ширше за спектр сигналу.

Значно більший інтерес має можливість обміну динамічного діапазону на смугу пропускання. Так, використовуючи широкосмугові завадостійкі види модуляції, можна передати повідомлення з динамічним діапазоном, наприклад 60 дБ, по каналу, у якому сигнал перевищує заваду лише на 20 дБ. При цьому використовується смуга пропускання каналу в кілька разів більш широка, ніж спектр повідомлення.

Відношення $r_k = V_k / V_c$ називається резервом ємності каналу зв'язку.

Тема 2. Математичні моделі сигналів

Математичне подання сигналу. Динамічне, геометричне подання сигналів, подання сигналів у виді ортогональних функцій. Енергетичні характеристики неперервних детермінованих сигналів. Комплексне представлення сигналів і перехід. Поняття аналітичного сигналу. Огинаюча, миттєва фаза і миттєва частота.

2.1 Математичний опис сигналу

Математичний опис та подання сигналів дозволяє створити математичну модель сигналу.

Математичним описом детермінованого сигналу є детермінована функція часу $S(t)$. Це означає, що будь-якому моменту t_i відповідає певне значення функції $S(t_i)$.

Математичне опис випадкових (чи нерегулярних) сигналів здійснюється з допомогою випадкових функцій. Випадковими сигналами є, наприклад, напруги або струми, що відповідають мовленню, музиці, послідовності телеграфних знаків, тощо. Для випадкової функції значення при фіксованому аргументі t_i – випадкова величина.

В електров'язку знаходять застосування гармонійні та імпульсні сигнали.

Гармонійний сигнал (рис. 2.1) записується у вигляді:

$$S(t) = A_0 \cos(\omega t + \varphi_0), \quad (2.1)$$

де A_0 – максимальне значення (амплітуда); $\omega = 2\pi f$ – кутова частота; $f = 1/T$ – циклічна частота; $\varphi_0 = 2\pi t_i/T$ – початкова фаза.

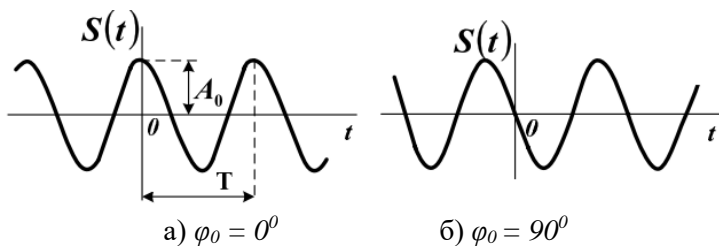


Рисунок 2.1 – Гармонійний сигнал

Імпульсними є сигнали, відмінні від нуля протягом обмеженого часу. Ці сигнали існують лише межах кінцевого відрізка (t_1, t_2) . При цьому розрізняють відеоімпульси (рис. 2.2 а) і радіоімпульси (рис. 2.2 б). Якщо $s_B(t)$ – відеоімпульс, то відповідний радіоімпульс описується виразом: $S_P(t) = s_B(t) \cos(\omega t + \varphi_0)$ (частота ω і початкова фаза φ_0 можуть бути довільними). У

радіоімпульсі $s_B(t)$ називається **огинаючою**, а функція $\cos(\omega t + \varphi_0)$ – **заповненням**. Параметрами відеоімпульсу прийнято вважати його амплітуду A_0 , тривалість t_s і тривалість фронту t_f , тривалість спаду t_c . Походження терміну «відеоімпульс» пов'язане з тим, що вперше такі імпульси почали застосовувати для опису сигналів телебачення.

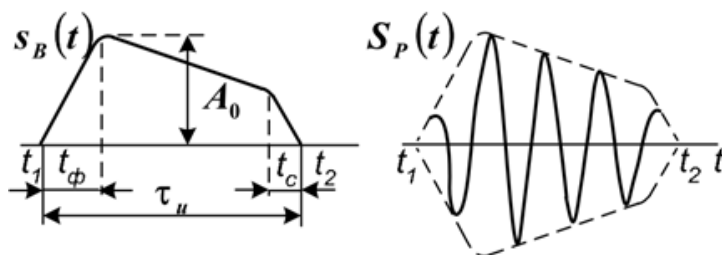


Рисунок 2.2 – Імпульсні сигнали: а) відеоімпульс; б) радіоімпульс

У електровз'язку найбільше застосування знаходять поодинокі імпульси чи періодичні послідовності імпульсів, форма яких наближається до прямокутної. Для періодичної послідовності імпульсів вводиться поняття **шпаруватості**, що визначається як відношення періоду до тривалості імпульсу: $Q = T/\tau_i$, T – період (часовий інтервал), так що $S(t_i + kT) = S(t_i)$, $k = 0, \pm 1, \pm 2, \dots$

2.2 Динамічне подання сигналів

Багато завдань теорії зв'язку та радіотехніки, наприклад, обчислення відгуку фізичної системи на відомий вхідний вплив, вимагають специфічної форми подання сигналів.

Спосіб отримання таких моделей сигналів полягає у наступному. Реальний сигнал приблизно представляється сумою деяких елементарних сигналів, що виникають у послідовні моменти часу. Якщо тепер спрямувати до нуля тривалість окремих елементарних сигналів, то, природно, межі буде отримано точне уявлення вихідного сигналу. У літературі цей спосіб опису сигналу отримав назву **динамічного опису**, підкреслюючи процес, що розвивається в часі.

Широке застосування знайшли два способи динамічного опису. Згідно з першим з них як елементарні сигнали використовуються ступінчасті функції, що виникають через рівні проміжки часу Δ .

Як такі функції використовуються функції включення (рис. 2.3.) або функції Хевісайда $\sigma(t)$.

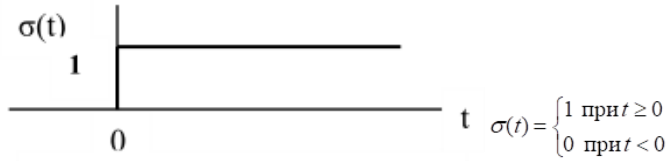


Рисунок 2.3 – Функція увімкнення

Інша можливість представлення сигналу полягає у використанні стандартних прямокутних функцій тривалістю Δ . На рис. 2.4 показані можливі способи представлення сигналів.

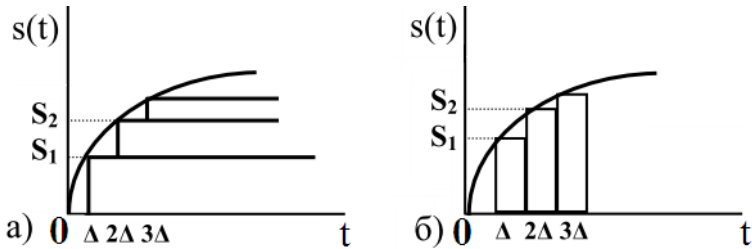


Рисунок 2.4 – Динамічне подання сигналів

Як видно (рис. 2.4 а), поточне значення сигналу при будь-якому t дорівнює сумі ступінчастих функцій:

$$s(t) \approx s_0 \sigma(t) + (s_1 - s_0) \sigma(t - \Delta) + (s_2 - s_1) \sigma(t - 2\Delta) + \dots \quad (2.2)$$

У разі подання аналогового сигналу сумою прямокутних імпульсів, що примикають один до одного, елементарний імпульс з номером k подається у вигляді:

$$u_k(t) = s_k [\sigma(t - t_k) - \sigma(t - t_k - \Delta)]. \quad (2.3)$$

Тоді вихідний сигнал є сумою елементарних імпульсів:

$$s(t) = \sum_{k=-\infty}^{+\infty} u_k(t). \quad (2.4)$$

Важливе значення при динамічному поданні сигналу відіграє й інша функція, яка називається дельта-функцією $\delta(t)$ або функцією Дірака. Такою функцією називається імпульсний сигнал, площа якого, наприклад, $A_m \cdot \tau$ дорівнює 1, причому тривалість імпульсу прагне до нуля, а амплітуда імпульсу A_m прагне до нескінченності.

Якщо у виразі (2.4) Δ спрямувати до нуля, то отримаємо формулу динамічного подання сигналу

$$s(t) = \int_{-\infty}^{+\infty} s(\tau) \delta(t - \tau) d\tau. \quad (2.5)$$

Таким чином, якщо безперервну функцію помножити на дельта-функцію і добуток проінтегрувати за часом, то результат дорівнюватиме значенню функції в тій точці, де існує δ -функція. У цьому полягає фільтруюча властивість дельта-функції.

2.3 Геометричне подання сигналів

Ідеї функціонального аналізу дали можливість створити теорію сигналів, основу якої лежить уявлення сигналу як вектора у деякому нескінченномірному просторі.

Якщо є певна сукупність сигналів $s_1(t), s_2(t), \dots$, що мають деякі загальні властивості, їх можна об'єднати в множини сигналів $M = \{s_1(t), s_2(t), \dots\}$.

Завдання розкладання сигналу складної форми на найпростіші складові подібна до розкладання звичайного вектора x тривимірного простору на його складові по координатному базису одиничних ортогональних векторів i, j, k . Таке уявлення можна записати як:

$$x = x_1 i + x_2 j + x_3 k \quad (2.6)$$

Коефіцієнти x_1, x_2, x_3 є проєкції вектора x на координатні осі i, j, k і називаються координатами вектора x . Інакше висловлюючись, вектор x у тривимірному просторі повністю визначається сукупністю його координат $x = (x_1, x_2, x_3)$.

Щоб перейти до узагальнення поняття вектора тривимірного простору для n -вимірного простору, функцію $x(t)$ за аналогією з (2.6) можна представити у вигляді суми:

$$x(t) = \sum_{i=0}^{n-1} x_i \cdot \psi_i(t), \quad (2.7)$$

де ψ_i – елементарні базові функції.

Множина векторів $\{\psi_i\}$ називається лінійно незалежною (**базисом**), якщо умова $\sum_{i=0}^{n-1} x_i \cdot \psi_i(t) = 0$ виконується лише тоді, коли всі $x_i = 0$. Лінійно незалежні вектори $\{\psi_i\}$ можна розглядати як координатні осі простору.

Метричним називається лінійний простір, де визначено відстань між елементами (векторами) простору (метрика), тобто. кожній парі елементів, скажімо, x і y може бути поставлене у відповідність деяке дійсне невід'ємне число $d(x,y)$ і спосіб, відповідно до якого знаходиться це число.

Серед лінійних метричних просторів важливе місце посідають **нормовані простори**.

Для цього вводиться нове поняття, що відповідає довжині вектора. У математиці довжину вектора називають його нормою. Простір сигналів називається нормованим, якщо кожному вектору $s(t)$ однозначно відповідає число $\|s\|$, що називається нормою. Для дійсних аналогових сигналів теорії сигналів норму сигналу вводять у вигляді:

$$\|s\| = \sqrt{\int_{-\infty}^{+\infty} s^2(t) dt}. \quad (2.8)$$

Для комплексних сигналів норма сигналу представляється:

$$\|s\| = \sqrt{\int_{-\infty}^{+\infty} s(t) s^*(t) dt}. \quad (2.9)$$

Квадрат норми називається енергією сигналу E_s :

$$E_s = \|s\|^2 = \int_{-\infty}^{+\infty} s^2(t) dt. \quad (2.10)$$

Така енергія сигналу виділяється на резисторі із опором 1 Ом. Вираз (2.10) є дуже зручним, оскільки відпадає необхідність розшифровувати розмірність сигналу, тобто сигнал заданий як струм чи напруга.

2.4 Подання сигналів у вигляді рядів ортогональних функцій

При надсиланні повідомлень одночасно існує велика кількість різноманітних сигналів. Припустимо, що є два сигнали s_i та s_j і визначимо енергію сумарного сигналу:

$$E = \int_{-\infty}^{+\infty} (s_i + s_j)^2 dt = E_i + E_j + 2 \int_{-\infty}^{+\infty} s_i \cdot s_j dy.$$

Видно, що на відміну від самих сигналів, їхня енергія неадитивна. Енергія сумарного сигналу містить так звану взаємну енергію, яка визначається як скалярний добуток двох дійсних сигналів:

$$E_{ij} = 2 \int_{-\infty}^{+\infty} s_i \cdot s_j dy. \quad (2.11)$$

Якщо взаємна енергія сигналів s_i і s_j дорівнює нулю, вони називаються **ортогональними**.

Для дослідження різних властивостей повідомлень, сигналів та перешкод зручно використовувати розкладання цих процесів у ряди. Будь-який процес (з деякими математичними обмеженнями) можна подати у вигляді ряду:

$$x(t) = \sum_{k=-\infty}^{\infty} C_k \phi_k(t) \quad (2.12)$$

$\phi_k(t)$ – ортогональні функції, тобто:

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \phi_k(t) \phi_n(t) dt = \begin{cases} E_k, & k = n \\ 0, & k \neq n \end{cases}$$

C_k – коефіцієнти розкладання, E_k – енергія ортогональних функцій.

$$C_k = \frac{1}{E_k} \int_{-T}^T x(t) \phi_k(t) dt$$

Якщо вибрати як ортогональні функції:

$$\phi_k(t) = \begin{cases} \cos k\Omega t \\ \sin k\Omega t \\ e^{jk\Omega t} \end{cases}$$

то цей ряд (2.12) називається рядом Фур'є (тригонометричний ряд).

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\Omega t + b_k \sin k\Omega t) \quad (2.13)$$

$$a_0 = \frac{2}{T} \int_0^T x(t) dt; \quad a_k = \frac{2}{T} \int_0^T x(t) \cos k\Omega t dt; \quad b_k = \frac{2}{T} \int_0^T x(t) \sin k\Omega t dt$$

$\Omega = 2\pi/T$ – частота першої гармоніки, що визначається періодом T (T – період функції $x(t)$).

Друга, еквівалентна формула записи тригонометричного ряду:

$$x(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} A_k \cos(k\Omega t - \phi_k), \quad (2.14)$$

$$A_k = \sqrt{a_k^2 + b_k^2}, \quad \phi_k = \arctg \frac{b_k}{a_k}, \quad a_k = A_k \cos \phi_k, \quad b_k = A_k \sin \phi_k.$$

На рис. 2.5 наведені графіки, представлення періодичної послідовності прямокутних імпульсів $s(t)$ кінцевим числом доданків ($k = 5$) ряду Фур'є. \

Для спектрального уявлення послідовності прямокутних імпульсів початок відліку доцільно брати у середині імпульсу.

Справді, у разі у розкладанні залишаться лише косинусоїдальні складові, оскільки інтеграли від непарних функцій у період дорівнюють нулю $b_k = 0$.

Для парного сигналу $x(t) = x(-t)$ коефіцієнти $a_k \neq 0$, $b_k = 0$.

Для непарного сигналу $x(t) = -x(-t)$ коефіцієнти $a_k = 0$, $b_k \neq 0$.

Для функції $s(t)$ (рис. 2.5) розклад має вигляд:

$$s(t) = \frac{A_m}{2} + \frac{2A_m}{\pi} + \frac{2A_m}{3\pi} \sin 3\Omega_1 t + \frac{2A_m}{5\pi} \sin 5\Omega_1 t + \dots \quad (2.15)$$

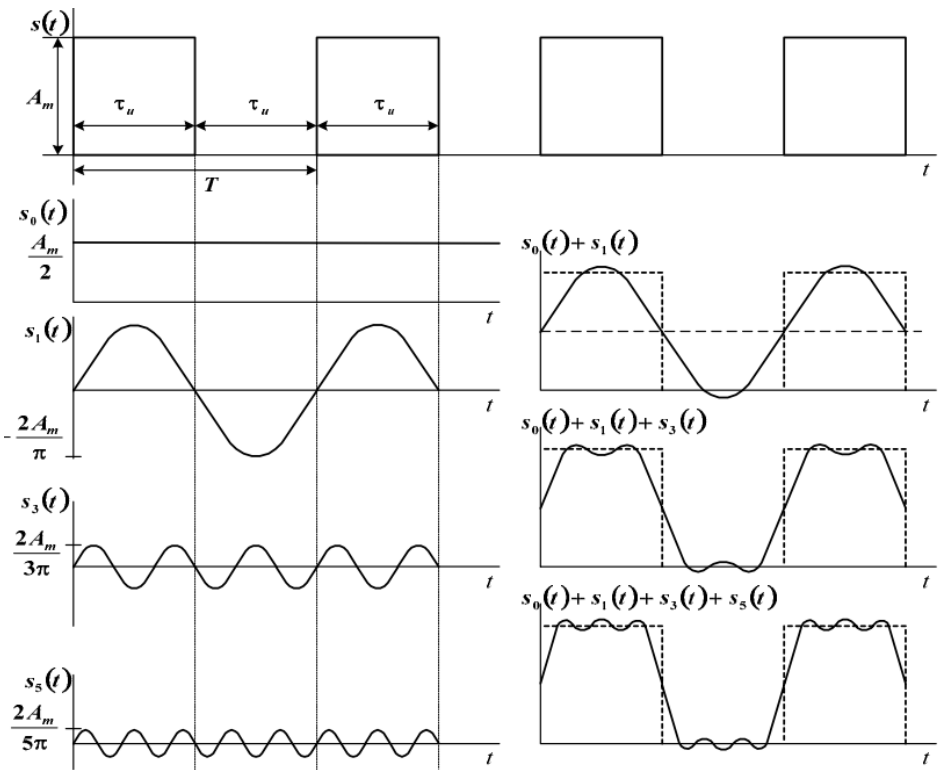


Рисунок 2.5 – Апроксимація прямокутних імпульсів сумою гармонік

Періодична послідовність прямокутних імпульсів $s(t)$ представляється як результат додавання постійної складової $A_m/2$ і синусоїдальних сигналів з частотами $F_1, 3F_1, 5F_1, \dots$, причому $F_1 = 1/T$.

Легко помітити, що графік суми двох перших доданків розкладання (2.15) відтворює форму графіка функції $s(t)$ дуже грубо, тільки в основних рисах. Зі збільшенням числа гармонік, що додаються, точність представлення $s(t)$ зростає. Сукупність всіх гармонійних складових розкладання функції ряд Фур'є називається **спектром функції**.

Наявність окремих гармонійних складових спектру та величини їх амплітуд можна наочно показати за допомогою **спектральної діаграми** (рис. 2.6), у якій горизонтальна вісь служить віссю частот, а вертикальна – віссю амплітуд.

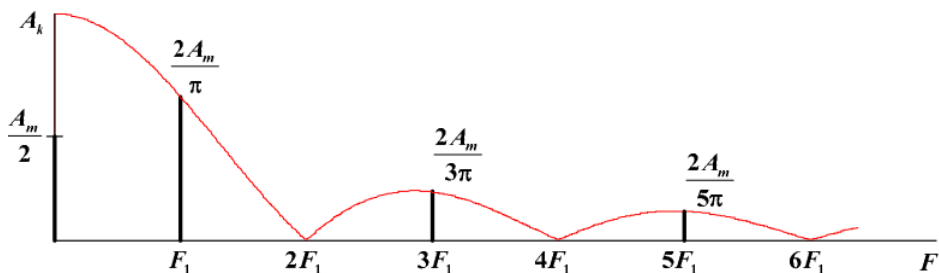


Рисунок 2.6 – Спектр амплітуд прямокутних імпульсів

У точках осі частот $F_1, 3F_1, 5F_1, \dots$ відображаються амплітуди відповідних гармонійних складових розкладання функції. За амплітудним спектром можна оцінити відсотковий зміст гармонік, наявність та рівні окремих гармонійних складових спектру.

2.5 Комплексне представлення сигналів та завад

Подання детермінованих сигналів рядами ортогональних функцій виявляється корисним при аналізі проходження сигналів через лінійні радіотехнічні пристрої. При аналізі нелінійних перетворень сигналів і, зокрема, модуляції та демодуляції, потрібен інший підхід. Цей підхід ґрунтується на понятті аналітичного сигналу.

Багато формул гармонічного аналізу записуються значно простіше і деякі завдання вирішуються легше, якщо використовувати в якості елементарних функцій експоненційні функції уявного аргументу, наприклад, комплексна огинаюча легко відокремлюється від множника з несучою частотою при вираженні сигналу в комплексній формі.

$$A \cdot \cos[j(\omega t + \varphi)] = [A \cdot \exp(j \varphi)] \cdot \exp(\omega t).$$

Якщо розкласти косинус суми за формулою Ейлера, то:

$$\cos(\omega t + \varphi) = 1/2[\exp\{j(\omega t + \varphi)\} + \exp\{-j(\omega t + \varphi)\}]. \quad (2.16)$$

Цьому запису можна дати геометричне трактування, користуючись представленням комплексних чисел як точок чи векторів на площині.

Вираз $\exp\{j(\omega t + \varphi)\}$ представляє собою вектор одиничної довжини, проведений під кутом $\omega t + \varphi$ до дійсної осі (рис. 2.7). При зміні часу t цей вектор одиничної довжини змінює положення, обертаючись в додатньому напрямку з кутовою швидкістю ω .

Зобразити синусоїду у формі (2.16), це означає уявити її сумою двох векторів, довжина кожного з яких дорівнює $1/2$, розташованих у будь-який момент часу симетрично щодо дійсної осі, які обертаються в різних напрямках з кутовими швидкостями ω і $-\omega$ (рис. 2.7).

Геометрична сума векторів завжди збігається у напрямку з дійсною віссю і представляє дійсну функцію $\cos(\omega t + \varphi)$.

При поданні косинусоїди у вигляді $\cos(\omega t + \varphi) = \text{Re}\{\exp\{j(\omega t + \varphi)\}\}$ можна обмежитися одним вектором, що обертається в додатньому напрямку і уявити косинусоїду його проекцією на дійсну вісь.

Довжина вектора представляє амплітуду косинусоїди, а кут, що утворюється ним в даний момент з дійсною віссю - повну фазу $(t + \varphi)$.

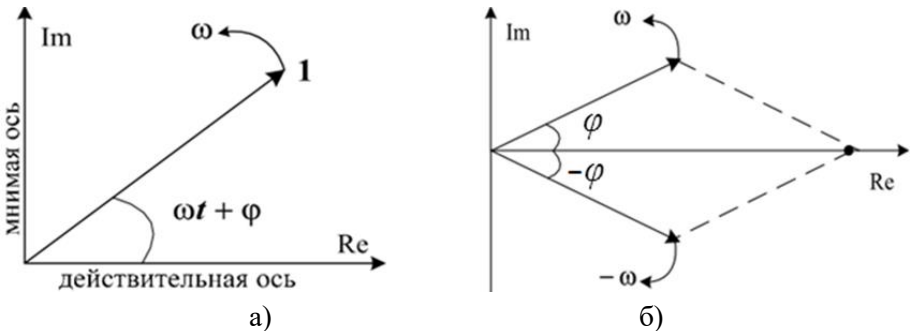


Рисунок 2.7 – а) Геометричне трактування експоненційної функції уявного аргументу; б) Експонентне трактування уявлення елементарної функції

Значну кількість сигналів, що застосовуються в системах електров'язку, можна представляти у вигляді:

$$s(t) = A(t) \cdot \cos[\omega t + \varphi(t)], \quad (2.17)$$

тобто як «квазігармонійну» функцію зі змінними «амплітудою» та «початковою фазою». Такий сигнал можна інтерпретувати геометрично як проекцію на дійсну вісь вектора, що обертається, але при цьому змінює свою довжину і кутову швидкість. Для опису властивостей сигналу, представленого у формі (2.17), вводять поняття **комплексного аналітичного сигналу**.

Комплексний сигнал можна подати у формі:

$$s_a(t) = A(t) \cdot e^{j\phi(t)} = A(t) \cdot (\cos \phi(t) + j \sin \phi(t)), \quad (2.18)$$

де $A(t) = \sqrt{s^2(t) + s^{*2}(t)}$ називають **огиноючою** сигналу,

$$\text{а } \phi(t) = \text{Arg } s_a(t) = \text{arctg} \frac{s^*(t)}{s(t)} + \begin{cases} 0, & s^*(t) \geq 0 \\ \pi, & s^*(t) < 0 \end{cases} \quad - \text{ миттєвою фазою}$$

сигналу.

$$\text{Тут } s(t) = A(t) \cdot \cos \phi(t); \quad s^*(t) = A(t) \cdot \sin \phi(t).$$

Функція $\phi(t)$ називається **миттєвою фазою** сигналу.

Похідна від миттєвої фази сигналу за часом називається **миттєвою частотою** сигналу:

$$\omega(t) = \frac{d\phi(t)}{dt} = \frac{1}{2\pi} \cdot \frac{s^{*'}(t) \cdot s(t) - s'(t) \cdot s^*(t)}{s^2(t) + s^{*2}(t)}. \quad (2.19)$$

Наприклад, для гармонійного сигналу:

$$\omega(t) = \frac{\cos \omega_0 t \cdot \cos \omega_0 t \cdot \omega_0 \cdot A^2 + \sin \omega_0 t \cdot \sin \omega_0 t \cdot \omega_0 \cdot A^2}{A^2 \cdot t + A^2 \cdot \sin^2 \omega_0 t} = \omega_0 = \text{const.}$$

Загалом миттєва частота змінюється у часі.

Відмітимо, що $A(t) \geq s(t)$, причому рівність досягається в моменти часу, коли $s^*(t) = 0$. У цих точках похідна $A(t)$ збігається з похідною сигналу $s(t)$:

$$A'(t) = \frac{s(t) \cdot s'(t) + s^*(t) \cdot s^{*'}(t)}{\sqrt{s^2(t) + s^{*2}(t)}}; \quad A'(t) = s'(t) \text{ нпу } s^*(t) = 0. \quad (2.20)$$

Отже, при $s^*(t) = 0$, огиноюча $A(t)$ дотикається сигналу $s(t)$.

Функція $\cos(\phi(t))$ називається **високочастотним заповненням** сигналу.

Процес формування сигналу на основі огиноючої $A(t)$ та фази $\phi(t)$ показаний на рис. 2.8.

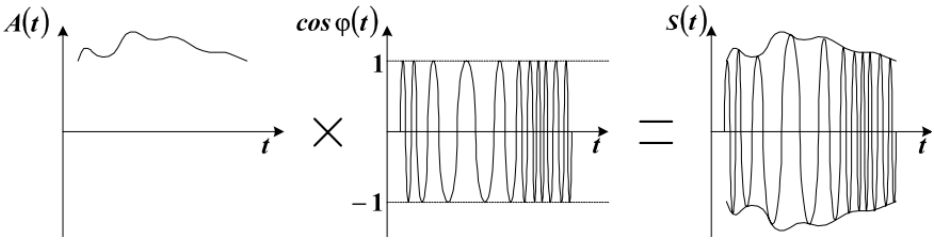


Рисунок 2.8 – Часове подання огинаючої та високочастотного заповнення

Якщо миттєва частота коливається навколо середнього значення ω_{cp} , можна записати:

$$\begin{aligned} \phi(t) &= \omega_{cp} t + \Theta(t); \\ s(t) &= A(t) \cos[\omega_{cp} t + \Theta(t)], \end{aligned} \quad (2.21)$$

де $\Theta(t)$ – називається миттєвою початковою фазою сигналу.

Вираз (2.21) є зручним для опису **вузькосмугових** сигналів. У цьому випадку основна частина спектру амплітуд зосереджена відносно вузькій, в порівнянні з $A(t)$ і $\phi(t)$, смузі частот. При цьому $A(t)$ і $\phi(t)$ змінюються повільно проти $\cos(\omega_{cp} t)$. Такі сигнали називаються **квazігармонійними**. У випадкових сигналів і завад $A(t)$, $\phi(t)$, $\omega(t)$, $\omega_{cp}(t)$ і $\Theta(t)$ є випадковими функціями часу.

Тема 3. Спектральні характеристики сигналів.

Спектральне представлення періодичних і неперіодичних сигналів. Основні властивості перетворення Фур'є. Теорема Котельнікова. Відновлення неперервного сигналу з відліків. Похибки дискретизації та відновлення неперервних сигналів.

3.1 Спектральне подання періодичних сигналів

Як відомо, розкладання періодичного сигналу по базису тригонометричних функцій – це розкладання його до ряду Фур'є. Розкладання сигналу до ряду Фур'є називається **спектром сигналу**.

У загальному випадку періодичний сигнал містить незалежну від часу постійну складову і нескінченний набір гармонійних коливань, чи гармонік, з частотами, кратними основній частоті послідовності. Графічне зображення коефіцієнтів Фур'є ряду для конкретного сигналу називається **спектральною діаграмою**. По горизонтальній осі відкладаються частоти гармонік, а по вертикалі – амплітуди (амплітудна діаграма) чи початкові фази (фазова діаграма).

При розкладанні у комплексний ряд Фур'є:

$$s(t) = \sum_{k=-\infty}^{\infty} C_k \cdot \exp(jk\Omega t), \quad C_k = \frac{1}{T} \int_0^T s(t) e^{-jk\Omega t} dt \quad (3.1)$$

Спектр сигналу містить компоненти на від'ємній півосі частот, причому $C_{-k} = C_k^*$ (* позначено комплексно-зв'язане число).

Між коефіцієнтами комплексного та тригонометричного ряду існує зв'язок:

$$A_k = 2|C_k|, \quad \phi_k = \arg C_k. \quad (3.2)$$

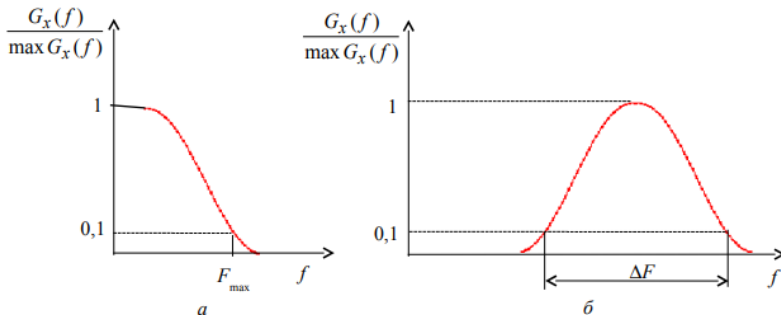


Рисунок 3.1 – Ширина спектру сигналу

Шириною спектру сигналу називають протяжність області додатних частот, поза якою значення функції не перевищують значення $0,1 \max\{G_X(\omega)\}$. Якщо спектр примикає до нуля, то ширину спектра позначають величиною F_{max} (рис. 3.1, а), а якщо спектр смуговий, то ширину спектра позначають величиною ΔF (рис. 3.1, б).

Як приклад розрахуємо спектр періодичної послідовності прямокутних імпульсів з амплітудою A (Часова діаграма періодичної послідовності імпульсів показана на рис. 3.1. а). Визначимо коефіцієнти розкладання в ряд Фур'є C_k :

$$a_k = \frac{2}{T} \int_{-\tau/2}^{\tau/2} A \cos k\Omega t dt = \frac{2A}{T} \frac{\sin k\Omega t}{k\Omega} = \frac{4A}{k\Omega T} \sin \frac{k\Omega \tau}{2}$$

$$b_k = \frac{2}{T} \int_{-\tau/2}^{\tau/2} A \cdot \sin k\Omega t dt = 0$$

Нехай $T = 2\tau$, тоді коефіцієнти a_k рівні:

$$a_0 = A, a_k = 2A / k\pi (\sin k\pi/2), \text{ при } k > 0.$$

Спектр цієї послідовності дискретний і показаний на рис. 3.2.

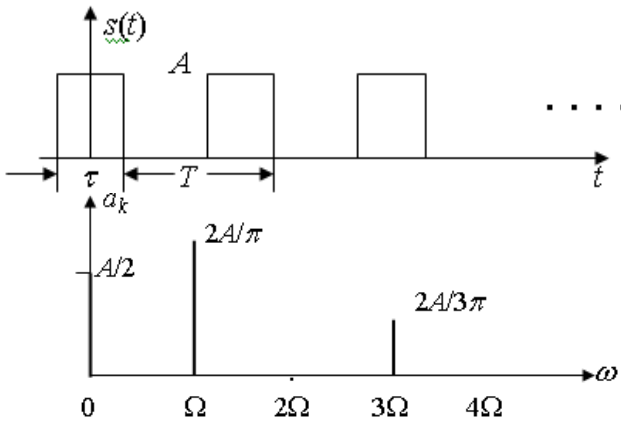


Рисунок 3.2 – а) Періодична послідовність прямокутних імпульсів;
 б) спектр послідовності

Ширина спектру сигналу рівна, в даному випадку, $\Delta F_c = 2\pi/\tau$.

3.2. Спектральне подання неперіодичних сигналів

Для спектрального подання неперіодичних (імпульсних) сигналів $s(t)$, заданих на кінцевому інтервалі (t_1, t_2) (рис. 3.3), безпосередньо скористатися рядом Фур'є не можна. Для гармонійного розкладання сигналу подумки доповнюють його такими самими імпульсними сигналами до періодичного з деяким інтервалом T (рис. 3.3).

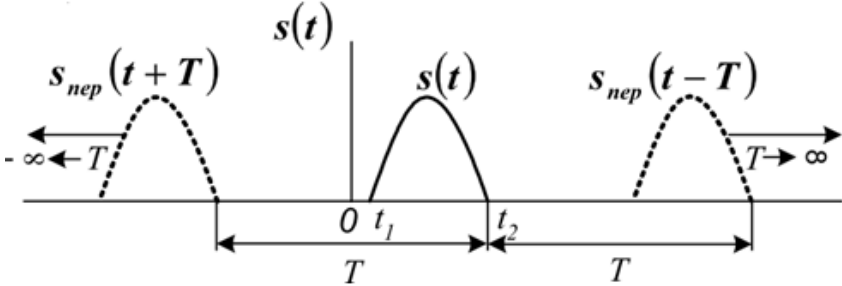


Рисунок 3.3 – Імпульсний сигнал $s(t)$ та його періодичне продовження $s_{nep}(t+kT)$

Для того щоб поза штучно введеним інтервалом вихідний сигнал дорівнював нулю, необхідно збільшити період повторення імпульсів.

У межі, при збільшенні періоду $\infty \rightarrow T$, всі імпульси підуть право і вліво в нескінченність і періодична послідовність знову стане одиночним імпульсом.

Для обчислення спектра зручна симетрична комплексна форма ряду Фур'є, але в ньому замість суми буде інтеграл з нескінченними межами (перетворення Фур'є):

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(j\Omega) \exp(j\Omega t) d\Omega. \quad (3.3)$$

При такому граничному переході основна частота сигналу $\Omega = 2\pi/T$ нескінченно збільшується кількість спектральних складових, частоти сусідніх гармонік $k\Omega$ та $(k+1)\Omega$ стають нерозрізнятими, а спектр буде суцільним.

Функція $G(j\Omega)$ називається **спектральною густиною (щільністю) сигналу** $x(t)$.

Функції $G(j\Omega)$ і $s(t)$ є дві математичні моделі одного й того ж фізичного процесу: одна відображає частотний склад сигналу, інша описує зміну сигналу з часом.

Спектральна густина сигналу визначається з використанням прямого перетворення Фур'є:

$$G(j\Omega) = \int_{-\infty}^{\infty} s(t) \exp(-j\Omega t) dt. \quad (3.4)$$

Таким чином, формули (3.3) і (3.4) називаються відповідно зворотним і прямим перетвореннями Фур'є. Він дає взаємозв'язок між сигналом $s(t)$ та його комплексною спектральною щільністю $G(j\Omega)$.

Для одиночного прямокутного імпульсу з амплітудою A та тривалістю τ на рис. 3.4 отримаємо спектр $S(j\Omega)$ на рис. 3.5:

$$S(j\Omega) = \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} A e^{-j\Omega t} dt = \frac{A}{-j\Omega} \left[\exp\left(-\frac{j\Omega\tau}{2}\right) - \exp\left(\frac{j\Omega\tau}{2}\right) \right]$$

Цей вираз з урахуванням формули Ейлера можна переписати у вигляді:

$$\sin a = \frac{\exp(ja) - \exp(-ja)}{2j}, \quad S(j\Omega) = A \tau \frac{\sin\left(\frac{\Omega\tau}{2}\right)}{\left(\frac{\Omega\tau}{2}\right)} \quad (3.5)$$

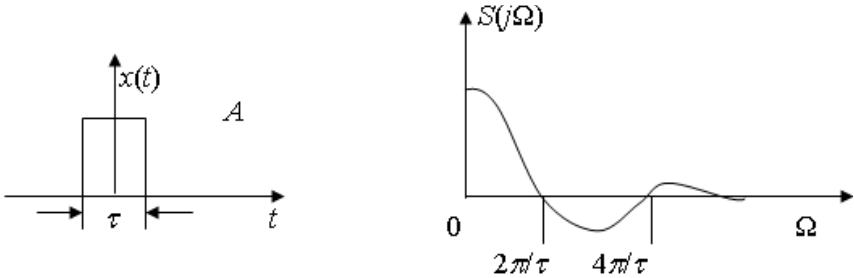


Рисунок 3.4 – Поодинокий прямокутний імпульс і його спектр

Спектр неперіодичного сигналу суцільний, нескінченний, ширина спектру визначається тривалістю сигналу і приблизно дорівнює $\Delta F_e \approx 2\pi/\tau$.

3.3 Основні властивості перетворення Фур'є.

Властивістю перетворень Фур'є визначається взаємна відповідність трансформації сигналів та їх спектрів.

1) Лінійність.

Перетворення Фур'є належить до лінійних інтегральних операцій, тобто спектр суми сигналів дорівнює сумі спектрів цих сигналів.

$$\sum_n a_n s_n(t) \Leftrightarrow \sum_n a_n s_n(\omega) \quad (3.6)$$

Приклад наведено на рис. 3.5: $s_0(k) = s_1(k) + s_2(k)$, $S_1(\omega) + S_2(\omega) = S_0(\omega)$.

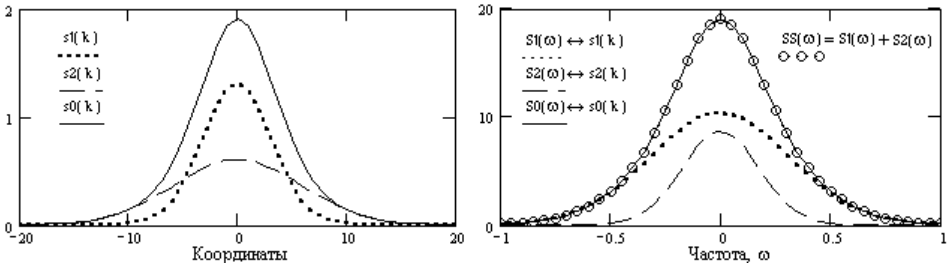


Рисунок 3.5 – Сигнали та їх спектри

2) **Властивості парності** перетворення визначаються косинусними (парними, дійсними) та синусними (непарними, уявними) частинами розкладання та подібністю прямого та зворотного перетворень.

На рис. 3.6. наведено приклади, що пояснюють властивості парності перетворення. Сигнал $s_1(k)$ є парним, $s_1(k) = s_1(-k)$, і має тільки дійсний парний спектр (уявна частина спектральної функції представлена нульовими значеннями). Сигнал $s_2(k) = -s_2(-k)$ непарний і має уявний непарний спектр, а нульовими значеннями представлена його дійсна частина. Сигнал $s_3(k)$ утворений сумою сигналів $s_1(k)$ та $s_2(k)$. Відповідно, спектральна функція сигналу представлена і дійсною парною частиною (що належить $s_1(k)$), і уявною непарною частиною (що належить $s_2(k)$). При зворотному перетворенні Фур'є окремо дійсної і уявної частини спектра $S_3(\omega)$, так само як і будь-яких інших комплексних спектрів, будуть окремо відновлені парна та непарна частини вихідного сигналу.

Зауважимо, що довільний вихідний сигнал може бути заданий в односторонньому варіанті (в інтервалі $0-T$), але парна та непарна частини цього сигналу займають інтервал від $-T$ до T , при цьому на лівій половині числової осі (від $-T$ до 0) ці два сигнали компенсують один одного, даючи нульові значення.

Сигнал $s(t)$, спектр $S(\omega)$. При цьому якщо:

$s(t)$ – парний, то $S(\omega)$ – дійсний, парний;

$s(t)$ – непарний, то $S(\omega)$ – уявний, непарний

$s(t)$ – довільний, то $S(\omega)$ – дійсна частина – парна, а уявна – непарна.

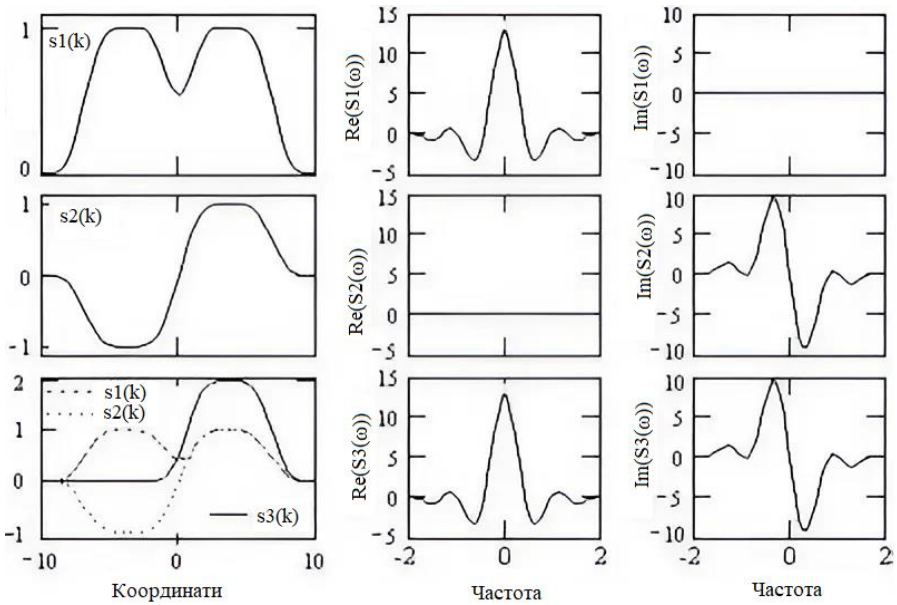


Рисунок 3.6 – Властивості парності перетворення

3) **Зміна аргументу функції** (стиснення або розширення сигналу) призводить до зворотної зміни аргументу її Фур'є-образу і обернено пропорційної зміни його модуля. Дійсно, якщо $s(t) \leftrightarrow S(\omega)$, то при зміні тривалості сигналу зі збереженням його форми (розтягування сигналу по часовій осі), тобто для сигналу з новим аргументом $s(x) = s(a \cdot t)$ при $x = a \cdot t$, отримуємо:

$$s(a \cdot t) \leftrightarrow \int_{-\infty}^{\infty} s(a \cdot t) \exp(-j\omega t) dt = \frac{1}{a} \int_{-\infty}^{\infty} s(x) \exp(-jx \omega/a) dx \quad (3.7')$$

$$s(a \cdot t) \leftrightarrow \left(\frac{1}{a}\right) \cdot S\left(\frac{\omega}{a}\right)$$

Вираз (3.7') дійсний при $a > 0$. При $a < 0$ відбувається дзеркальний поворот сигналу щодо вертикальної осі, а заміна змінної $t = x/a$ викликає перестановку меж інтегрування і, відповідно, зміну знака спектру:

$$s(a \cdot t) \leftrightarrow -(1/a)S(\omega/a). \quad (3.7'')$$

Узагальнена формула зміни аргументу:

$$s(a \cdot t) \leftrightarrow -(1/|a|)S(\omega/a), \quad a \neq 0. \quad (3.7)$$

Якщо під аргументом функції та її спектру розуміти певні фізичні одиниці, наприклад, час – частота, то звідси випливає: що коротший за своєю тривалістю сигнал, то ширший за частотою його спектр, і навпаки. Це можна наочно бачити на рис. 3.5. для сигналів $s_1(k)$ та $s_2(k)$ та їх спектрів $S_1(\omega)$ та $S_2(\omega)$.

Від зміни аргументу функцій слід відрізняти зміну масштабу функцій. Зміна масштабу аргументів змінює лише оцифрування числових осей відображення сигналів та їх спектрів, але не змінює самих сигналів та спектрів. Так, при масштабі осі часу $t = 1$ секунда, масштаб осі частот $f = 1/t = 1$ герц, а при $t = 1$ мксек $f = 1/t = 1$ МГц .

4) Теорема запізнення.

Запізнення (зсув, зсув) сигналу за аргументом функції на інтервал t_0 призводить до зміни фазочастотної функції спектра (фазового кута всіх гармонік) на величину $-\omega t_0$ без зміни модуля (амплітудної функції) спектра. Застосовуючи заміну змінної $t - t_0 = x$, отримуємо:

$$s(t - t_0) \Leftrightarrow \int_{-\infty}^{\infty} s(t - t_0) \exp(-j\omega t) dt = \int_{-\infty}^{\infty} s(x) \exp(-j\omega x) \exp(-j\omega t_0) dx = S(\omega) \exp(-j\omega t_0) \quad (3.8)$$

Цілком очевидно, що амплітуди гармонік сигналу при його зрушенні не повинні змінюватися. З огляду на те, що $|\exp(-j\omega t_0)| = 1$, це випливає і з (3.8):

$$|S(\omega) \exp(-j\omega t_0)| = |S(\omega)|.$$

Фазовий спектр зсувається на $-\omega t_0$ з лінійною залежністю від частоти:

$$S(\omega) \exp(-j\omega t_0) = R(\omega) \exp[j(j(\omega))] \exp(-j\omega t_0) = R(\omega) \exp[j(j(\omega) - \omega t_0)] \quad (3.9)$$

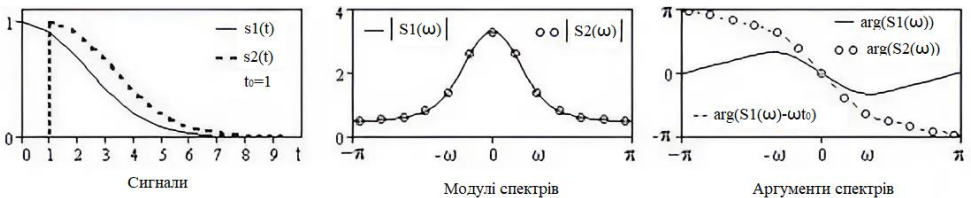


Рисунок 3.7 – Зміна спектра сигналу за його зсуву

Приклад двох однакових сигналів, зсунутих один до одного на $t_0=1$, і відповідних даним сигналів спектрів наведено на рис. 3.7.

5) Спектральна щільність (пряме перетворення Фур'є):

а) гармонійної функції $s(t) = \cos(\omega_0 t)$

$$G(j\Omega) = \pi(\delta(\Omega - \omega_0) + \delta(\Omega + \omega_0))$$

б) радіоімпульсу (властивість зміщення спектру) дозволяє розрахувати спектральну щільність сигналу $s(t)$, помноженого на гармонійне коливання $s_1(t) = s(t)\cos(\omega_0 t + \varphi_0)$.

$$G(j\Omega) = \frac{1}{2}(G_A(j(\Omega - \omega_0))e^{j\varphi} + G_A(j(\Omega + \omega_0))e^{-j\varphi}),$$

де $G_A(j\Omega)$ – спектральна щільність огинаючої $A(t)$.

Отже, виникає розщеплення спектру $G(j\Omega)$ на дві частини максимуми яких виникають на частотах $(+\omega_0)$ і $(-\omega_0)$.

в) δ -функції $s(t) = \delta(t)$:

$$G(j\Omega) = 1.$$

б) Спектри потужності.

Часова функція потужності сигналу у загальній формі визначається виразом: $w(t) = s(t) \cdot s^*(t) = |s(t)|^2$.

Спектральна щільність потужності, відповідно, дорівнює перетворення Фур'є добутку $s(t) \cdot s^*(t)$, яке відобразиться в спектральному поданні згорточкою Фур'є-образів цих функцій:

$$W(f) = \int_{-\infty}^{\infty} S(f) * S^*(f - \nu) d\nu. \quad (3.10)$$

Але для всіх поточних значень частоти f інтеграл у правій частині цього виразу дорівнює добутку $S(f) \cdot S^*(f)$, так як для всіх значень зсуву $\nu \neq 0$ через ортогональність гармонік $S(f)$ і $S^*(f-\nu)$ значення їх добутку дорівнюють нулю. Звідси:

$$W(f) = S(f) * S^*(f) = |S(f)|^2 \quad (3.11)$$

Спектр потужності – дійсна невід'ємна парна функція, яку часто називають **енергетичним спектром**. Спектр потужності, як квадрат модуля спектра сигналу, не містить фазової інформації про частотні складові, а, отже, відновлення сигналу по спектру потужності неможливе. Це також означає, що сигнали з різними фазовими характеристиками можуть мати однакові спектри потужності. Зокрема, зсув сигналу не відображається на його спектрі потужності.

7) **Рівність Парсеваля**. Повна енергія спектру сигналу:

$$E_s = \int_{-\infty}^{\infty} W(f) df = \int_{-\infty}^{\infty} |S(f)|^2 df. \quad (3.12)$$

Так як координатне і частотне представлення по суті лише різні математичні відображення одного і того ж сигналу, то рівною має бути і енергія сигналу у двох представленнях, звідки слідує рівність Парсеваля:

$$\int_{-\infty}^{\infty} |S(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df, \quad (3.13)$$

тобто. енергія сигналу дорівнює інтегралу модуля його частотного спектру – сумі енергій всіх частотних складових сигналу. Аналогічно енергії взаємодії сигналів:

$$\int_{-\infty}^{\infty} x(t)y^*(t)dt = \int_{-\infty}^{\infty} X(f)Y^*(f)df. \quad (3.14)$$

З рівності Парсеваля випливає інваріантність скалярного добутку сигналів та норми щодо перетворення Фур'є:

$$(x(t),y(t)) = (X(f),Y(f)), \|x(t)\|^2 = \|X(f)\|^2.$$

Не слід забувати, що при поданні спектрів у колових частотах (ω) у правій частині наведених рівностей повинен стояти множник $1/2\pi$.

3.4. Сигнали з обмеженим діапазоном. Теорема Котельникова

3.4.1. Розкладання неперервних сигналів в ряд Котельникова.

Як зазначено раніше, будь-які сигнали кінцевої тривалості мають теоретично нескінченно широкий спектр частот. У той самий час частка енергії, передається на високих частотах, дуже мала і нею при розрахунку повної енергії сигналу можна знехтувати. Отже сигнали з обмеженим спектром є зручними математичними моделями реальних сигналів.

В 1933 В. А. Котельников довів, що сигнал $x(t)$ з обмеженою смугою частот, що не має спектральних компонент з частотами, які перевищують значення $\omega_s = 2\pi F_s$, однозначно визначається значеннями, обраними через рівні проміжки часу $\Delta t = \pi/\omega_s = 1/2F_s$.

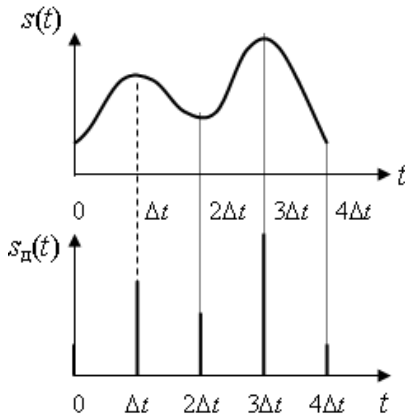


Рисунок 3.8 – Часові діаграми неперервного $s(t)$ та дискретизованого $s_{\text{д}}(t)$ сигналу

Важливо, що не треба безперервно передавати вихідний сигнал $s(t)$, достатньо передавати відліки $s(k\Delta t)$. Це перший крок переходу від неперервного сигналу до цифрового. З погляду математики теорема Котельникова означає подання сигналу як ряду:

$$s(t) = \sum_{k=-\infty}^{\infty} s(k\Delta t) \frac{\sin \omega_g(t - k\Delta t)}{\omega_g(t - k\Delta t)}, \quad (3.15)$$

де $s(k\Delta t)$ – відліки, $(\sin \omega_g(t - k\Delta t)) / \omega_g(t - k\Delta t)$ – функції відліків.

Ряд Котельникова – це розкладання сигналу $s(t)$ в ряд по ортогональних функціях $\phi_k(t)$.

$$\phi_k(t) = (\sin \omega_g(t - k\Delta t)) / \omega_g(t - k\Delta t) \quad (3.16)$$

Теоретично дискретизація здійснюється за допомогою δ -імпульсів.

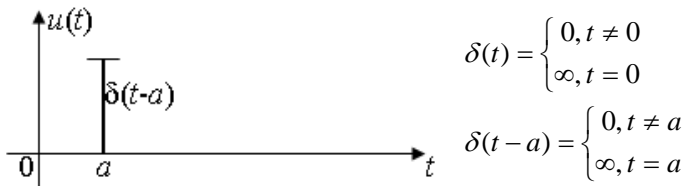


Рисунок 3.8 – Часова діаграма одиночного δ -імпульсу

Спектр одиночного δ -імпульсу отримаємо, використовуючи перетворення Фур'є:

$$\dot{S}_{\delta}(j\Omega) = \int_{-\infty}^{\infty} \delta(t) e^{-j\Omega t} dt = 1$$

Використана «фільтруюча» властивість дельта-функцій:

$$\int_{-\infty}^{\infty} \delta(t-a) f(t) dt = f(a)$$

Отже, спектр одиночного дельта-імпульсу має вигляд:

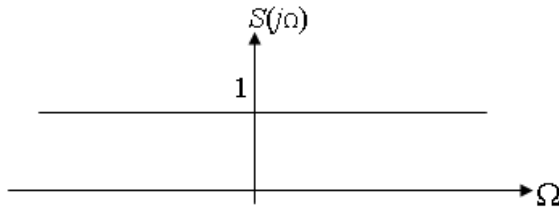


Рисунок 3.9 – Спектр одиночного δ -імпульсу

Щоб отримати відліки функції $s(t)$, перемножимо функцію $s(t)$ на періодичну послідовність дельта-імпульсів з періодом $T = \Delta t$.

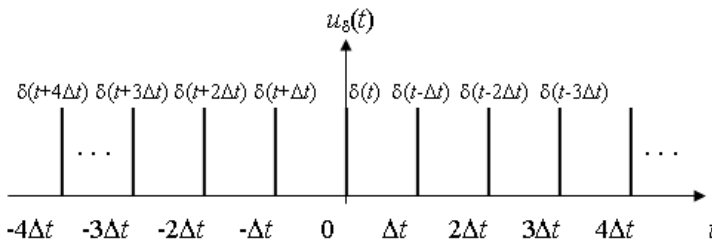


Рисунок 3.10 – Часова діаграма послідовності δ -імпульсів

Оскільки сигнал періодичний, його спектр буде дискретним.

$$u_{\delta}(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\Omega t} = \dots + \frac{1}{\Delta t} e^{-2jk\Omega t} + \frac{1}{\Delta t} e^{-jk\Omega t} + \frac{1}{\Delta t} + \frac{1}{\Delta t} e^{jk\Omega t} + \dots$$

$$C_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \delta(t) e^{-jk\Omega t} dt = \frac{1}{\Delta t}$$

$$\Omega = \frac{2\pi}{T} = \frac{2\pi}{\Delta t} = \frac{2\pi\omega_s}{\pi} = 2\omega_s = \omega_d \quad (3.17)$$

$T = \Delta t$; ω_d – частота дискретизації.

Спектр періодичної послідовності дельта-імпульсів відповідно до формули для $U(t)$ має такий вигляд:

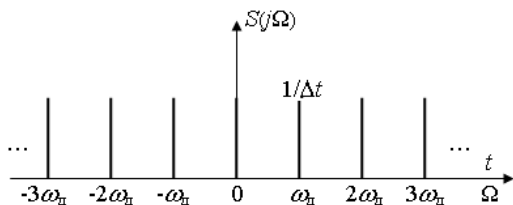


Рисунок 3.11 – Спектр періодичної послідовності δ -імпульсів.

3.4.2 Спектр дискретизованого сигналу

Розглянемо часові діаграми вихідного неперервного та дискретизованого сигналів:

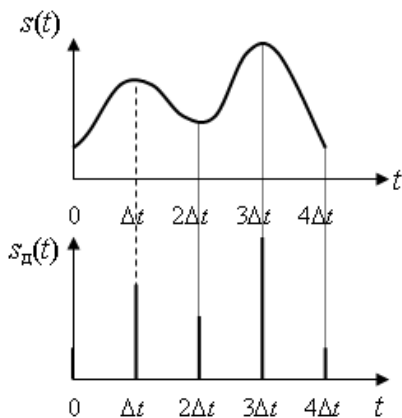


Рисунок 3.12 – Часові діаграми неперервного та дискретизованого сигналів

$s_d(t) = s(t)U_\delta(t)$ – дискретизований сигнал

$s(t)$ – вихідний сигнал.

$U_\delta(t)$ – періодична послідовність δ -імпульсів

Розкладемо періодичну послідовність δ -імпульсів у ряд Фур'є, як ми це робили вище:

$$U_{\delta}(t) = \dots + \frac{1}{\Delta t} e^{-j\omega_{\Delta}t} + \frac{1}{\Delta t} + \frac{1}{\Delta t} e^{j\omega_{\Delta}t} + \dots$$

$$s_{\Delta}(t) = s(t)U_{\delta}(t) = s(t)[\dots + \frac{1}{\Delta t} e^{-j\omega_{\Delta}t} + \frac{1}{\Delta t} + \frac{1}{\Delta t} e^{j\omega_{\Delta}t} + \dots]$$

Знайдемо спектр дискретизованого сигналу.

$$\begin{aligned} \dot{S}_{\Delta}(\Omega) &= \int_{-\infty}^{\infty} s_{\Delta}(t)e^{-j\Omega t} dt = \dots + \frac{1}{\Delta t} \int_{-\infty}^{\infty} s(t)e^{-j(\Omega+\omega_{\Delta})t} dt + \frac{1}{\Delta t} \int_{-\infty}^{\infty} s(t)e^{-j\Omega t} dt + \\ &+ \frac{1}{\Delta t} \int_{-\infty}^{\infty} s(t)e^{-j(\Omega-\omega_{\Delta})t} dt + \dots = \dots + \frac{1}{\Delta t} \dot{S}_x(\Omega + \omega_{\Delta}) + \frac{1}{\Delta t} \dot{S}_x(\Omega) + \\ &+ \frac{1}{\Delta t} \dot{S}_x(\Omega - \omega_{\Delta}) + \frac{1}{\Delta t} \dot{S}_x(\Omega - 2\omega_{\Delta}) + \dots \end{aligned} \quad (3.18)$$

Бачимо, що спектр дискретизованого сигналу містить спектр вихідного сигналу $S_x(\Omega)$, спектр вихідного сигналу зміщений на величину частоти дискретизації вправо $S_x(\Omega - \omega_{\Delta})$, той же спектр зміщений на величину частоти дискретизації вліво $S_x(\Omega + \omega_{\Delta})$, той же спектр зміщений на величину $2\omega_{\Delta}$ і т.д.

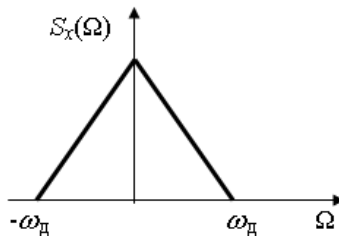


Рисунок 3.13 – Спектр вихідного неперервного сигналу

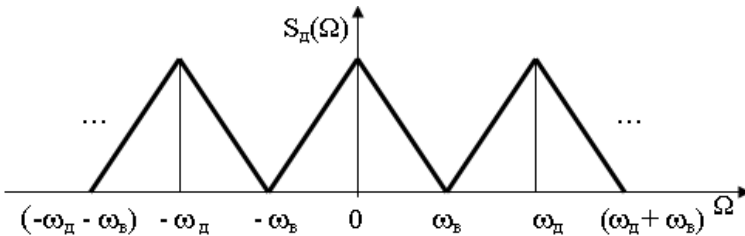


Рисунок 3.14 – Спектр дискретизованого сигналу

3.5 Відновлення неперервного сигналу з відліків

У лінію зв'язку передаються імпульси-відліки, які надходять вхід приймача.

Для відновлення вихідного неперервного сигналу з імпульсів-відліків ці імпульси треба подати на вхід ідеального фільтра низьких частот (ІФНЧ), який має наступні характеристики.

Амплітудно-частотна характеристика ідеального ФНЧ (АЧХ ІФНЧ) має вигляд:

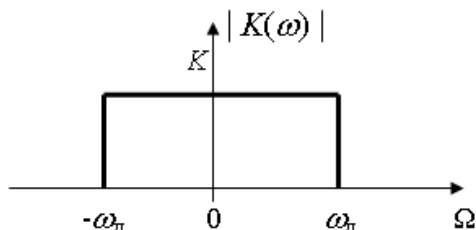


Рисунок 3.15 – АЧХ ідеального фільтра низьких частот

Імпульсна реакція ІФНЧ, тобто, реакція на дельта-імпульс має вигляд:

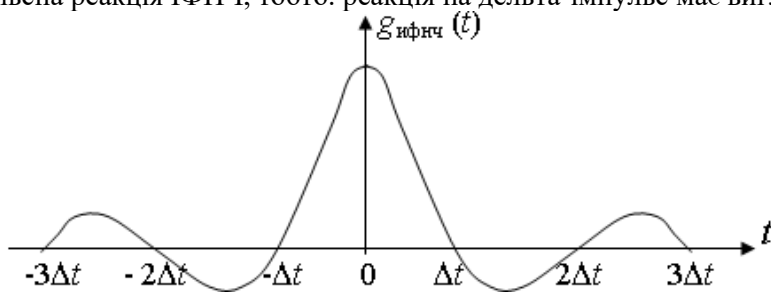


Рисунок 3.16 – Імпульсна реакція ІФНЧ на δ -імпульс

$$g_{\text{ІФНЧ}}(t) = K \frac{\sin \omega_b t}{\omega_b t}$$

$$\omega_b t = k\pi$$

$$t = k \frac{\pi}{\omega_b} = k\Delta t$$
(3.19)

Перша формула – це вираз імпульсної реакції ІФНЧ, друга і третя формули визначають моменти часу, для яких $g_{\text{ІФНЧ}}(t)$ перетворюється на нуль.

Зі спектральної точки зору ми пропускаємо дискретизований сигнал, що має спектр відповідно до рис. 3.14 через ІФНЧ з АЧХ рис. 3.15. Очевидно, що на виході ІФНЧ отримаємо спектр:

$$S(\omega) = K S_o(\omega) = K S_x(\omega) / \Delta t.$$

Таким чином, з точністю до постійного множника ми отримали на виході ІФНЧ спектр вихідного сигналу $s(t)$. З точки зору ми отримали вихідний неперервний сигнал $s(t)$.

3.6 Похибки дискретизації та відновлення неперервних сигналів

Теорема Котельникова точно справедлива лише сигналів з фінітним (кінцевим) спектром. На рис. 3.17 показані деякі варіанти фінітних спектрів:

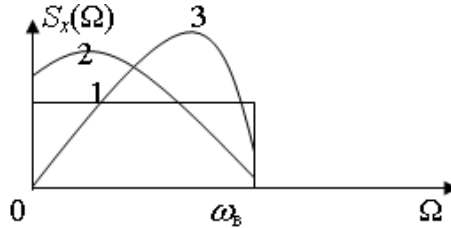


Рисунок 3.17 – Приклади фінітних спектрів

Проте спектри реальних інформаційних сигналів нескінченні. І тут теорема Котельникова справедлива з похибкою. **Похибка дискретизації** визначається енергією спектральних складових сигналу, що лежать поза частотою $\omega_в$.

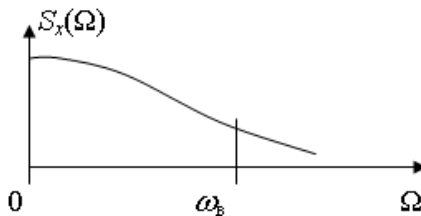


Рисунок 3.18 – Обмеження спектра реального сигналу

$$\overline{\Delta E_d^2} = \int_{\omega_в}^{\infty} \left| \dot{S}_x(\Omega) \right|^2 d\Omega \quad (3.20)$$

Друга причина виникнення похибок - неідеальність відновлюючого ФНЧ. Тбто похибка дискретизації та відновлення неперервного сигналу визначається такими причинами:

- 1) Спектри реальних сигналів не є фінітними.
- 2) АЧХ реальних ФНЧ неідеальні.

Висновок: що вище і що ближче характеристики ФНЧ до ідеальних, то ближче відновлений сигнал до вихідного.

Тема 4. Випадкові процеси в системах електрозв'язку

Характеристики випадкових процесів. Числові характеристики й кореляційна функція стаціонарних процесів. Нормальний випадковий процес (Гаусівський процес). Спектральна густина потужності стаціонарного випадкового процесу. Теорема Вінера-Хінчіна. Білий шум.

4.1 Характеристики випадкових процесів

Процеси, що розглядаються в теорії зв'язку, можуть бути детермінованими чи випадковими.

Детерміновані процеси – це процеси, протікання яких у часі відоме заздалегідь і практично точно. Наприклад, гармонійний сигнал $U(t) = U_m \cos(\omega_0 t + \varphi_0)$, де U_m , ω_0 , φ_0 – задані. Це найпростіша модель інформаційного сигналу, вона не забезпечує точного опису його основних перетворень, що зумовлює появу похибок у розрахунках.

Тому вводиться нова, складніша модель – **випадкові процеси** (ВП). Опис випадкових (чи нерегулярних) процесів здійснюється за допомогою випадкових (імовірнісних) функцій. Приклад - тепловий шум, випадковий процес, оскільки його параметри невідомі.

Випадкова функція – це така функція, поведінку якої не можна передбачити при проведенні експерименту. Випадкова функція часу називається випадковим (стохастичним) процесом. Окремі спостереження над ВП, проведені в однакових контрольованих умовах досліду, дають щораз різні функції $x(t)$ – різні реалізації випадкового процесу. Сукупність $\{x_k(t)\}$ всіх можливих реалізацій даного випадкового процесу називається ансамблем (рис. 4.1).

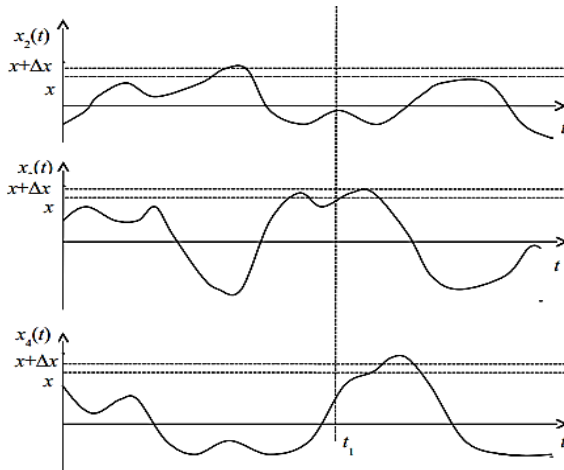


Рисунок 4.1 – Ансамбль реалізацій випадкового процесу

Реалізації відрізняються між собою але мають певні спільні характеристики.

Статистичний підхід до опису випадкового процесу полягає в тому, що визначають деякі усереднені характеристики для ансамблю $\{x_k(t)\}$ у цілому.

Функція розподілу імовірностей ВП (ФРІ).

Функція розподілу імовірностей позначається $F(x)$, що характеризує ймовірність того, що випадковий процес у певний момент часу t_1 набуває значення менше x_1 . Повне позначення одновимірної ФРІ:

$$F(x_1, t_1) = P(x < x_1, t = t_1) \quad (4.1)$$

Двовимірна ФРІ.

$$F_2(x_1 t_1, x_2 t_2) = P(x < x_1, t = t_1, x < x_2; t = t_2) \quad (4.2)$$

Найбільш повна характеристика n-мірна ФРІ:

$$F_n(x_1 t_1 \dots x_n t_n) = P(x < x_1; t < t_1; \dots x < x_n; t = t_n) \quad (4.3)$$

Функція густини ймовірностей випадкового процесу (ФГІ)

У найпростішому випадку одновимірної ФГІ дорівнює:

$$W(x_1 t_1) = \lim_{\Delta x \rightarrow 0} \frac{P(x_1 < x < x_1 + \Delta x; t = t_1)}{\Delta x} \quad (4.4)$$

Одновимірні ФГІ дорівнює межі відношення ймовірності потрапляння випадкового процесу в інтервал від x_1 до $x_1 + \Delta x$, при $t = t_1$, до Δx при Δx , що прямує до нуля.

Найбільш повною характеристикою є n-вимірні ФГІ.

ФРІ та ФГІ пов'язані один з одним. ФГІ – це перша похідна ФРІ по x_1 . Відповідно, ФРІ дорівнює інтегралу від $-\infty$ до x_1 від ФГІ:

$$F(x_1 t_1) = \int_{-\infty}^{x_1} W(x_1 t_1) dx_1, \text{ де } \int_{-\infty}^{\infty} W(x_1 t_1) dx_1 = 1 \quad (4.5)$$

4.2 Числові характеристики випадкового процесу

1. Середнє значення (математичне очікування чи перший початковий момент):

$$m_1 = \bar{x} = \int_{-\infty}^{+\infty} xW(x,t)dx \quad (4.6)$$

Фізичний зміст m_1 – це стала складова випадкового процесу.

2. Другий початковий момент:

$$m_2 = \bar{x}^2 = \int_{-\infty}^{+\infty} x^2W(x,t)dx \quad (4.7)$$

Фізичний зміст m_2 - це повна середня потужність випадкового процесу на одиничному опорі.

3. Дисперсія (другий центральний момент):

$$\sigma^2 = M_2 = \overline{(x-m_1)^2} = \int_{-\infty}^{+\infty} (x-m_1)^2W(x,t)dx \quad (4.8)$$

Фізичний зміст σ^2 – це середня потужність змінної складової випадкового процесу на одиничному опорі.

Числові характеристики пов'язані між собою:

$$\sigma^2 = m_2 - m_1^2 \quad (4.9)$$

Стационарність.

1. Нестационарний випадковий процес – ФГІ і ФРІ залежить від початку відліку часу.

2. Стационарний у вузькому значенні – ФГІ і ФРІ не залежить від початку відліку часу.

3. Стационарний у сенсі одно- і двовимірні ФГІ і ФРІ не залежить від початку відліку часу.

Для стационарного випадкового процесу m_1 , m_2 , σ^2 – не залежить від часу.

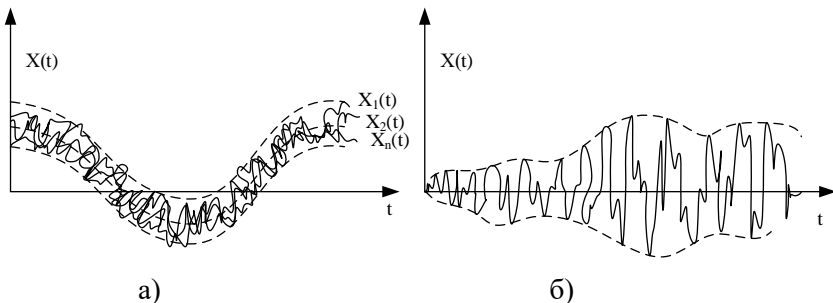


Рисунок 4.2 – Нестационарні процеси: а)змінне мат. очікування, б)змінна дисперсія

Ергодичність.

Випадковий процес називається ергодичним, якщо для нього усереднення за часом реалізації і усереднення по множині реалізацій дає той самий результат. Ця властивість має велике значення практично, оскільки усереднення за часом однієї реалізації технічно реалізувати простіше, але не завжди дає істинний результат. Тому підтвердження ергодичності процесу дозволяє значно спростити знаходження його параметрів.

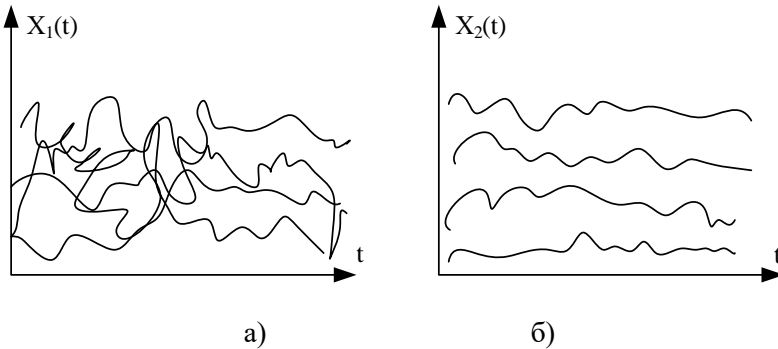


Рисунок 4.3 – а) Ергодичний та б) неергодичний процес

4.3 Нормальний випадковий процес (Гаусівський процес)

Процес називається нормальним або Гаусівським, якщо його одновимірний ФПІ має вигляд:

$$W(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2\sigma^2}} \quad (4.10)$$

Графіки нормальної ФПВ побудовано рис. 4.4:

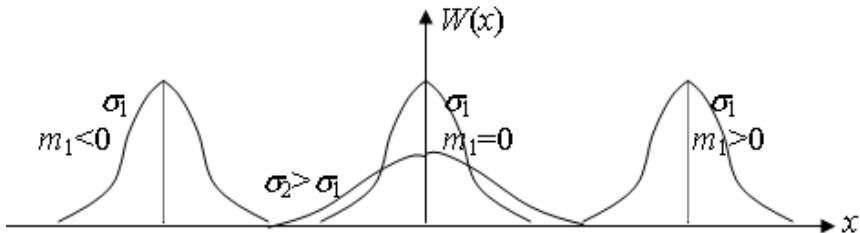


Рисунок 4.4 – Графіки нормальної функції щільності імовірності

Властивості нормального випадкового процесу.

1. $W(x) \geq 0$.

2. Нормальна ФПІ симетрична щодо $x = m_1$.
3. $W(x) - \max$ при $x = m_1$.
4. Площа під кривою $W(x)$ дорівнює 1.
5. При зміні m_1 форма кривої не змінюється, але крива зміщується вздовж осі x .
6. Чим більша дисперсія σ^2 , тим крива нижча та ширша.
7. З ймовірністю близькою до 1 ($P \cong 0,997$) миттєві значення нормального випадкового процесу лежать у межах: $m_1 - 3\sigma < x < m_1 + 3\sigma$.
8. ФРІ для нормального ВП – табульована функція (інтеграл імовірності Лапласа).

$$F(0) = 0,5$$

$$F(3,9) = 0,99995$$

$$F(-x) = 1 - F(x)$$

$$F(-\infty) = 0; F(\infty) = 1.$$

ФРІ для нормального процесу має вид:

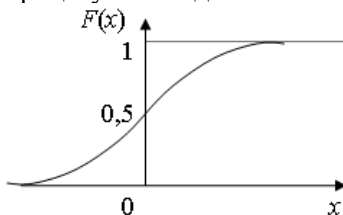


Рисунок 4.5 – Графік ФРІ для нормального процесу

4.4 Кореляційна функція стаціонарних процесів

В теорії зв'язку кореляційний теорія використовується при дослідженні випадкових процесів, дозволяючи встановити зв'язок між кореляційними і спектральними властивостями випадкових сигналів. Часто виникає завдання надійного виявлення одного сигналу, що передається в іншому або в перешкодах. Кореляційний аналіз дає уявлення про швидкість зміни в часі, а також тривалості сигналу без розкладання його на гармонійні складові.

Грецьке слово «кореляція» означає зв'язок між різними процесами та явищами. **Автокореляцією** називається узгодження сигналу $s(t)$ зі своєю запізненою версією $s(t-\tau)$. **Автокореляційна функція (АКФ)** дійсного діючого сигналу $s(t)$ кінцевої тривалості визначається наступним чином:

$$B_s(\tau) = \int_{-\infty}^{\infty} s(t) \cdot s(t-\tau) dt, \quad B_s(0) = \int_{-\infty}^{\infty} s^2(t) dt = E. \quad (4.11.)$$

Автокореляційна функція $B_s(\tau)$ показує міру схожості сигналу з власною копією, зміщеною τ одиниць часу. Значення автокореляційної функції в нулі періодичного сигналу дорівнює середній потужності (енергії) сигналу.

Змінна τ відіграє роль параметра сканування чи пошуку. Отже, АКФ є мірою статистичного зв'язку між значеннями процесу, час між якими τ . Зі збільшенням τ , АКФ спадає монотонно або коливаючись до нуля, якщо середнє значення процесу $=0$. Якщо інакше – прямує до величини $\overline{X(t)}^2$

Якщо сигнал $s(t)$ є періодичним із періодом T_0 , то автокореляційну функцію обчислюють за період, таким чином:

$$B_S(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t) \cdot s(t - \tau) dt \quad (4.12)$$

Розділивши $B_S(\tau)$ на $B_S(0)$, отримаємо **нормовану кореляційну функцію**, яка має позначення $R_S(\tau)$.

У системах зв'язку передачі сигналів широко використовуються відеоімпульси прямокутної форми. Для таких сигналів АКФ має вигляд

$$B_S(\tau) = \begin{cases} A_m^2 \cdot \tau_{и} (1 - \frac{|\tau|}{\tau_{и}}), & |\tau| \leq \tau_{и} \\ 0, & |\tau| > \tau_{и} \end{cases} \quad (4.13)$$

Енергія такого відеоімпульсу дорівнює

$$E_S = A_m^2 \tau_{и} \quad (4.14)$$

Важливим параметром сигналу $s(t)$ є тривалість його АКФ, що називається **інтервалом кореляції** k . Він визначається як відношення площі, обмеженої АКФ сигналу $s(t)$, до енергії сигналу

$$\tau_k = \frac{\int_{-\infty}^{\infty} B_S(\tau) d\tau}{B_S(0)} = \int_{-\infty}^{\infty} R_S(\tau) d\tau.$$

Умовно вважають, що значення процесу, які віддалені між собою на $\tau \leq \tau_k$ суттєво корельовані між собою, а значення процесу, які віддалені на $\tau > \tau_k$, некорельовані або корельовані несуттєво.

Властивості автокореляційної функції діючого сигналу:

1. $B_S(\tau)$ – парна; $B_S(\tau) = B_S(-\tau)$ (симетрія щодо τ щодо 0).

2. $B_S(0) = \max; B_S(0) = \int_{-\infty}^{\infty} s(t)^2 dt$, тобто в нулі дорівнює енергії сигналу.
3. $B_S(\tau) \leq B_S(0)$, кореляційна функція є спадаючою функцією модуля τ , тобто $\lim_{\tau \rightarrow \pm\infty} B_S(\tau) = 0$.
4. τ_k – інтервал кореляції випадкового процесу, що характеризує ширину графіка функції кореляції:

$$\tau_k = \begin{cases} \int_{-\infty}^{\infty} B(\tau) d\tau & |\tau| \leq \tau_k \text{ – то значення корельовані,} \\ 0 & |\tau| > \tau_k \text{ – значення не корельовані.} \end{cases}$$

5. $R(\tau) = B(\tau) / B(0)$ – коефіцієнт кореляції, $|R(\tau)| \leq 1$.
6. АКФ практично будь-якого сигналу не залежить від його фазового спектру. Отже, сигнали, амплітудні спектри яких повністю збігаються, а фазові розрізняються, матимуть однакову АКФ. Ще одне зауваження полягає в тому, що по АКФ можна відновити вихідний сигнал (знову ж з втратою інформації про фазу).

Процедура знаходження АКФ представлена на рис. 4.6. де видно, що АКФ прямокутного відеосигналу є рівнобедреним трикутником, і її тривалість дорівнює подвоєній тривалості імпульсу.

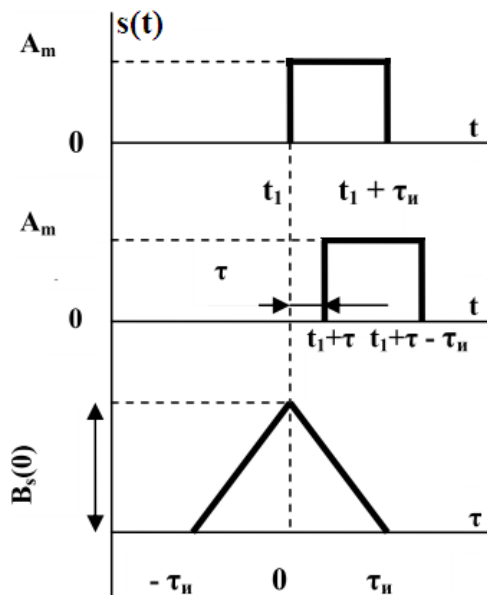


Рисунок 4.6 – Знаходження АКФ прямокутного відеоімпульсу

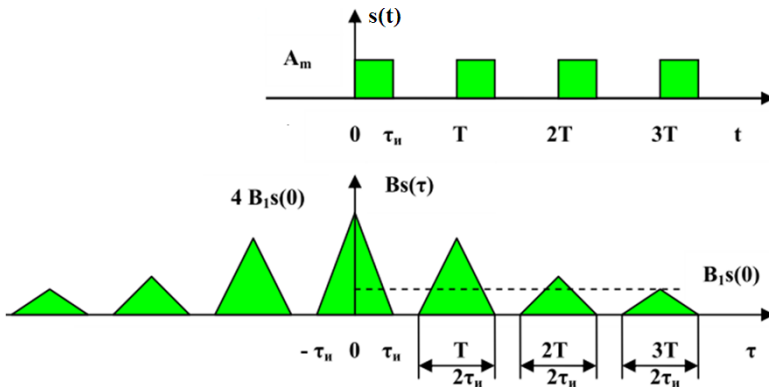


Рисунок 4.7 – Автокореляційна функція пачки імпульсів

Якщо ж сигнал $s(t)$ задається у вигляді пачки n імпульсів, зсунутих один щодо іншого на час T , то максимальне значення $B_S(\tau)$ при $\tau = 0$ дорівнює добутку енергії одного імпульсу на кількість імпульсів (див. рис. 4.7).

Для оцінки ступеня зв'язку між двома різними сигналами $s_1(t)$ та $s_2(t)$ використовується **взаємна кореляційна функція (ВКФ) $B_{12}(\tau)$**

$$B_{12}(\tau) = \int_{-\infty}^{\infty} s_1(t) \cdot s_2(t - \tau) dt \quad (4.15)$$

Розглянута вище автокореляційна функція $B_S(\tau)$ є окремим випадком функції $B_{12}(\tau)$, коли $s_1(t) = s_2(t) = s(t)$. На відміну від АКФ ВКФ двох сигналів не є парною і необов'язково максимальна при $\tau = 0$, тобто при відсутності часового зсуву сигналів.

Необхідно також відзначити, що $B_{12}(\tau)$ є асиметричною функцією щодо осі ординат і не досягає максимуму при $\tau = 0$.

Автокореляція дискретного сигналу

За аналогією до формули (7.1) АКФ дискретного сигналу $\{s_k\}$ та його затриманої копії на час $\tau = n\Delta t_n$ $\{s_{k-n}\}$ може бути представлена у вигляді

$$\widehat{B}_S(n) = \sum_{k=-\infty}^{\infty} s_k \cdot s_{k-n} \quad (4.16)$$

Ця функція, яка залежить від числа тактів зсуву дискретної послідовності n , як і звичайна АКФ, є парною, тобто при нульовому зсуві $n = 0$ дискретна АКФ визначає енергію дискретного сигналу

$$\widehat{B}_s(0) = \sum_{k=-\infty}^{\infty} s^2_k. \quad (4.17)$$

Як приклад розрахуємо АКФ дискретного аналога імпульсу прямокутної форми з одиничними амплітудами $\{1, 1, 1\}$. Копії цього сигналу та значення дискретної АКФ мають вигляд рис. 4.8:

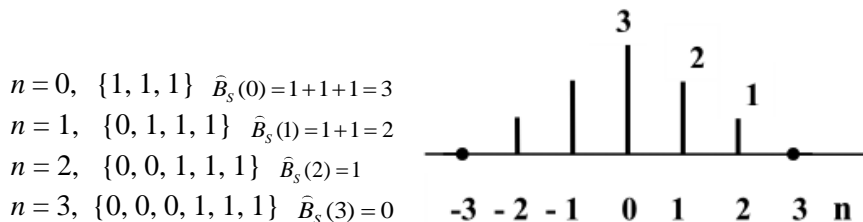


Рисунок 4.8 – Автокореляційна функція дискретного сигналу

Як і у випадку аналогових відеоімпульсів, пелюстки дискретної АКФ із збільшенням зсуву n зменшуються за лінійним законом.

Змінимо форму дискретного сигналу, отже він буде описуватися у вигляді послідовності $\{1, 1, -1\}$. Його АКФ матиме вигляд рис. 4.9:

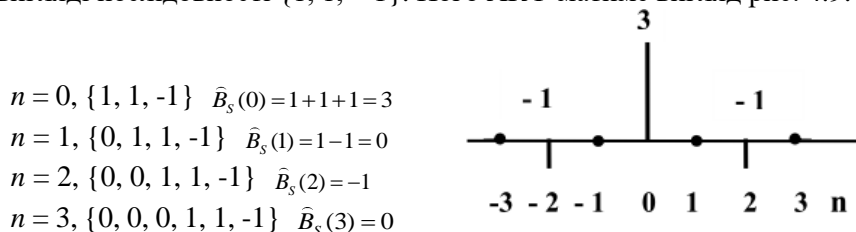


Рисунок 4.9 – Автокореляційна функція дискретного сигналу

Порівнюючи дискретні АКФ на рис. 4.8. та рис. 4.9. можна відзначити, що саме сигнал $\{1, 1, -1\}$ має найбільш досконалу з точки зору рівня бічних пелюсток кореляційну функцію. Цей сигнал є найпростішим із сімейства сигналів Баркера, що представляють собою М-позиційні сигнали, у яких значення бічних пелюсток АКФ при $n \neq 0$ не перевищують одиниці. Енергія цих сигналів $\widehat{B}_s(0) = E_s$ завжди дорівнює числу позицій М. Доведено, що число М у сигналах Баркера не може перевищувати числа 13. На цей час відомі сигнали, у яких число позицій М дорівнює 3, 4, 5, 7, 11 і 13.

Зв'язок кореляційної функції з енергетичним спектром

Перетворення Фур'є для випадкового процесу неможливе, оскільки існує нескінченна кількість його реалізацій. Але можливо здійснити перетворення Фур'є для АКФ. Скористаємося виразом теореми Парсеваля:

$$\int_{-\infty}^{\infty} s(t) \cdot u(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{U}(\omega) \cdot \dot{S}^*(\omega) d\omega, \quad (4.18)$$

де зсув функції виразимо як $u(t)=s(t+\tau)$ і відповідно $\dot{U}(\omega) = \dot{S}(\omega) \cdot e^{j\omega\tau}$.

Тоді:

$$B_S(\tau) = \int_{-\infty}^{\infty} s(t) \cdot s(t-\tau) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) \cdot \dot{S}^*(\omega) \cdot e^{j\omega\tau} d\omega. \quad (4.19)$$

Враховуючи, що $\dot{S}(\omega) \cdot \dot{S}^*(\omega) = |\dot{S}(\omega)|^2 = W_S(\omega)$, маємо:

$$B_S(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} W_S(\omega) \cdot e^{j\omega\tau} d\omega, \quad (4.20)$$

$$W_S(\tau) = \int_{-\infty}^{\infty} B_S(\omega) \cdot e^{-j\omega\tau} d\tau. \quad (4.21)$$

Отже, пряме перетворення Фур'є (4.21) кореляційної функції $B_S(\tau)$ дозволяє отримати енергетичний спектр сигналу $W_S(\omega)$ (спектральну щільність потужності), що показує розподіл енергії по частоті. Зворотне перетворення (4.20) визначає кореляційну функцію. Таким чином, спектральна густина потужності й кореляційна функція стаціонарного випадкового процесу зв'язані перетвореннями Фур'є. Це твердження відоме як **теорема Хінчина-Вінера**.

З виразів (4.20) та (4.21) можна зробити такі висновки.

1. Чим ширший спектр $|\dot{S}(\omega)|$ сигналу $s(t)$, тим менше інтервал кореляції τ_k . Інакше кажучи, інтервал кореляції й ширина спектру процесів обернено пропорційні величини, а їх добуток є величиною порядку 0,5:

$$\tau_k * F_{max} \sim 0,5.$$

2. Чим більший інтервал кореляції τ_k заданого сигналу, тим менше ширина його спектру.

3. Кореляційна функція $B_S(\tau)$ не залежить від ФЧХ спектра сигналу. Але оскільки форма сигналу $s(t)$ при заданому амплітудному спектрі істотно залежить від ФЧХ, можна зробити висновок, що різним за формою сигналам, що мають однакові амплітудні спектри, відповідають однакові кореляційні функції $B_S(\tau)$.

Необхідно відзначити, що отриманий зв'язок між автокореляційною функцією та енергетичним спектром дозволяє встановити критерій існування сигналу із заданими характеристиками. Відомо, що енергетичний спектр будь-якого сигналу, за визначенням, завжди додатний. Це означає, що кореляційна функція не може мати, наприклад, вид прямокутника, тому що в цьому випадку енергетичний спектр повинен описуватися знакозмінною функцією, що суперечить фізичним уявленням.

4.5 Білий шум

Спектр повільно змінюваних сигналів містить більше компонентів в лівій частині осі в області низьких частот і навпаки.

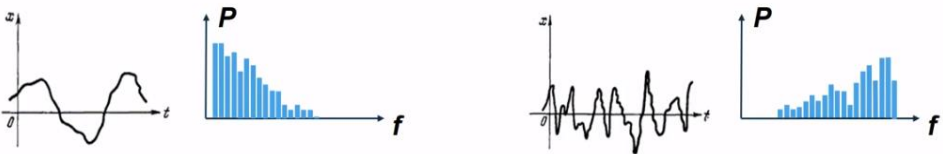


Рисунок 4.10 – Приклад спектрів стаціонарних процесів

Випадковий процес, в якого область частот заповнена рівномірно називається білим шумом. Це стаціонарний ВП, який отримав свою назву по аналогії з білим кольором як сума кольорів всіх частот. Тепловий шум, дробовий шум, випромінювання Землі і небесних тіл, тощо в сумі будуть мати нормальний розподіл. Аналогічно отримали назви так звані кольорові шуми в залежності від спектрального розподілу.

Основні характеристики білого шуму:

- рівномірна спектральна густина потужності;
- нормальний розподіл значень в часі;
- додається до корисного сигналу;
- статистично незалежний від корисного сигналу.

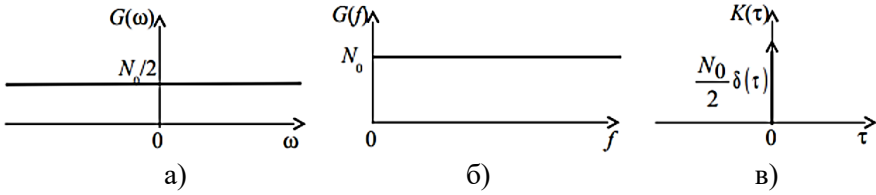


Рисунок 4.10 –Спектральна густина потужності білого шуму а) двосторонній спектр; б) односторонній спектр; в) кореляційна функція білого шуму

Кореляційна функція білого шуму визначається:

$$K(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{N_0}{2} \cdot e^{j\omega\tau} d\omega = \frac{N_0}{2} \delta(\tau) \quad (4.22)$$

З виразу видно, що інтервал кореляції прямує до нуля, тобто будь які два значення білого шуму є незалежними. Очевидно, що білий шум має нескінченну потужність і фізично існувати не може, проте є зручною моделлю, що часто використовується в теорії і техніці зв'язку.

Тема 5. Методи формування та перетворення сигналів

Загальні відомості про модуляцію. Амплітудна модуляція. Кутова модуляція. Квадратурна модуляція. Формування та детектування модульованих сигналів. Дискретна модуляція гармонічного носія (маніпуляція). Імпульсна модуляція. Модуляція шумової несучої. Кодування та декодування сигналів.

Перетворення дискретного повідомлення в сигнал звичайно здійснюється у вигляді двох операцій – кодування і модуляції (рис. 5.1.). Кодування являє собою перетворення повідомлення в послідовність кодових символів, а модуляція – перетворення цих символів у сигнали, придатні для передачі по каналу. За допомогою кодування і модуляції джерело повідомлень узгоджується із каналом.

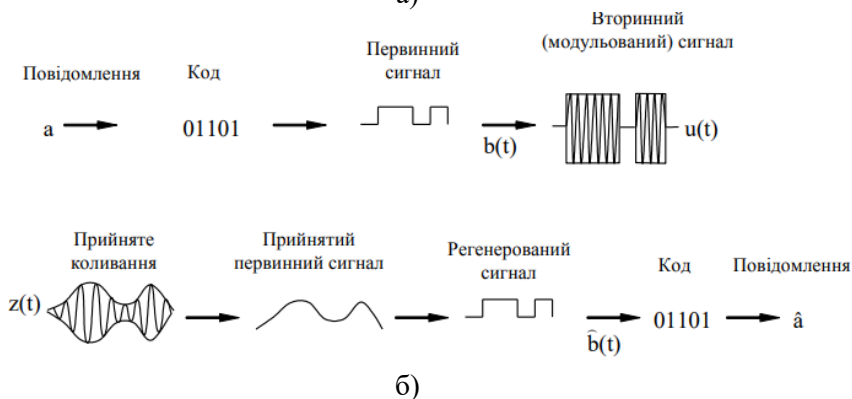
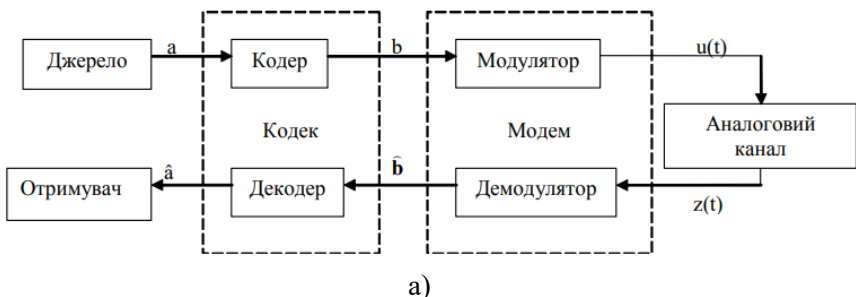


Рисунок 5.1 – а) Структурна схема системи; б) процес передачі повідомлень

5.1 Модуляція сигналів

Сигнали від вимірювальних датчиків та будь-яких інших джерел інформації передаються по лініях зв'язку до приймачів - вимірювальних приладів, у вимірювально-обчислювальні системи реєстрації та обробки

даних, у будь-які інші центри накопичення та зберігання даних. Як правило, інформаційні сигнали є низькочастотними і обмеженими по ширині спектру, на відміну від ширококутових високочастотних каналів зв'язку, розрахованих на передачу сигналів від множини джерел одночасно з частотним поділом каналів. Перенесення спектра сигналів з низькочастотної області у виділену їх передачі область високих частот виконується операцією модуляції.

Формування модульованих сигналів (**модуляція**) передбачає взаємодію двох сигналів: керуючого модулюючого та допоміжного несучого. Суть управляючого впливу модулюючого сигналу $s_c(t)$ полягає в тому, що деякі параметри коливання змінюються відповідно до модулюючого коливання.

У системах зв'язку як управляючі коливання використовуються різноманітні первинні електричні сигнали (ПЕС): телефонні, телеграфні, телевізійні та ін.

Як несучі широко застосовуються гармонічні сигнали, власна частота яких ω_0 значно перевищує верхню частоту спектра модулюючого коливання Ω_{\max} . Це означає, що по відношенню до несучого коливання, модулююче коливання повільно змінює свої значення в часі. Повільність зміни $s_c(t)$ підкреслює, що на період модулюючого коливання, припадають тисячі, сотні тисяч і більше періодів несучого коливання. При цьому з одного боку забезпечується досить повне відображення модулюючого коливання в несучому коливанні, а з іншого, обумовлюється вузькосмуговий спектр модульованого коливання.

Таким чином, для передачі інформації, що міститься в ПЕС, використовується допоміжне несуче коливання, що виконує роль переносника повідомлення:

$$S_n(t) = U \cdot \cos(\omega_n t + \varphi_n), \quad (5.1)$$

Зазвичай вважають $f_n \gg kF_1$, де F_1 – найвища гармоніка ПЕС.

Процес зміни одного або кількох параметрів високочастотного (несучого) коливання відповідно до первинного (модулюючого) сигналу називається модуляцією.

При модуляції інформаційними параметрами несучого коливання $S_n(t) = U \cdot \cos(\omega_n t + \varphi_n)$ можуть бути амплітуда U , частота ω_n або фаза φ_n , які змінюються відповідно до модулюючого сигналу $s_c(t)$, тому розрізняють **амплітудну модуляцію (АМ)**, **частотну модуляцію (ЧМ)** та **фазову модуляцію (ФМ)**. У каналах передачі цифрової інформації набула також поширення **квадратурна модуляція**, при якій одночасно змінюються

амплітуда і фаза коливань. Дискретну модуляцію (модуляція дискретних даних за допомогою високочастотного носія) зазвичай називають **маніпуляцією**.

При використанні в якості несучих сигналів періодичних послідовностей імпульсів (наприклад, прямокутних) вільними параметрами модуляції можуть бути амплітуда, тривалість, частота слідування і фаза (положення імпульсу відносно тактової точки) імпульсів. Це дає чотири основні види **імпульсної модуляції**: АІМ, ДІМ, ЧІМ та ФІМ.

В якості несучих сигналів можна використовувати як періодичні коливання, а й стаціонарні випадкові процеси (**модуляція шумової несучої**). Як модульовані параметри випадкових сигналів використовуються моменти випадкових процесів. Так, наприклад, модуляція другого моменту випадкових послідовностей (модуляція за потужністю) є аналогією амплітудної модуляції.

У модульованих коливаннях змінні параметри мають вигляд:

- при амплітудній модуляції – $U(t) = U_m + \Delta U(t) = U_m + a \cdot s_c(t)$;
- при частотній модуляції – $\omega_n(t) = \omega_0 + \Delta\omega(t) = \omega_0 + a \cdot s_c(t)$;
- при фазовій модуляції – $\varphi_n(t) = \varphi_0 + \Delta\varphi(t) = \varphi_0 + a \cdot s_c(t)$;

де $\Delta A(t)$, $\Delta\omega(t)$, $\Delta\varphi(t)$ – прирости, пропорційні до модулюючого коливання $s_c(t)$; a – коефіцієнт пропорційності.

Пристрій для отримання результуючого (модульованого) сигналу $S_m(S_c, t)$, називається **модулятором**, на один вхід якого подається несуче коливання $S_n(t)$, (модульоване), на другий вхід первинний (модулюючий) сигнал $s_c(t)$.

5.1.1 Амплітудна модуляція гармонійного коливання

Амплітудна модуляція – процес зміни амплітуди коливання, що відповідає зміні неперервного інформаційного сигналу. Історично була першим видом модуляції, освоєним практично. В даний час АМ застосовується в основному тільки для радіомовлення на порівняно низьких частотах (не вище за короткі хвили) і для передачі зображення в телевізійному мовленні. Це обумовлено низьким ККД використання енергії модульованих сигналів.

При амплітудній модуляції миттєва амплітуда несучого коливання:

$$U(t) = U_m + a \cdot s_c(t), \quad (5.2)$$

де U_m – амплітуда несучої; a – коефіцієнт пропорційності, що вибирається так, щоб амплітуда $U(t)$ завжди була додатньою. Частота і фаза гармонійного коливання при АМ залишаються незмінними.

Для математичного опису АМ сигналу (5.2) замість коефіцієнта a , що залежить від конкретної схеми модулятора, вводиться індекс модуляції:

$$m_{AM} = \frac{S_0}{U_{max}} \quad (5.3)$$

тобто відношення амплітуд модулюючого і несучого коливання. Фізично індекс модуляції характеризує **глибину амплітудної модуляції** і може змінюватися в межах $0 \leq m_{AM} \leq 1$.

При значенні $m_{AM} < 1$ форма огинаючої несучого коливання повністю повторює форму модулюючого сигналу $s(t)$, малу глибину модуляції для основних гармонік модулюючого сигналу ($m_{AM} \ll 1$) застосовувати недоцільно, оскільки при цьому потужність переданого інформаційного сигналу буде набагато менше потужності несучого коливання, і потужність передавача використовується неекономічно

Стовідсоткова (глибока) модуляція ($m_{AM} = 1$) може призводити до спотворень сигналів при перевантаженнях передавача, якщо останній має обмежений динамічний діапазон по амплітуді несучих частот або обмежену потужність передавача (збільшення амплітуди несучих коливань, в пікових інтервалах сигналу $U(t)$ в два рази вимагає збільшення потужності передавача вчетверо).

При $m_{AM} > 1$ виникає так звана перемодуляція, Форма огинаючої при перемодуляції спотворюється щодо форми модулюючого сигналу і після демодуляції, якщо застосовуються її найпростіші методи, інформація може спотворюватися.

Таким чином, для будь-якого АМ сигналу справедливо:

$$S_{AM}(s_c, t) = U_m [1 + m_{AM} \cdot s_c] \cos(\omega_0 t + \varphi_0). \quad (5.4)$$

Амплітудна модуляція гармонійним коливанням. У найпростішому випадку модулюючий сигнал є гармонійним коливанням із частотою $\Omega \ll \omega_0$. При цьому вираз

$$S_{AM}(s_c, t) = U_m [1 + m_{AM} \cdot \cos \Omega t] \cos(\omega_0 t + \varphi_0), \quad (5.5)$$

відповідає однотональному АМ сигналу, представленою на рис. 5.2.

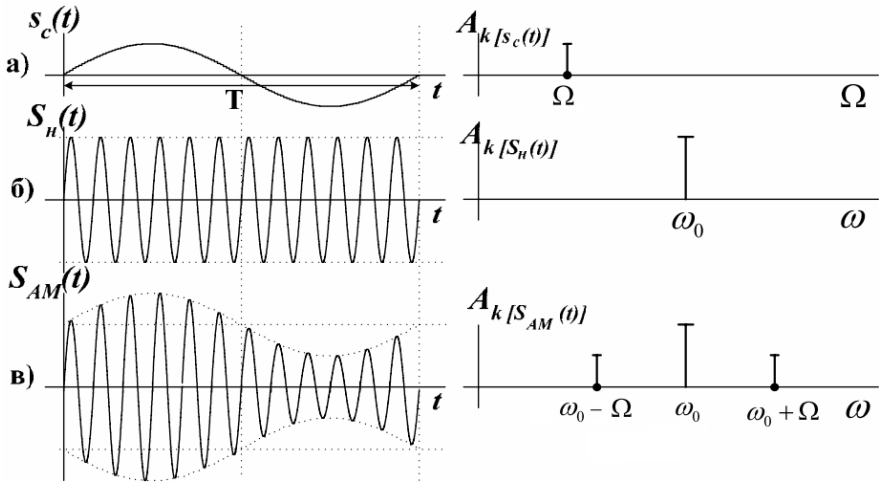


Рисунок 5.2 – Часові та спектральні діаграми процесу формування АМ гармонійного коливання

Однотональний АМ сигнал можна подати у вигляді суми трьох гармонійних складових із частотами: ω_0 – несучої; $\omega_0 + \Omega$ – верхньої бокової і $\omega_0 - \Omega$ – нижньої бокової:

$$S_{AM}(s_c, t) = U_m \cdot \cos(\omega_0 t + \varphi_0) + (U_m \cdot m_{AM}/2)\cos[(\omega_0 + \Omega)t + \varphi_0] + (U_m \cdot m_{AM}/2)\cos[(\omega_0 - \Omega)t + \varphi_0] \quad (5.6)$$

Спектральна діаграма однотонального АМ сигналу, побудована (5.7), симетрична щодо несучої частоти ω_0 (рис. 5.2 в). Амплітуди бічних коливань із частотами $\omega_0 - \Omega$ і $\omega_0 + \Omega$ однакові і навіть при $m_{AM} = 1$ не перевищують половини амплітуди несучого коливання U_m .

Гармонічні модулюючі сигнали і відповідно однотональний АМ сигнал на практиці зустрічаються рідко. Найчастіше модулюючі первинні сигнали $s_c(t)$ є складними функціями часу (рис. 5.3,а). Будь-який складний сигнал $s_c(t)$ можна подати у вигляді кінцевої або нескінченної суми гармонійних складових, скориставшись рядом або інтегралом Фур'є. Кожна гармонійна складова сигналу $s_c(t)$ із частотою Ω_i призведе до появи в АМ сигналі двох бічних складових із частотами $\omega_0 \pm \Omega_i$.

Множині гармонійних складових у модулюючому сигналі з частотами $\Omega_i, i = 1, 2, \dots, N$ буде відповідати множина бічних складових з частотами $\omega_0 \pm \Omega_i, i = 1, 2, \dots, N$. Для наочності таке перетворення спектру при АМ показано на рис. 5.3 б. Спектр складномодульованого АМ сигналу, крім несучого

коливання з частотою ω_0 , містить групи верхніх і нижніх бічних коливань, що утворюють відповідно верхню бічну смугу і нижню бічну смугу АМ сигналу.

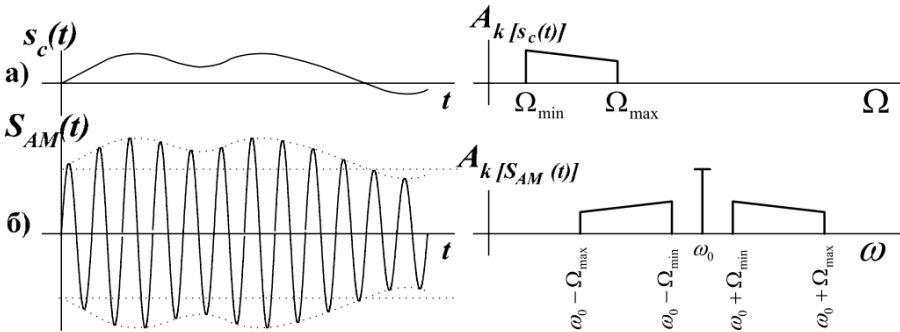


Рисунок 5.3 – Часові і спектральні діаграми АМ сигналу

При цьому верхня бічна смуга частот є масштабною копією спектра інформаційного сигналу, зсунутого в область високих частот величину ω_0 . Нижня бічна смуга частот також повторює спектральну діаграму сигналу $s_c(t)$, але частоти у ній розташовуються у дзеркальному порядку щодо несучої частоти ω_0 .

Ширина спектру АМ сигналу $\Delta\omega_{AM}$ дорівнює подвоєному значенню найбільш високої частоти Ω_{\max} спектру модулюючого низькочастотного сигналу, тобто $\Delta\omega_{AM} = 2\Omega_{\max}$.

Наявність двох бічних смуг обумовлює розширення займаної смуги частот приблизно удвічі, проти спектру інформаційного сигналу. Потужність, що припадає на коливання несучої частоти, постійна. Потужність, укладена у бічних смугах, залежить від індексу модуляції та збільшується зі збільшенням глибини модуляції. Однак навіть у крайньому випадку, коли $m_{AM} = 1$, тільки 1/3 всієї потужності коливання припадає на дві бічні смуги.

Повна середня потужність сигналу АМ на опорі R дорівнює:

$$\frac{U_m^2}{2R} \left(1 + \frac{m_{AM}^2}{2}\right) \quad (5.7)$$

Якщо амплітуда несучої в спектрі АМ сигналу дорівнює U_m , глибина модуляції дорівнює m_{AM} , то амплітуди бічних частот дорівнюють:

$$m_{AM}U_m/2. \quad (5.8)$$

Під корисною потужністю модульованих сигналів розуміють потужність, що несе інформацію, тобто в даному випадку потужність бічних частот. Коефіцієнт корисної дії даного типу модуляції визначається відношенням потужності бічних частот до загальної середньої потужності модульованого сигналу:

$$\eta_{AM} = (U_m^2 m_{AM}^2 / 4) / P_u = m_{AM}^2 (m_{AM}^2 + 2) \quad (5.9)$$

Навіть при $m_{AM}^2 = 1$ ККД амплітудної модуляції становить лише 33%, а при практичному використанні зазвичай менше 20%.

Балансна модуляція

Аналіз спектрального складу АМ сигналу показав, що первинний модулюючий сигнал знаходить своє відображення лише складових бічних смуг спектру АМ сигналу. У процесі відображення первинного сигналу в модульованому коливанні складова спектра частоти ω_0 виконує роль своєрідного початку відліку для частот бічних спектральних складових. Тому її можна виключити зі спектру сигналу, що передається, і відновити на приймальному кінці.

Якщо модульоване коливання не містить складової несучої частоти ω_0 , то модуляцію називають балансною (БМ). Два сигнали з фазовим зсувом на 180 градусів одночасно подаються на дві різні частини схеми, що призводить до того, що несуча хвиля компенсується (балансується), залишаючи тільки амплітудно модульовані бічні смуги. Це зменшує кількість потужності, що витрачається на несучу частоту, та підвищує ефективність передачі даних. Такий вид модуляції доцільний з енергетичної точки зору, оскільки на несучу частоту, припадає 2/3 всієї потужності модульованого коливання. За інших рівних умов потужність, що вивільнилася, дозволить реалізувати велику дальність зв'язку, або при колишній дальності поліпшити її якість.

Односмугова модуляція

Балансна модуляція дозволяє більш раціонально розподілити енергію сигналу, проте ширина спектра $\Delta\Omega_{БМ}$ залишається такою самою, як і для звичайної амплітудної модуляції. У той же час симетрія спектру сигналу АМ означає, що верхня бічна смуга і нижня бічна смуга кожна окремо, повністю відображають модулююче коливання. При цьому друга бічна смуга не несе жодної додаткової інформації, удвічі розширюючи спектр. Вид модуляції, при якому в спектрі амплітудно-модульованого коливання зберігається лише одна бічна смуга (верхня або нижня) називається односмуговою модуляцією.

5.1.2 Методи кутової модуляції

При фазовій та частотній модуляції сигнал має постійну амплітуду і може бути записаний у наступному вигляді:

$$S_{\text{ФМ(ЧМ)}}(t) = U \cdot \cos(\varphi(t)). \quad (5.10)$$

За відсутності модуляції аргумент гармонійного коливання миттєва (повна) фаза $\varphi(t) = \omega_0 t$ змінюється із постійною швидкістю ω_0 , тобто є лінійною функцією часу. І фазова, і частотна модуляція передбачають залежність зміни фази $\varphi(t)$ від інформаційного сигналу $s_c(t)$. Ця спільність дозволяє об'єднати обидва види модуляції однією назвою – кутова модуляція.

При кутовій модуляції лінійність зміни $\varphi(t)$ порушується і в кожний момент часу t швидкість зміни $\varphi(t)$ визначається миттєвою частотою $\omega(t)$, причому:

$$\omega(t) = \frac{d\varphi(t)}{dt}; \quad \varphi(t) = \int_0^t \omega(t) dt. \quad (5.11)$$

Принципи частотної та фазової (кутової) модуляції

Фазова модуляція - процес зміни миттєвої фази коливання пропорційно зміні неперервного інформаційного сигналу:

$$\varphi(t) = \omega_0 t + \Delta\varphi(t) = \omega_0 t + a \cdot s_c(t) \quad (5.12)$$

Таким чином

$$S_{\text{ФМ}}(t) = U_m \cdot \cos[\omega_0 t + a \cdot s_c(t)]. \quad (5.13)$$

Максимальне відхилення фази називається індексом модуляції:

$$a |s_c(t)|_{\max} = m_{\text{ФМ}}. \quad (5.14)$$

Якщо модуляція здійснюється гармонічним коливанням (**тональна модуляція**) $s_c(t) = U_{m\Omega} \cdot \cos\Omega t$ із частотою Ω , то

$$S_{\text{ФМ}}(t) = U_m \cdot \cos(\omega_0 t + a \cdot U_{m\Omega} \cdot \cos\Omega t) = U_m \cdot \cos(\omega_0 t + m_{\text{ФМ}} \cdot \cos\Omega t).$$

Зауважимо, що індекс модуляції $m_{\text{ФМ}} = a \cdot U_{m\Omega}$ пропорційний амплітуді модулюючого коливання.

На рис. 5.4. показано, як змінюються миттєва частота та фаза при тональній фазовій модуляції.

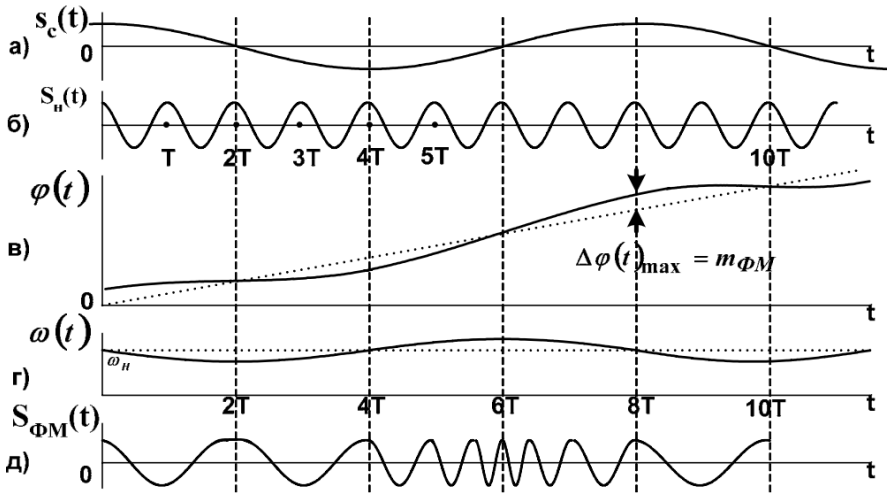


Рисунок 5.4 – Часові діаграми процесу формування ФМ сигналів

Інформаційний однотоновий сигнал $s_c(t) = U_{m\Omega} \cdot \cos \Omega t$ (рис. 5.4. а) модулює несуче коливання $s_n(t)$ (рис. 5.4. б), при цьому закон зміни миттєвої фази коливання $\varphi(t) = \omega_0 t + a \cdot s_c(t)$ повторює закон зміни $s_c(t)$ (рис. 5.4.), тобто на лінійну зміну фази (пунктир на малюнку) накладається змінний приріст $\Delta\varphi(t) = a \cdot s_c(t)$, а закон зміни миттєвої частоти коливання $\omega(t)$ (рис. 5.4 г) визначається похідною:

$$\omega(t) = \frac{d\varphi(t)}{dt} = \frac{d}{dt}(\omega_0 t + a \cdot U_{m\Omega} \cos \Omega t) = \omega_0 - a \cdot U_{m\Omega} \sin \Omega t.$$

Фазомодульоване коливання (рис. 9.1 д) у моменти часу $t = 2T$ і $t = 10T$ сигнал $S_{\text{ФМ}}(t)$ має мінімальну, а в момент $t = 6T$ максимальну миттєву частоту.

Частотна модуляція – процес зміни миттєвої частоти коливання відповідно до зміни інформаційного сигналу: $\omega(t) = \omega_0 + a \cdot s_c(t)$.

Розглянемо найпростіший спосіб однотонової частотної модуляції.

На рис. 5.5 зображено часові діаграми зміни миттєвої частоти та фази для однотонової частотної модуляції.

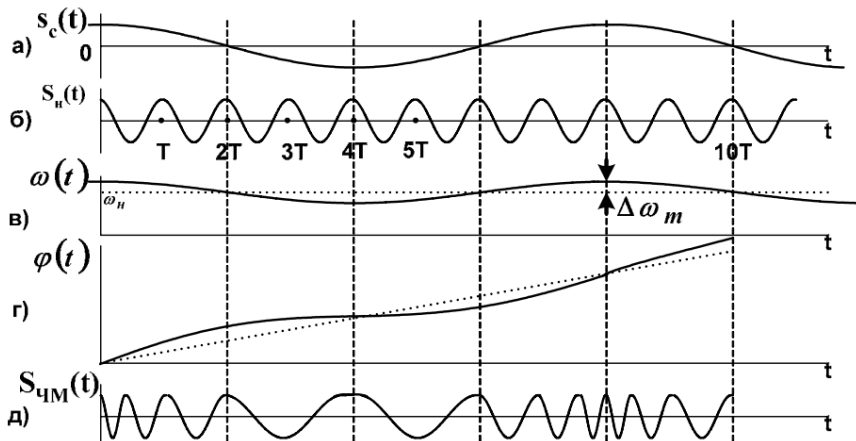


Рисунок 5.5 – Часові діаграми процесу формування ЧС сигналів

Інформаційний однотональний сигнал $s_c(t) = U_{m\Omega} \cdot \cos\Omega t$ (рис. 5.5. а) модулює несуче коливання $s_n(t)$ (рис. 5.5. б), при цьому закон зміни миттєвої частоти коливання $\omega(t) = \omega_0 + U_m \cdot a \cdot \cos\Omega t$ повторює закон зміни $s_c(t)$ (рис. 5.5.). Тут $a \cdot U_{m\Omega} = \Delta\omega_m$ – девіація частоти, пропорційна амплітуді модулюючого коливання $U_{m\Omega}$.

Девіацією частоти називається максимальне відхилення частоти від середнього значення ω_0 :

$$a|s_c(t)|_{\max} = \Delta\omega_m. \quad (5.15)$$

Відношення девіації частоти $\Delta\omega_m$ до частоти модулюючого коливання Ω називається індексом частотної модуляції:

$$m_{\text{ЧМ}} = \Delta\omega_m / \Omega. \quad (5.16)$$

У моменти часу $t = 0, t = 8T$ миттєва частота максимальна, у момент $t = 4T$ – мінімальна. Закон зміни миттєвої фази коливання $\varphi(t)$ (рис. 5.5. г) визначається інтегруванням

$$\varphi(t) = \int_0^t \omega(t) dt; \quad \varphi(t) = \omega_0 t + m_{\text{ЧМ}} \sin \Omega t..$$

Враховуючи зв'язок частоти та фази, вираз для частотно-модульованого сигналу запишеться наступним чином:

$$S_{\text{ЧМ}}(t) = U_m \cos\left[\int_0^t \omega(t) dt\right] = U_m \cos\left[\omega_0 t + a \int_0^t s_c(t) dt\right]. \quad (5.17)$$

Для тональної частотної модуляції:

$$S_{\text{ЧМ}}(t) = U_m \cdot \cos(\omega_0 t + m_{\text{ЧМ}} \cdot \sin \Omega t). \quad (5.18)$$

Порівняння виразів (5.13) і (5.17) показує, що при ФМ збільшення фази пропорційно до модулюючого коливання $s_c(t)$, а при ЧС – інтегралу від $s_c(t)$. Якщо спочатку проінтегрувати $s_c(t)$, а потім цим коливанням модулювати несучу фазу, то вийде ЧС сигнал. Такий спосіб формування ЧС сигналу застосовується практично. Подібним чином, якщо продиференціювати $s_c(t)$ і це коливання використовувати для модуляції частоти, то отримаємо ФМ сигнал.

Спектр сигналів кутової модуляції

Сигнали з кутовою модуляцією, як і за АМ, можуть бути представлені у вигляді суми гармонійних коливань. Порівняно просто це можна зробити для тональної модуляції. При тональній модуляції спектри ФМ та ЧС однакові, якщо $m_{\text{ФМ}} = m_{\text{ЧМ}} = m$, тому розглядатимемо лише спектр ЧС сигналу.

Перетворимо (5.18) за формулою косинуса суми двох аргументів:

$$S_{\text{ЧМ}}(t) = U_m \cdot \cos(\omega_0 t + m \cdot \sin \Omega t) = U_m \cdot \cos(\omega_0 t) \cdot \cos(m \cdot \sin \Omega t) - U_m \cdot \sin(\omega_0 t) \cdot \sin(m \cdot \sin \Omega t). \quad (5.19)$$

З теорії Бесселевих функцій відомі такі співвідношення:

$$\begin{aligned} \cos(m \sin x) &= J_0(m) + 2 \sum_{k=1}^{\infty} J_{2k}(m) \cos 2kx, \\ \sin(m \sin x) &= 2 \sum_{k=1}^{\infty} J_{2k-1}(m) \sin((2k-1)x), \end{aligned} \quad (5.20)$$

де $J_k(m)$ – функція Бесселя k -го порядку аргументу m . Підставляючи (5.20) в (5.19), виконуючи звичайні алгебраїчні перетворення і розкриваючи добуток тригонометричних функцій, отримуємо:

$$S_{\text{ЧМ}}(t) = U_m J_0(m) \cos(\omega_0 t) + \sum_{k=1}^{\infty} U_m J_k(m) \cos(\omega_0 + k\Omega)t + \sum_{k=1}^{\infty} (-1)^k U_m J_k(m) \cos(\omega_0 - k\Omega)t. \quad (5.21)$$

Таким чином, спектр навіть для однотонольної кутової модуляції досить складний. У формулі (5.21) перший член – гармонійна складова із частотою несучої. Група гармонійних складових із частотами $(\omega_0 + k\Omega)$, $k = 1, 2, \dots$, визначає верхню бічну смугу частот, а група складових із частотами $(\omega_0 - k\Omega)$, $k = 1, 2, \dots$, нижню бічну смугу частот. Число верхніх та нижніх гармонік бічних частот теоретично нескінченне.

Бічні гармонічні коливання розташовані симетрично щодо ω_0 з відривом Ω . Амплітуди всіх компонентів спектра, в тому числі і з частотою ω_0 , пропорційні значенням функцій Бесселя $J_k(m)$.

Формулу (5.21) можна у більш компактному вигляді. Дійсно враховуючи $(-1)^k J_k(m) = J_k(m)$, отримуємо:

$$S_{\text{ЧМ}}(t) = U_m \sum_{k=-\infty}^{\infty} J_k(m) \cos(\omega_0 + k\Omega)t. \quad (5.22)$$

Для побудови спектральних діаграм необхідно знання функцій Бесселя $J_k(m)$ за різних значень k і m . Ці відомості є у математичних довідниках. На рис. 5.6. наведено графіки функцій Бесселя при $k = 0, 1, 2, \dots, 7$. Значення функцій Бесселя, які відсутні на графіках, можна знайти за рекурентною формулою:

$$J_{k+1}(m) = \left(\frac{2k}{m}\right)J_k(m) - J_{k-1}(m).$$

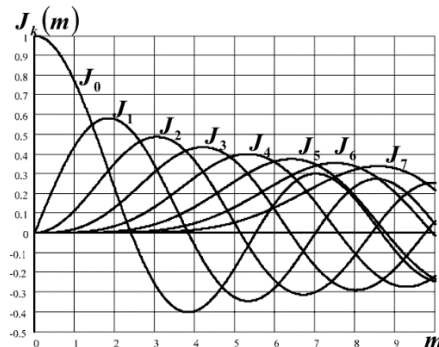


Рисунок 5.6 – Графіки функцій Бесселя

Аналіз графіків функцій Бесселя показує, що більший порядок k функції Бесселя, то за більших аргументів m спостерігається її максимум, проте за $k > m$ значення функцій Бесселя виявляються малою величиною. Отже, малими будуть відповідні складові спектра; ними можна знехтувати. Тому ширину спектра сигналів з кутовою модуляцією можна визначити за формулою:

$$\Delta\omega_{\text{ум}} \approx 2(m + 1)\Omega, \quad (5.23)$$

де Ω – частота модулюючого сигналу.

Звідси випливає, що порівняно з АМ–сигналами, смуга частот яких дорівнює 2Ω , для передачі сигналів з кутовою модуляцією потрібна смуга частот, у m разів більша. З іншого боку, саме широкосмуговість ЧС і ФМ сигналів забезпечує їхню велику стійкість до перешкод у порівнянні з АМ сигналами.

5.1.3 Квадратурна модуляція

Квадратурна модуляція дозволяє модулювати несучу частоту одночасно двома сигналами шляхом модуляції амплітуди несучої одним сигналом, і фази несучої іншим сигналом. Рівняння результуючих коливань амплітудно-фазової модуляції: $s(t) = u(t) \cos(\omega_0 t + \varphi(t))$.

Сигнал $s(t)$ зазвичай формують в дещо іншій послідовності, з урахуванням наступного виконання демодуляції. Представимо сигнал у вигляді суми двох АМ-коливань.

$$s(t) = u(t) \cos(\omega_0 t) \cos \varphi(t) - u(t) \sin(\omega_0 t) \sin \varphi(t).$$

При $a(t) = u(t) \cos \varphi(t)$ і $b(t) = -u(t) \sin \varphi(t)$, сигнали $a(t)$ і $b(t)$ можуть бути використані як модулюючі сигнали для несучих коливань $\cos(\omega_0 t)$ і $\sin(\omega_0 t)$, зсунутих по фазі на 90° відносно один одного:

$$s(t) = a(t) \cos(\omega_0 t) + b(t) \sin(\omega_0 t).$$

Отриманий сигнал називають **квадратурним** (quadrature), а спосіб модуляції – **квадратною модуляцією** (КАМ).

Спектр квадратурного сигналу може бути отриманий безпосередньо за рівнянням балансової модуляції для суми двох сигналів:

$$S(\omega) = \frac{1}{2} A(\omega + \omega_0) + \frac{1}{2} A(\omega - \omega_0) - \frac{1}{2} j B(\omega + \omega_0) + \frac{1}{2} j B(\omega - \omega_0).$$

Демодуляція квадратурного сигналу відповідно виконується множенням на два опорні коливання, зсунуті один відносно одного на 90° :

$$s_1(t) = s(t) \cos \omega_0 t = \frac{1}{2} a(t) + \frac{1}{2} a(t) \cos 2\omega_0 t + \frac{1}{2} b(t) \sin 2\omega_0 t,$$

$$s_2(t) = s(t) \sin \omega_0 t = \frac{1}{2} b(t) + \frac{1}{2} a(t) \sin 2\omega_0 t - \frac{1}{2} b(t) \cos 2\omega_0 t.$$

Низькочастотні складові $a(t)$ та $b(t)$ виділяються фільтром низьких частот. Як і при балансній амплітудній модуляції, для точної демодуляції сигналів потрібно дотримання частоти і початкової фази опорного коливання.

5.2 Формування та детектування модульованих сигналів

5.2.1 Формування та детектування сигналів амплітудної та односмугової амплітудної модуляції

Пристрої формування та демодуляції радіосигналів можуть бути різними в залежності від активних елементів, способу подачі на них несучої і модулюючого сигналу.

Розглянемо основні засади побудови модуляторів для АМ сигналів. При вхідному сигналі щодо невеликої потужності одним із методів формування АМ сигналів $S_{AM}(t)$ є операція перемноження двох коливань: інформаційного $s_{ПЕС}(t)$ і несучого $S_n(t)$, де як перемножувач може використовуватися спеціальна мікросхема.

Операція амплітудного детектування протилежна амплітудної модуляції. Якщо детектування АМ коливання проводиться без опорної напруги, то як перетворюючий використовується нелінійний елемент (НЕ). Другий необхідний елемент детектора – фільтр низьких частот (ФНЧ), який здійснює придушення спектральних складових кратних несучих частот (рис. 5.7 а).

Нелінійний елемент амплітудного детектора зазвичай виконується на напівпровідниковому діоді. Залежно від амплітуди АМ сигналу та ступеня нелінійності характеристик НЕ можливі два режими детектування: квадратичне детектування при малих амплітудах вхідного сигналу та лінійне детектування у режимі великих амплітуд

Поряд з амплітудним, використовується синхронне детектування, яке засноване на перемноженні АМ сигналу і коливань опорного генератора, що збігається за частотою і фазою з несучою АМ сигналу, з подальшим виділенням низькочастотних складових за допомогою ФНЧ. Головні труднощі при синхронному детектуванні полягають в отриманні синфазного з опорного несучого коливання. Таке коливання формується за допомогою системи фазового автопідстроювання частоти (ФАПЧ) (рис. 5.7. б).

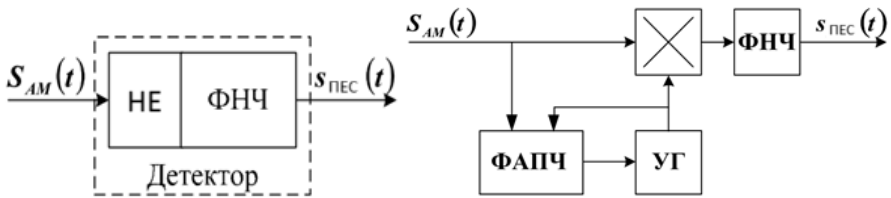


Рисунок 5.7 – а) Амплітудний детектор; б) Синхронний детектор

У системі ФАПЧ здійснюється порівняння фази вхідного сигналу та керованого генератора (КГ). За наявності відхилень здійснюється зміна фази КГ доти, доки не забезпечиться синфазність вхідного сигналу та КГ.

При синхронному детектуванні сигналів з балансною і односмуговою модуляцією з'являються важливі проблеми у отриманні синфазної опорної напруги. Це пов'язано з тим, що у спектрі цих сигналів несуча відсутня. Тут використовуються два технічні рішення.

У першому випадку разом з балансним або односмуговим модульованими сигналами передається так званий пілот-сигнал, що являє собою залишок несучої. Пілот-сигнал використовується у приймачі для системи ФАПЧ опорного генератора.

Друге рішення полягає в тому, що для детектування використовується високостабільний генератор несучої, що відрізняється за частотою від несучої.

Внаслідок несинхронності несучого та опорного колювання в системах передачі з односмуговою модуляцією спектр відновленого сигналу зміщується. Ці спотворення знижують якість передачі первинних сигналів. Проте, як свідчать експериментальні дослідження, невеликий частотний зсув помітного впливу якість не надає. При телефонному зв'язку абонент мало помічає зсув частот до 10 ... 20 Гц. При передачі радіомовних програм допустимим є зсув частот до 2 Гц. Приблизно такий самий зсув (± 1 Гц) не позначається як факсимільного зв'язку. Проте, звідси випливають дуже жорсткі вимоги до стабільності генераторного устаткування систем зв'язку з ОМ.

5.2.2 Формування та детектування сигналів кутової модуляції

Частотна модуляція (ЧМ) є основним видом модуляції у сучасних системах передачі інформації НВЧ діапазону, у тому числі системах супутникового радіозв'язку та телебачення. При ЧС забезпечується висока стійкість до перешкод і висока якість передачі інформації, допускається

можливість одночасної роботи в загальному каналі зв'язку великої кількості кореспондентів і реалізується більш повне використання за енергетичними показниками радіопередаючого пристрою в силу сталості амплітуди сигналу в порівнянні з амплітудною модуляцією.

Способи здійснення частотної та фазової модуляції можна розділити на дві групи: прямі та непрямі (рис. 5.8).

Прямий метод при ЧС означає безпосередній вплив на автогенератор або, точніше, на коливальну систему, що визначає частоту коливаль. Непрямий метод ЧС полягає у перетворенні фазової модуляції на частотну.

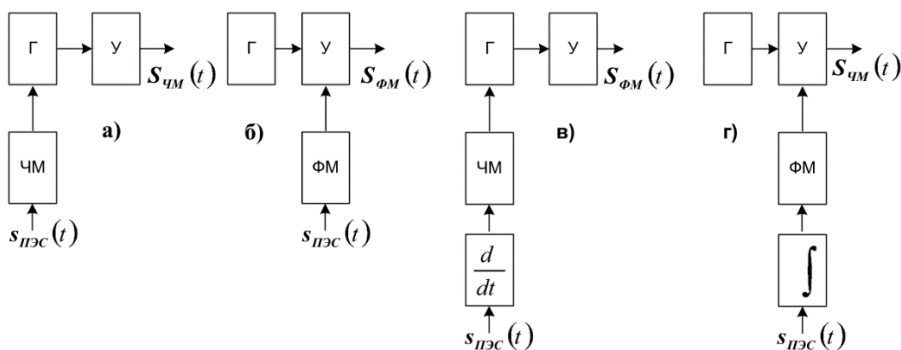


Рисунок 5.8 – Прямі (а, б) та непрямі (в, г) методи частотної та фазової модуляції

Прямий метод при ФМ означає вплив на високочастотний підсилювач або помножувач частоти, тобто електричні кола, що визначають фазу високочастотних коливаль. Непрямий метод ФМ полягає у перетворенні частотної модуляції на фазову.

Для перетворення фазової модуляції в частотну на вході фазового модулятора включається інтегратор (рис. 5.8 г). Для перетворення частотної модуляції у фазову на вході частотного модулятора включається диференціююча ланка кола (рис. 5.8 в).

Залежно від характеру перетворень розрізняють частотно-амплітудні, частотно-фазові та частотно-імпульсні детектори.

У частотно-амплітудних детекторах зміна частоти сигналу перетворюється на зміну амплітуди яка потім виділяється амплітудним детектором. Для того, щоб на виході детектора не виникли спотворення за рахунок можливих змін амплітуди вхідної напруги, перед детектуванням зазвичай проводять обмеження.

У частотно-фазових детекторах зміна частоти перетворюється на зміну фазового зсуву між двома напругами з наступним фазовим детектуванням.

Фазові детектори перетворюють вхідний фазомодульований сигнал вихідну напругу, що змінюється за законом модулюючого сигналу. Виявити фазовий зсув ФМ в сигналі можна шляхом порівняння з когерентним немодульованим коливанням, яке називають опорним.

Структурна схема фазового детектора аналогічна до схеми синхронного детектора (рис. 5.7). Усі фазові детектори розрізняються за типом використовуваного перемножувача, наявності або відсутності обмежувача та методів створення опорної напруги. Як перемножувачі можна використовувати будь-які нелінійні або параметричні елементи – діоди, транзистори, диференціальні та операційні підсилювачі з керованим зворотним зв'язком, спеціальні аналогові перемножувачі, ключові схеми та ін.

5.3 Маніпуляція сигналів

При дискретній зміні керуючого коливання модульовані параметри несучої будуть змінюватися стрибком. В цьому випадку замість терміну «модуляція» застосовується термін «маніпуляція», а саме коливання називається маніпульованим. Зокрема маніпуляція – це модуляція несучого коливання імпульсами постійного струму прямокутної форми.

Дискретне маніпулююче коливання може мати вигляд уніполярних або біполярних прямокутних імпульсів. Для опису двох можливих станів широко використовуються терміни «посилка» та «пауза». Ці стани зазвичай позначають символами +1 і -1 або 1 і 0.

5.3.1 Часові та спектральні характеристики амплітудно-маніпульованих сигналів.

Амплітудною маніпуляцією (АМн) називається процес зміни амплітуди несучого (високочастотного, маніпульованого) коливання відповідно до закону зміни амплітуди дискретного інформаційного (первинного електричного, маніпулюючого) сигналу.

Структурну схему отримання АМн сигналу можна представити як ключ, керований первинним сигналом $s_c(t)$, на вхід якого надходить несучий сигнал $S_n(t)$ (рис. 5.9).

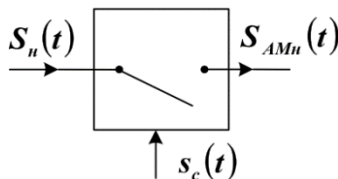


Рисунок 5.9 – Структурна схема амплітудного модулятора

Амплітудно-маніпульований сигнал має вигляд послідовності радіоімпульсів з прямокутною огибаючою (рис. 5.10. в).

При цьому первинний сигнал можна подати у вигляді відрізка ряду Фур'є:

$$S_c(t) = \sum_{k=1}^{\infty} A_k \cdot \cos(\Omega_k t + \phi_k) - \text{сигнал (рис. 5.10.а);}$$

$$S_H(t) = A_m \cdot \cos(\omega_H t + \phi_0) - \text{несучий сигнал (рис. 5.10.б).}$$

Результуючий сигнал з тривалістю інтервалів τ_H , що відповідають символам кодової комбінації (1 і 0 або +1 і -1), має на вигляд:

$$S_{AMH}(t) = \frac{1}{2} A_m \cdot [1 + x_c(t)] \cdot \cos(\omega_H t + \phi_0), \quad (5.24)$$

де $x_c(t)$ – нормована функція, що повторює закон зміни $s_c(t)$ (рис. 5.10. а) і набуває значень ± 1 .

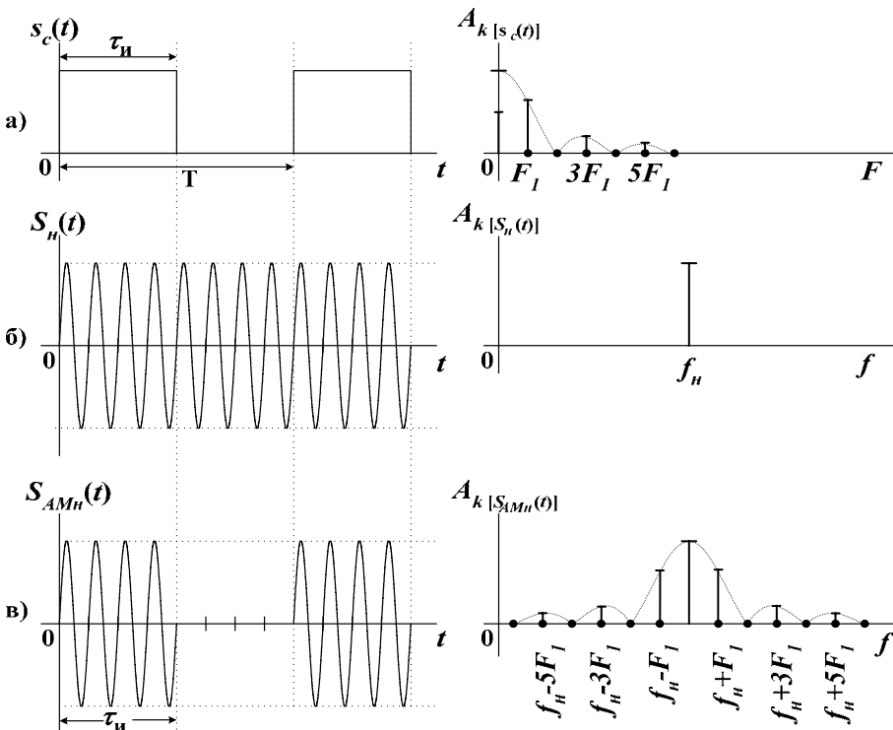


Рисунок 5.10 – Часові та спектральні характеристики формування АМн сигналу

Спектральний розклад періодичної послідовності АМн сигналів визначається наступним виразом:

$$S_{\text{АМн}}(t) = \frac{A_m \cdot \tau_u}{T} [1 + 2 \sum_{k=1}^{\infty} A_k \cdot \cos(k 2\pi F_1 t)] \cdot \cos(2\pi f_n t) = \frac{A_m \cdot \tau_u}{T} \cdot \cos(2\pi f_n t) + \quad (5.25)$$

$$+ \frac{A_m \cdot \tau}{T} \sum_{k=1}^{\infty} \left| \frac{\sin(k\pi F_1 \tau_u)}{k\pi F_1 \tau_u} \right| \times [\cos(f_n + kF_1)2\pi + \cos(f_n - kF_1)2\pi].$$

Спектр модульованого сигналу містить у своєму складі складову з амплітудою $A_m \cdot \tau_u / T$ на несучій частоті f_n та дві симетричні бічні смуги з частотами складових $(f_n + kF_1)$; $(f_n - kF_1)$ – і амплітудами $\frac{A_m \cdot \tau}{T} \cdot \frac{\sin(k\pi F_1 \tau_u)}{k\pi F_1 \tau_u}$.

Для періодичних сигналів – спектр дискретний, а при випадковому поступанні кодових символів (неперіодичних сигналів) – спектр стає суцільним.

Ширина спектру АМн коливання: $\Delta F_{\text{АМн}} = 2kF_1$,

де k – номер гармоніки, що враховується; $F_1 = 1/T$ – частота першої гармоніки інформаційного сигналу.

У реальних каналах ширину спектра беруть з урахуванням третьої або п'ятої гармоніки, наприклад, при необхідності передати цифровий сигнал зі швидкістю $V = 50$ Бод, ширина спектра $\Delta F_{\text{АМн}} = 2 \cdot 5 \cdot F_1 = 5 \cdot V = 250$ Гц.

В даний час двійкова амплітудна маніпуляція використовується в низькошвидкісних системах передачі інформації, багатоканальних системах зв'язку з часовим поділом, в радіолокаційних системах, а також в ряді оптичних систем.

5.3.2 Часові та спектральні характеристики частотно-маніпульованих сигналів.

При частотній маніпуляції (ЧМн) частота високочастотного коливання змінюється стрибком на величину $\pm \Delta f_m$ щодо несучої f_n (рис. 5.11. а). Таким чином, на виході ЧМн здійснюються коливання на частотах f_1 і f_2 .

Різницю частот $f_1 - f_2 = \Delta f_{\text{дв}}$ називають частотним зсувом. Максимальне відхилення частоти Δf_m від несучої називають девіацією. Відношення девіації частоти Δf_m до частоти маніпулюючого коливання F називається індексом частотної маніпуляції. Індекс ЧМн прямо пропорційний девіації і обернено пропорційний частоті інформаційного сигналу: $m_{\text{ЧМн}} = \Delta f_m / F$.

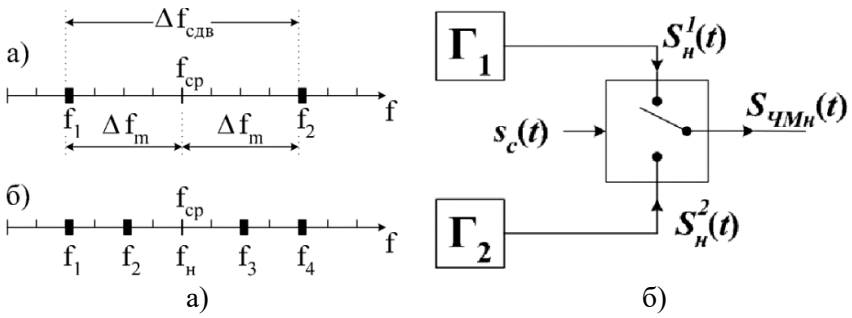


Рисунок 5.11 – а) Параметри сигналів ЧМн; б) Структурна схема формування ЧМн коливань з розривом фази

Розрізняють частотну маніпуляцію: з **розривом фази та без розриву фази**. Загальний вигляд ЧМн сигналу з розривом фази можна подати у вигляді суми двох АМн сигналів з різними несучими частотами f_1 і f_2 . Технічно такий вид маніпуляції реалізується за допомогою двох генераторів (рис. 5.11. б), що управляються ключем під впливом інформаційного сигналу: $S_{\text{ЧМн}}(t) = S_{1\text{АМн}}(t) + S_{2\text{АМн}}(t)$.

Формування ЧМн сигналу із розривом фази показано на рис. 5.12.

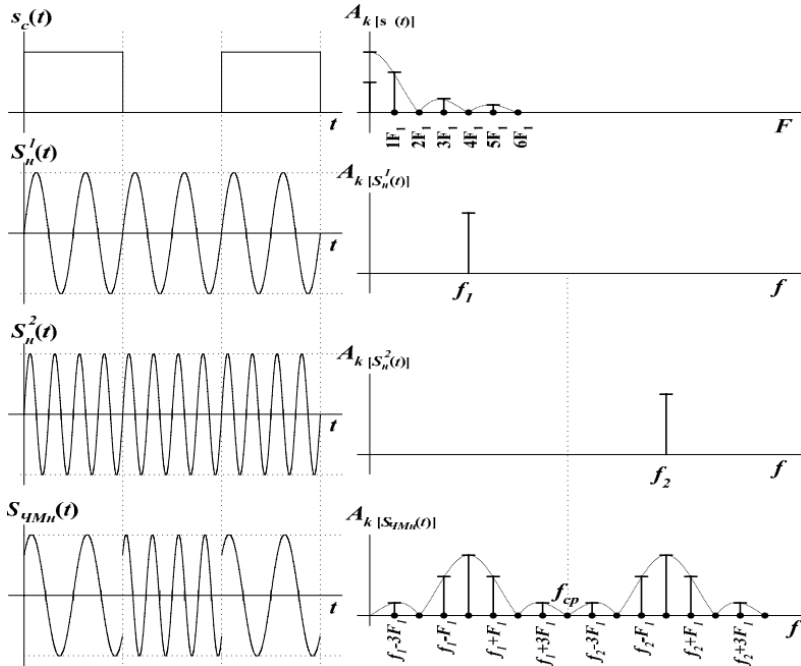


Рисунок 5.12 – Формування ЧМн сигналу із розривом фази

З рис. 5.12 видно, що ширина спектру ЧМн сигналу відрізняється від спектру сигналу АМн на величину $2\Delta f_m$: $\Delta F_{\text{ЧМн}} = 2kF_1 + 2\Delta f_m$, де k – номер гармоніки, що враховується.

Наприклад, при необхідності передати цифровий сигнал зі швидкістю $V = 75$ біт/с, $\Delta f_m = 250$ Гц, $k = 3$, ширина спектру $\Delta F_{\text{ЧМн}} = 2 \cdot 3 \cdot (75/2) + 2 \cdot 250 = 725$ Гц.

Загальний вигляд ЧМн сигналу без розриву фази (рис. 5.13.) можна записати у вигляді:

$$S_{\text{ЧМн}}(t) = A_m \cos[\omega_n t + \Delta\varphi(t)],$$

де $\Delta\varphi(t)$ – приріст фази, обумовлений збільшенням частоти $\Delta\omega(t)$.

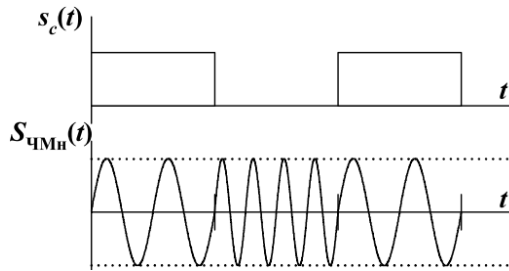


Рисунок 5.13 – Часові характеристики формування ЧМн коливань без розриву фази

Цей вид маніпуляції передбачає використання одного джерела коливань, частота якого змінюється у вигляді керованої реактивності (у разі фаза змінюється безупинно – без розриву).

5.4 Імпульсно-модульовані сигнали

В імпульсній модуляції як носій модульованих сигналів використовуються послідовності імпульсів, як правило – прямокутних. У бездротових системах передачі даних (в радіозв'язку) ці послідовності заповнюються високочастотними коливаннями, створюючи тим самим подвійну модуляцію. Як правило, ці види модуляції застосовуються при передачі дискретизованих даних. Для прямокутних імпульсів найбільш широко використовуються **амплітудно-імпульсна (АІМ)** та **широтно-імпульсна (ШІМ)** модуляція.

Амплітудно-імпульсна модуляція (АІМ) полягає у зміні збільшення амплітуди імпульсів пропорційно функції керуючого сигналу при постійній тривалості імпульсів та періоді їхнього формування:

$$U(t) = U_o + k \cdot s(t), \quad \tau_i = \text{const}, \quad T = \text{const}. \quad (5.26)$$

Спектр АІМ розглянемо з прикладу модулювання однотонального сигналу $s(t)$. Напишемо рівняння модульованого сигналу у наступній формі:

$$u(t) = (1 + m \cos \Omega t) \cdot f(t), \quad (5.27)$$

де $f(t)$ – періодична послідовність прямокутних імпульсів із частотою ω_o , яку можна апроксимувати рядом Фур'є (без урахування фази): $f(t) = U_o + U_n \cos n\omega_o t$.

Отримуємо:

$$\begin{aligned} u(t) &= (1 + m \cos \Omega t) U_o + U_n \cos n\omega_o t \cdot (1 + m \cos \Omega t) = \\ &= U_o + U_o m \cos \Omega t + U_n \cos n\omega_o t + \\ &+ 0.5m U_n \cos (n\omega_o + \Omega)t + 0.5m U_n \cos (n\omega_o - \Omega)t. \end{aligned} \quad (5.27)$$

Форма спектру, у початковій частині спектрального діапазону, наведено на рис. 5.14. б) В цілому, спектр нескінченний, що визначається нескінченністю спектра прямокутних імпульсів. При кожній гармоніці $n\omega_o$ спектра прямокутних імпульсів з'являються бічні складові $n\omega_o \pm \Omega$, відповідні спектру моделюючої функції (при багатотональному сигналі – бічні смуги спектрів). При додатковому високочастотному заповненні імпульсів весь спектр зміщується в область високих частот на частоту заповнення.

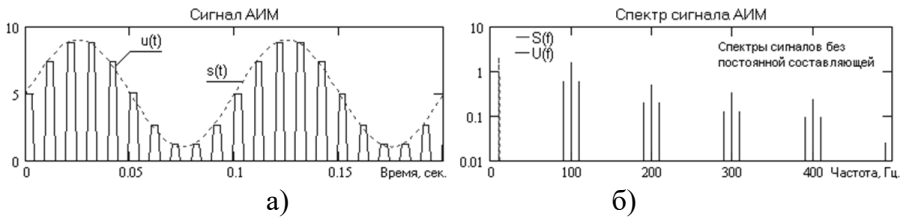


Рисунок 5.14 – а) Графік АІМ сигналу; б) спектр АІМ сигналу без постійної складової

Широтно-імпульсна модуляція

Широтно-імпульсна модуляція (ШІМ, в англійській термінології pulse width modulation, PWM), яку іноді називають модуляцією за тривалістю імпульсів, полягає в управлінні тривалістю імпульсів пропорційно функції керуючого сигналу при постійній амплітуді імпульсів і періоді слідування по пульсів:

$$\tau(t) = t_0 + k \cdot s(t), \quad U = \text{const}, \quad T = \text{const}. \quad (5.28)$$

Розглянемо виконання ШІМ у найпростішому варіанті на прикладі гармонійного коливання, наведеного на рис. 5.15.

Переданий сигнал дискретизується, при цьому має значення як інтервал дискретизації, так і кількість рівнів квантування. При передачі даних прямокутні імпульси починаються в моменти дискретних відліків даних, а тривалість імпульсів встановлюється пропорційно значенню відліків, при цьому максимальна тривалість імпульсів не повинна перевищувати інтервалу дискретизації даних. При цьому число рівнів квантування гармоніки прийнято рівним 8. Природно, що при малому числі рівнів квантування похибка відновлення вихідного гармонійного сигналу дуже велика.

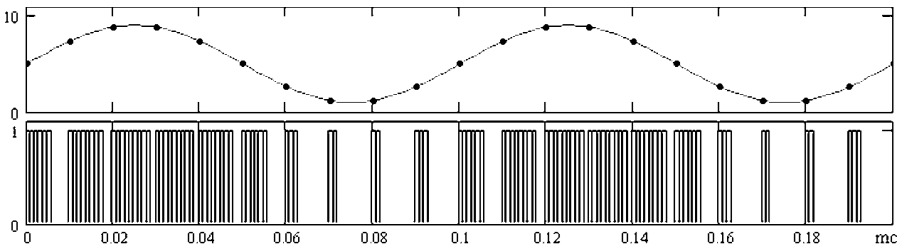


Рисунок 5.15 – Широтно-імпульсна модуляція

Принагідно зауважимо, що широтно-імпульсна модуляція з подальшим виділенням постійної складової може дуже ефективно використовуватися (і використовується) для стеження за середнім рівнем сигналу та автоматичного регулювання його динамічного діапазону, як, наприклад, у системах встановлення гучності звуку та яскравості кольорів та зображення загалом у сучасних телевізійних установках.

Часова імпульсна модуляція (ЧІМ) являє собою девіацію імпульсів по часовій осі за законом модулюючого сигналу, і по суті аналогічна кутовий модуляції гармонійної несучої. Вона також може бути фазовою (ФІМ) або частотною (ЧІМ).

Кодово-імпульсна модуляція полягає в тому, що в точках дискретизації модулюючого сигналу проводиться квантування його значень і кодування квантованих значень, як правило, у двійковій системі обчислення. Кодовані значення потім передаються за допомогою відповідної кодової послідовності стандартних символів.

5.5 Кодування сигналів

При кодуванні елементів повідомлення перетворюються у відповідні їм числа (кодові символи). Кожному елементу повідомлення привласнюється певна сукупність кодових символів, яка називається кодовою комбінацією.

Сукупність кодових комбінацій, що позначають дискретні повідомлення, утворює код. Правило кодування може бути виражене кодовою таблицею, у якій наводяться алфавіт кодованих повідомлень і відповідні їм кодові комбінації. Множина можливих кодових символів називається кодовим алфавітом, а їх кількість m — основою коду.

У загальному випадку при основі коду m правила кодування N елементів повідомлення зводяться до правила запису N різних чисел в m -ій системі числення. Число розрядів n , що утворюють кодову комбінацію, називається значністю коду, або довжиною кодової комбінації. Залежно від системи числення, що використовується при кодуванні, розрізняють двійкові й m -і (недвійкові) коди.

Коди, у яких усі комбінації мають однакову довжину, називають рівномірними. Для рівномірного коду число можливих комбінацій рівне m^n . Прикладом такого коду є п'ятизначний код Бодо, що містить п'ять двійкових елементів ($m=2, n=5$).

Число можливих кодових комбінацій рівне $2^5=32$, що достатньо для кодування всіх букв алфавіту. Застосування рівномірних кодів спрощує побудову автоматичних літеродрукувальних пристроїв і не вимагає передачі символів розділу між кодовими комбінаціями.

В нерівномірних кодах кодові комбінації відрізняються один від одного як взаємним розташуванням символів, так і їх кількістю. Це призводить до того, що різні комбінації мають різну тривалість. Типовим прикладом нерівномірних кодів є код Морзе, у якому символи 0 і 1 використовуються тільки у двох комбінаціях — як одиночні (1 і 0) або як потрійні (111 і 000). Сигнал, що відповідає одній одиниці, називається крапкою, а трьома одиницями — тире. Символ 0 використовується як знак, що відокремлює крапку від тире, крапку від крапки й тире від тире.

Сукупність 000 використовується як знак розділу між кодовими комбінаціями.

По завадостійкості коди ділять на прості (примітивні) і коригувальні. Коди, у яких усі можливі кодові комбінації використовуються для передачі інформації, називаються простими, або кодами без надмірності. У сучасних системах передачі дискретних повідомлень кодеком називаються пристрої, що перетворюють повідомлення в код (кодер) і код у повідомлення (декодер).

При прийомі сигналів для відновлення аналогового сигналу необхідно перетворити цифровий сигнал (послідовність двійкових імпульсів) квантований АІМ сигнал (таке перетворення називається декодуванням) і потім здійснити операцію демодуляції, тобто. виділення з сигналу АІМ аналогового сигналу $s(t)$.

При використанні ІКМ виконуються такі перетворення аналогового сигналу: у пункті передачі амплітудно-імпульсна модуляція, квантування та кодування; у пункті прийому декодування та демодуляція квантованого АІМ

сигналу. Отриманий прийому аналоговий сигнал відрізняється від переданого, оскільки утворюється з квантованих імпульсів, амплітуди яких рівні миттєвим значенням сигналу $s(t)$, а найближчим дозволеним значенням.

Таким чином, операція квантування вносить у процес передачі сигналу неперекривну помилку, яка тим менша, чим більше рівнів квантування.

До складу декодера входить перетворювач послідовного коду в паралельний (рис. 5.16), на виходах якого з'являється набір одиниць та нулів, що відповідає прийнятій кодовій комбінації. Кожна одиниця (струмовий імпульс) надходить на вхід суматора з вагою, де збільшується в 2^k разів. На виході суматора виникає імпульс, амплітуда якого визначається кодовою комбінацією на вході декодера. Наприклад, при проходженні кодової комбінації 0100110 на перший, четвертий, п'ятий і сьомий входи суматора напруга не подається (безструмові імпульси), а на другий, третій і шостий входи подається напруга, яка збільшується відповідно в 2^1 , 2^2 і 2^5 разів. На виході суматора виникає напруга, пропорційна $2^1 + 2^2 + 2^5 = 38$, тобто. квантований АІМ сигнал.

На наступному кроці необхідно з відлікових значень, наприклад струму отримати безперервний струм. Зробити це нам допоможе звичайний конденсатор невеликої ємності, який при короткочасному впливі на нього струму (тобто відлікового значення) миттєво зарядиться і утримуватиме заряд до наступного короткочасного впливу.

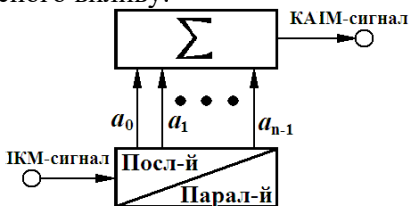


Рисунок 5.15 – Декодер ІКМ-сигналу.

Зазначимо ще раз, що відновлена таким шляхом крива безперервного струму дещо відрізнятиметься від тієї, яка була отримана, наприклад, на клемках мікрофона: вона матиме плоскі сходинки між відліковими значеннями. Можна сказати, що процес взяття відлікових значень та подальшого відновлення безперервної кривої струму мікрофона супроводжується специфічними спотвореннями, які можуть вплинути на якість відтворення звуку. Однак на практиці для відновлення струму використовують не конденсатор, а складніші схеми, що роблять форму відновленого струму схожою на форму вихідного струму і тим самим зводять нанівець дії зазначених спотворень.

Тема 6. Проходження сигналів через канали зв'язку

Імпульсні і частотні характеристики систем. Перетворення випадкових сигналів у лінійних та нелінійних каналах з постійними параметрами. Проходження сигналів через канали зв'язку з випадковими параметрами. Моделі неперервних каналів. Моделі дискретних каналів.

6.1 Імпульсна перехідна функція

Будемо називати канал зв'язку ідеальною системою, якщо вона: а) фізично здійсненна, б) стійка, в) має постійні параметри і г) лінійна. Основні властивості такої ідеальної системи описуються її **імпульсною перехідною функцією, або ваговою функцією**, яка представляє реакцію системи на збурення у вигляді дельта-функції.

Нехай на вхід системи надходить деяка гладка функція $x(t)$, а на виході спостерігається гладка функція $y(t)$. Імпульсна перехідна функція системи визначається рівнянням:

$$h(t) = y(t) \text{ при } h(t) = \delta(t), \quad (6.1)$$

Причому відлік часу t починається з моменту, коли на вхід системи подається дельта-функція.

Реакція $y(t)$ системи на довільний вхідний процес $x(t)$ визначається інтегралом згортки (інтегралом Дюамеля):

$$y(t) = \int_{-\infty}^{\infty} h(t)x(t - \tau)d\tau. \quad (6.2)$$

Інакше кажучи, реакція системи $y(t)$ це зважена лінійна сума всіх минулих і майбутніх значень вхідного процесу $x(t)$.

Фізично здійсненна система не може реагувати на збурювання доти, поки воно не поступило на вхід системи. Це означає, що

$$h(\tau) = 0 \text{ при } \tau < 0. \quad (6.3)$$

Отже, для фізично здійснених систем нижня межа інтегрування в рівнянні (6.2) дорівнює нулю, а не мінус нескінченності. Імпульсна характеристика, що відповідає виразу (6.3) називається казуальною імпульсною характеристикою.

Говорять, що фізична система має **постійні параметри**, якщо її імпульсна перехідна функція не залежить від моменту поступання змушуючої сили на вхід системи:

$$h(t, \tau) = h(\tau) \text{ при } -\infty < t \quad (6.4)$$

Якщо параметри системи постійні, то стаціонарний вхідний процес завжди генерує стаціонарний процес на виході системи (після загасання перехідних процесів, що виникають при включенні системи).

Фізична система називається **стійкою**, якщо при довільній припустимій обмеженій функції на її вході функція на виході також є обмеженою. Ця умова виконується, якщо

$$\int_{-\infty}^{\infty} |h(t)| d\tau < \infty \quad (6.5)$$

Розглянуті тут співвідношення між процесами на вході й виході фізичних систем слушні лише у випадку стійкої системи

Лінійна система має властивості адитивності і однорідності. Нехай вхідні процеси x_1 і x_2 генерують вихідні процеси y_1 і y_2 відповідно. Система називається адитивною, якщо вхідний процес $(x_1 + x_2)$ генерує на виході системи процес $(y_1 + y_2)$, і однорідною, якщо вхідному процесу sx_1 відповідає вихідний процес sy_1 (тут s – довільна постійна). Все це означає, що вагова функція $h(\tau)$ не залежить від процесу $x(t)$, тобто:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t - \tau)d\tau \quad (6.6)$$

Якщо система лінійна, то випадковий вхідний процес із гаусівським розподілом імовірностей породжує процес на виході також з гаусівським розподілом імовірностей. У практичних ситуаціях лінійність відноситься до числа найбільш рідко виконуваних властивостей. Однак, якщо досліджувана система не є сильно нелінійною, то методи кореляційного і спектрального аналізу приведуть у більшості випадків до цілком осмислених результатів, що описують найкращі (у середньоквадратичному сенсі) лінійні наближення для досліджуваних систем.

6.2 Частотні характеристики

Динамічні властивості фізичних систем прийнято описувати, як правило, не самою імпульсною перехідною функцією $h(\tau)$, а деяким її лінійним перетворенням, причому вид перетворення залежить від конкретного завдання. Однак у випадку ідеальної системи зручніше за все користуватися перетворенням Фур'є, яке дозволяє безпосередньо описати динамічні характеристики системи в частотній області. Перетворення Фур'є імпульсної перехідної функції $h(\tau)$, що задовольняє умові $h(\tau) = 0$ при $\tau < 0$ має вигляд:

$$H(f) = \int_0^{\infty} h(\tau) \exp(-i2\pi f \tau) d\tau \quad (6.7)$$

і називається **частотною характеристикою системи**. У загальному випадку частотна характеристика є комплексною функцією частоти, так що її можна представити у вигляді

$$H(f) = H_R(f) + iH_I(f), \quad (6.8)$$

де $H_R(f)$ і $H_I(f)$ – дійсна і уявна частини функції $H(f)$ відповідно:

$$H_R(f) = \int_0^{\infty} h(\tau) \cos(2\pi f \tau) d\tau, \quad H_I(f) = \int_0^{\infty} h(\tau) \sin(2\pi f \tau) d\tau$$

Надалі ми також будемо користуватися полярною формою запису для частотної характеристики $H(f)$:

$$|H(f)| = \left[H_R^2(f) + H_I^2(f) \right]^{1/2}, \quad \varphi(f) = \arctg [H_I(f) / H_R(f)] \quad (6.9)$$

Модуль $|H(f)|$ звичайно називається амплітудною характеристикою, а аргумент $\varphi(f)$ – фазовою характеристикою. Частотна характеристика має очевидну фізичну інтерпретацію: якщо на вхід ідеальної системи надходить гармонійне коливання із частотою f , то на виході теж буде спостерігатися гармонійне коливання із частотою f . Однак амплітуда вихідного сигналу буде в загальному випадку відмінна від амплітуди вхідного сигналу й, крім того, вихідний сигнал буде зсунутий по фазі щодо вхідного. Інакше кажучи, вхідний і вихідний процеси визначаються в цьому випадку рівностями $x(t) = X \sin 2\pi f t$, $y(t) = Y \sin(2\pi f t - \theta)$.

Відношення амплітуд вихідного й вхідного сигналів задає амплітудну характеристику, а зсув по фазі між $x(t)$ і $y(t)$ – фазову характеристику системи:

$$|H(f)| = |Y(f)| / |X(f)|, \quad \varphi(f) = \theta(f).$$

У теорії керування частотну характеристику (6.7) називають передатною функцією. Однак, строго кажучи, передатна функція визначається перетворенням Лапласа імпульсної перехідної функції, тобто має вигляд:

$$H(p) = \int_0^{\infty} h(\tau) \exp(-p\tau) d\tau, \quad p = a + jb,$$

де a (тобто дійсна частина комплексного числа p) не обов'язково дорівнює нулю.

6.3 Проходження сигналів через канали з детермінованими характеристиками

Передача сигналів по реальних каналах зв'язку завжди супроводжується змінами (перетвореннями) цих сигналів, у результаті чого прийняті сигнали відрізняються від переданих.

Відмінності ці обумовлені, насамперед, лінійними й нелінійними перетвореннями вхідних сигналів, а також наявністю шумів у каналі, що існують найчастіше незалежно від переданих сигналів. З погляду передачі інформації по каналу важливим є поділ перетворень сигналу на оборотні й необоротні.

Оборотні перетворення не спричиняють втрати інформації. При **необоротних перетвореннях** втрати інформації неминучі. Для оборотних перетворень сигналу часто використовують термін «перекручування», а необоротні перетворення називають завадами (адитивними і неадитивними).

Прикладом найпростішого детермінованого оборотного перетворення вхідного сигналу $X(t)$, яке не міняє його форми, є

$$Y(t) = kX(t-\tau). \quad (6.10)$$

В цьому випадку вихідний сигнал каналу $Y(t)$ відрізняється від вхідного лише відомим масштабом k , який легко компенсується відповідним посиленням або ослабленням сигналу й постійною затримкою в часі τ . Вона найчастіше невелика. По суті, лише при створення зв'язку в масштабах космосу або при дуже великій кількості реактивних елементів в лінії зв'язку затримка може виявитися відчутною.

Проходження випадкових сигналів через детерміновані лінійні кола. При стаціонарному вхідному процесі й вихідний процес виявляється стаціонарним, тоді кореляційну функцію вихідного сигналу можна записати як

$$B_Y(\tau) = \int_{-\infty}^{\infty} B_X(\nu)B_h(\tau-\nu)d\nu, \quad (6.11)$$

де $B_X(\nu)$ і $B_h(\tau-\nu)$ – кореляційні функції вхідного сигналу й імпульсної функції відповідно.

З теорії відомо, що перетворення Фур'є від згортки двох функцій дорівнює добутку перетворень Фур'є від цих функцій. Застосувавши це правило до (6.11), одержимо просте співвідношення між спектральними

щільностями стаціонарних процесів на вході й на виході лінійного кола з постійною передатною функцією (частотною характеристикою) $H(f)$:

$$S_Y(f) = S_X(f) / |H(f)|^2. \quad (6.12)$$

Звідси випливає, що функція кореляції і спектр процесу на виході кола повністю визначаються функцією кореляції або спектром процесу на вході й АЧХ кола, тобто не залежать ні від розподілу ймовірностей вхідного процесу, ні від фазочастотної характеристики кола.

Проходження вузькосмугових випадкових сигналів через лінійні смугові кола. Відносно вузькосмугові процеси (тобто такі, у яких ширина спектра значно вужче середньої частоти) зручно представляти у квазігармонійній формі. Якщо середня частота ω_0 задана, то такий вузькосмуговий сигнал повністю визначається своєю комплексною огинаючою $A(t)$ або її дійсною і уявною частинами (квадратурними складовими) $A_c(t)$ і $A_s(t)$, які є низькочастотними процесами. Вираз для низькочастотного еквівалента вузькосмугового сигналу $\dot{A}_x(t) = \exp(-j\omega_0 t)$ має спектр: $S_A(\omega) = SX(\omega + \omega_0)$.

Основна частина реальних неперервних каналів зв'язку відносяться до лінійних і вузькосмугових, тому сигнали на їхньому виході можуть розглядатися як реакція на вузькосмуговий сигнал $X(t)$ смугового фільтра з передатною функцією $H(f)$.

Проходження випадкових сигналів через нелінійні кола.

Обмежимося розглядом тільки безінерційних нелінійних систем з регулярними параметрами, у яких вхід і вихід зв'язані деякою нелінійною залежністю, що називається характеристикою системи: $y(t) = \varphi[x(t)]$.

Цим співвідношенням досить точно може бути охарактеризована робота ланок реальних каналів зв'язку, наприклад, демодуляторів, модуляторів тощо.

Перетворення $x(t) \rightarrow y(t)$, як правило, однозначне, що не завжди можна сказати про зворотнє перетворення $y(t) \rightarrow x(t)$ (наприклад, квадратичний ланцюг з характеристикою $y = kx^2$). У силу незастосовності суперпозиції до нелінійних систем розгляд складного впливу (наприклад, суми детермінованого й випадкового доданків) не можна звести до розгляду проходження кожної зі складових окремо.

При нелінійних перетвореннях виникає трансформація (зміна) спектра вхідного впливу. Так, якщо на вхід нелінійної системи впливає суміш регулярного сигналу й адитивного шуму $X(t) = u(t) + N(t)$ у вузькій смузі частот F_c , що групується близько середньої частоти f_0 , то в загальному випадку на виході будуть присутні складові комбінаційних частот трьох видів, що групуються близько частот nf_0 ($n=0, 1, \dots$): продукти биття складових вхідного

сигналу між собою (схс), продукти биття складових вхідного шуму (шхш); продукти биття сигналу й шуму (схш). Розділити їх на виході системи зазвичай неможливо.

Якщо відомі характеристика $y=\varphi(x)$ нелінійної системи й двовимірна функція розподілу вхідного впливу $w(x_1, x_2; t_1, t_2)$, то основні статистичні характеристики вихідного процесу, у принципі, завжди можна визначити. Так, математичне очікування відгуку

$$\overline{Y(t)} = \overline{\phi(X)} = \int_{-\infty}^{\infty} \phi(x)w(x)dx \quad (6.13)$$

а його кореляційна функція

$$B_Y(t, \tau) = \int_{-\infty}^{\infty} \left[\phi(x_1(t)) - \overline{Y(t)} \right] \times \left[\phi(x_1(t - \tau)) - \overline{Y(t - \tau)} \right] \times \\ \times w(x_1, x_2, t, t + \tau) dx_1 dx_2 \quad (6.14)$$

Перетворенням Фур'є по (6.14) можна знайти і енергетичний спектр.

Проходження сигналів через випадкові канали зв'язку. Крім детермінованих перетворень сигналу, в окремих ланках каналу (зокрема, у лінії зв'язку) мають місце випадкові перетворення сигналу. У найпростішому випадку це перетворення зводиться до підсумовування сигналу з незалежним від нього випадковим процесом, що називається адитивною завадою або адитивним шумом. У більш складних каналах до цього додаються випадкові зміни параметрів каналу, в результаті яких навіть під час відсутності адитивних завад прийнятий сигнал не визначається однозначно переданим.

Адитивні завади в каналі зв'язку викликаються досить різними причинами і можуть приймати всілякі форми, індивідуальні реалізації яких важко врахувати. Саме ці завади частіше викликають необоротні перетворення переданих сигналів.

6.4 Моделі неперервних каналів

Неперервними називаються канали, вхідні та вихідні сигнали яких набувають довільних значень з деякого інтервалу.

Неперервні канали можна класифікувати по вигляду перешкод і характеру перетворення вхідного сигналу $S_p(t, \lambda_0)$ в корисний прийнятий $S_p(t, \lambda)$. Якщо обмежитися припущенням, що у каналі діє адитивний нормальний білий шум $n(t)$, то неперервні канали поділяються по вигляду перетворення $S_p(t, \lambda_0)$ в $S_p(t, \lambda)$, тобто по вигляду спотворень сигналу. У більшості радіотехнічних систем випромінювані сигнали є вузькосмуговими

$S_p(t, \lambda_0) = A(t) \cos[\omega_0 t + \varphi(t)]$, де $A(t)$ і $\varphi(t)$ – функції, що відображають закони амплітудної та кутової модуляції; ω_0 – несуча частота сигналу.

Спотворення випромінюваного сигналу $S_p(t, \lambda_0)$ прийнято розглядати окремо для однопроменевих та багатопроменевих каналів. В однопроменевих каналах електромагнітні коливання поширюються по одному шляху. Однопроменевими каналами є лінії зв'язку на відстанях прямої видимості: лінії ближнього радіозв'язку на коротких і ультракоротких хвилях, лінії зв'язку Земля-повітря, повітря-Земля, повітря-повітря тощо.

Прийнятий корисний сигнал по відношенню до випромінюваного характеризується додатковими параметрами: випадковим ослабленням $\alpha(t)$, середнім часом запізнення τ_3 , доплерівським зміщенням частоти Ω і випадковою початковою фазою θ і може бути записаний у вигляді:

$$S_p(t, \lambda) = \alpha(t) A(t - \tau_3) \cos[(\omega_0 + \Omega)t + \varphi(t - \tau_3) - \theta]. \quad (6.15)$$

Ефект Д'оплера — явище зміни частоти хвилі, яку реєструє приймач, викликане переміщенням джерела або приймача. Ефект названий на честь австрійського фізика Крістіана Доплера. Виходячи з власних спостережень за хвилями на воді, Доплер припустив, що подібні явища відбуваються в повітрі з іншими хвилями.

Залежно від того, чи відомі на приймальній стороні значення додаткових параметрів $\alpha(t)$, τ_3 , Ω , θ можна виділити кілька моделей неперервних каналів.

Гауссівським каналом називається канал, в якому діє адитивний нормальний білий шум, а спотворення корисного сигналу несуттєві і можуть бути компенсовані. Компенсація спотворень можлива, якщо на приймальній стороні додаткові параметри повністю відомі або вимірюються досить точно. Тому вважається, що $S_p(t, \lambda_0) = S_p(t, \lambda)$. Вихідний сигнал гауссівського каналу:

$$Y_p(t) = S_p(t, \lambda_0) + n(t). \quad (6.16)$$

Тобто, вихідний сигнал подається як сума корисного сигналу і нормального білого шуму $n(t)$.

Гауссівський канал з невідомою фазою сигналу визначається параметрами τ_3 , Ω , $\alpha(t) = \alpha$, які є постійними та відомими. Фаза θ вважається рівномірно розподіленою величиною в інтервалі $[0, 2\pi]$. Така модель добре описує процеси в лініях радіозв'язку на відстанях прямої видимості.

Канал з амплітудними завмираннями є подальшим ускладненням каналу з невідомою фазою у припущенні, що $\alpha(t)$ – випадкова функція часу. Вихідний корисний сигнал каналу із завмираннями:

$$S_p(t, \lambda) = A(t) \cos[(\omega_0)t + \varphi(t) - \theta]. \quad (6.17)$$

Випадкова функція $\alpha(t)$ перемножується із сигналом і тому називається мультиплікативною перешкодою, яку можна розглядати як функцію, що модулює по амплітуді випромінюваний сигнал. Модуляція призводить до розширення спектра прийнятого сигналу щодо спектра випромінюваного сигналу. Тому такий канал називають каналом із розсіюванням енергії за частотою.

За часом кореляції мультиплікативні перешкоди поділяються на повільні та швидкі. Про повільні завмирання говорять у разі, якщо час кореляції $\alpha(t)$ значно перевищує інтервал спостереження сигналу. Причинами повільних замирань є зміни властивостей середовища поширення радіохвиль залежно від метеорологічних умов, часу доби, року, сонячної активності тощо.

Швидка мультиплікативна перешкода має час кореляції менше, ніж інтервал спостереження сигналу. Основною причиною швидких замирань є наявність багатьох шляхів, якими поширюються електромагнітні хвилі.

Багатопроменеве поширення виникає при передачі інформації на далекій відстані при відображенні радіохвиль від протяжних поверхонь суші та моря, при відображенні від іоносфери та тропосфери. Через різні шляхи поширення часи запізнення окремих сигналів, що приймаються, різні. Тому багатопроменеві канали називають каналами з розсіюванням енергії в часі.

Результуючий сигнал на виході багатопроменевого каналу

$$S(t, \lambda) = \sum_{i=1}^l S_{pi}(t, \lambda_{0i}), \quad (6.18)$$

де $S_{pi}(t, \lambda_{0i}) = \alpha_i(t) A(t) \cos[(\omega_0 + \Omega)t + \varphi(t) - \theta]$.

При великій кількості шляхів можна вважати, що $S_p(t, \lambda)$ є реалізацією нормального ВП. Зазвичай середнє значення процесу дорівнює нулю, тоді моделі багатопромених каналів відрізняються по вигляду КФ.

Таким чином, неперервний канал вважається заданим, якщо вказані потужність сигналів, смуга частот, дано опис моделей перешкод та спотворень сигналів.

6.5 Моделі дискретних каналів

Дискретними називаються канали, вхідні та вихідні сигнали яких приймають кінцеве число миттєвих значень. Поняття дискретного каналу природно виникає під час передачі дискретних повідомлень і визначається як сукупність технічних засобів, включених між кодером і декодером каналу.

Перехід від дискретних сигналів до неперервних здійснюється на передаючій стороні, при маніпуляції параметрами неперервної несучої. На приймальній стороні дискретні сигнали з'являються виході першої вирішальної схеми (демодулятора).

Властивості дискретного каналу визначаються неперервним каналом та структурою модему. Дискретний канал задається множиною вхідних $\{s_i\}$, $i = \overline{1, L_s}$, і вихідних $\{y_j\}$, $j = \overline{1, L_y}$, символів (сигналів), тривалістю символів τ і умовними ймовірностями $P(y_j/s_i)$ перетворення вхідних символів у вихідні. Зазвичай тривалості всіх вхідних та вихідних символів однакові. Об'єми алфавітів вхідних L_s та вихідних L_y сигналів у загальному випадку можуть бути різними, причому $L_y \geq L_s$. Однак у більшості випадків $L_y = L_s$. Для дискретних каналів широко використовується представлення прийнятої послідовності символів $Y = (y_1, y_2, \dots, y_n)$ у вигляді суми переданої послідовності $S = (s_1, s_2, \dots, s_n)$ та комбінації перешкоди (вектора помилки) $E = (e_1, e_2, \dots, e_n)$

$$Y = S \oplus E, \quad (6.19)$$

де \oplus – розуміється як порозрядне додавання S і E по модулю L_s . У разі двійкових послідовностей ($L_s = 2$) нульовий символ вектора помилки $e_i = 0$ означає, що i -й символ прийнятий правильно ($y_i = s_i$), а $e_i = 1$, помилку в прийомі ($y_i \neq s_i$).

Класифікацію дискретних каналів зручно вести по вектору помилки E . Різні моделі каналів відрізняються розподілом ймовірностей вектора E . Найбільш поширені наступні моделі:

Канал без пам'яті – це канал, у якому символи $e_i \in$ незалежними ВВ. Прийом кожного сигнального символу у такому каналі залежить від результату прийому попередніх символів. За наявності такої залежності має місце **канал із пам'яттю**. Дискретний канал називається **стаціонарним**, якщо ймовірність помилкового прийому символів не змінюється з часом.

Через простоту технічної реалізації найбільше застосування знаходять канали, сигнали у яких представляються двійковим кодом. Такі канали називаються двійковими (бінарними) та задаються за допомогою графа (рис. 6.1.). Імовірності $P(0/0)$ до $P(1/1)$ характеризують правильний прийом символів 0 та 1 відповідно, а $P(1/0)$ та $P(0/1)$ – ймовірності помилок при прийомі символів 0 та 1 .

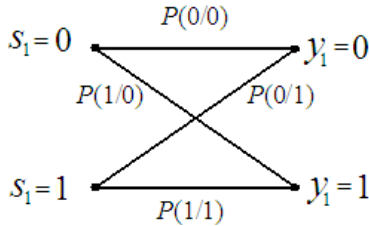


Рисунок 6.1 – Граф двійкового каналу

Симетричним двійковим називається канал, в якому ймовірності помилок при прийомі 0 і 1 однакові, $P(1/0) = P(0/1)$, а отже, рівні ймовірності правильного прийому символів $P(0/0) = P(1/1) = 1 - p$. Для симетричного стаціонарного каналу без пам'яті ймовірність спотворення i -го символу $P(e_i = 1) = p$, а ймовірність правильного прийому $P(e_i = 0) = 1 - p$.

Двійковий канал без пам'яті зі стиранням відрізняється від розглянутого тим, що вихідний алфавіт крім 0 і 1 містить третій символ «?» – символ стирання. Він з'являється у випадках, коли демодулятор неспроможний надійно розпізнати переданий символ. Такий канал часто використовується в системах передачі інформації із зворотним зв'язком, коли при прийомі символу "?" проводиться повторення передачі. Це дозволяє значно знизити можливість помилкового прийому за рахунок зменшення швидкості передачі.

Марківський канал є найпростішою моделлю дискретного каналу з пам'яттю. Він характеризується вектором помилки, символи якого утворюють простий ланцюг Маркова. Ймовірність спотворення символу в цьому каналі залежить від результату прийому попереднього символу.

Марківська модель задається матрицею перехідних ймовірностей:

$$P = \begin{bmatrix} 1 - p_1 & p_1 \\ p_2 & 1 - p_2 \end{bmatrix}, \quad (6.20)$$

де p_1 – умовна ймовірність прийняти $(i + 1)$ -й символ помилково, якщо i -й прийнятий правильно; $1 - p_1$ – умовна ймовірність прийняти $(i + 1)$ -й символ правильно, якщо i -й прийнятий правильно; p_2 – умовна ймовірність прийняти $(i + 1)$ -й символ помилково, якщо i -й прийнятий помилково; $1 - p_2$ – умовна ймовірність прийняти $(i + 1)$ -й символ правильно, якщо i -й прийнятий помилково.

Безумовна (середня) ймовірність помилки в каналі, що розглядається, повинна задовольняти рівняння:

$$P(x_{i+1} / x_i) = p_2 \cdot P_{\text{пом}}(x_i) + p_1 \cdot P_{\text{прав}}(x_i) \text{ або}$$

$$p(x_{i+1} / x_i) = p_1 / (1 + p_1 + p_2).$$

Ця модель має перевагу – простоту використання, але завжди адекватно відтворює властивості реальних каналів.

Високу точність дозволяє отримати **модель Гільберта для дискретного каналу з пам'яттю**. У такій моделі канал може перебувати у двох станах S_1 та S_2 . У стані S_1 помилок немає; у стані S_2 помилки виникають незалежно з ймовірністю p_2 . Також вважаються відомими ймовірності переходу $p(S_1 / S_2)$ зі стану S_1 в S_2 та ймовірності переходу $p(S_2 / S_1)$ зі стану S_2 у стан S_1 . У цьому випадку простий марківський ланцюг утворює не послідовність помилок, а послідовність переходів:

$$P = \begin{bmatrix} 1 - p(S_2 / S_1) & p(S_2 / S_1) \\ p(S_1 / S_2) & 1 - p(S_1 / S_2) \end{bmatrix}, \quad (6.21)$$

Безумовні ймовірності знаходження каналу станах S_1 і S_2 :

$$p(S_1) = \frac{p(S_1 / S_2)}{p(S_1 / S_2) + p(S_2 / S_1)}, \quad p(S_2) = \frac{p(S_2 / S_1)}{p(S_2 / S_1) + p(S_1 / S_2)}, \quad (6.22)$$

Безумовна ймовірність помилки в цьому випадку може бути визначена за такою формулою:

$$p = p_2 \cdot p(S_2) = p_2 \cdot \frac{p(S_2 / S_1)}{p(S_2 / S_1) + p(S_1 / S_2)}. \quad (6.23)$$

Найчастіше під час використання моделі Гільберта для двійкового каналу вважають $p_2 = 1/2$, тобто стан S_2 сприймається як повний обрив зв'язку. Це узгоджується з уявленням про канал, у якому діють комутаційні перешкоди.

З інших моделей симетричних двійкових каналів слід відзначити канал з пакетами помилок, який характеризується тим, що спотворюючі символи (одиниці) вектора помилки групуються в пакети. Таке групування відбувається, якщо в неперервному каналі, що входить до дискретного, діють сильні завмирання сигналів на час тривалості декількох символів або імпульсні перешкоди великої тривалості. Подібні канали задаються ймовірностями спотворень серій з q символів поспіль.

Змістовий модуль 2. Елементи теорії інформації в електрозв'язку

Тема 7. Інформаційні характеристики джерел повідомлень та каналів зв'язку

Джерела повідомлень та канали зв'язку. Основні поняття теорії інформації: кількісна міра інформації, ентропія джерела повідомлень, продуктивність, надмірність. Швидкість і пропускна здатність дискретного каналу зв'язку без завад. Інформаційні характеристики дискретних каналів зв'язку. Інформація в неперервних повідомленнях. Теорема Шеннона для дискретного та неперервного каналів зв'язку. Потенційні можливості неперервних каналів зв'язку при передачі дискретних повідомлень.

7.1 Основні поняття теорії інформації

Кількісна міра інформації використовується для оцінки кількості інформації, що міститься в повідомленні. Кількість інформації визначає, наскільки повідомлення зменшує невизначеність щодо стану об'єкта чи явища.

Кількість інформації $I(x_i)$ в окремому одиничному повідомленні x_i визначається величиною, зворотною до ймовірності появи повідомлення $p(x_i)$ і обчислюється в логарифмічних одиницях:

$$I(x_i) = \log_b \left(\frac{1}{p(x_i)} \right) = -\log_b p(x_i). \quad (7.1)$$

При $p(x_i)=1$, $I(x_i)=0$, що відповідає прийнятому визначенню інформації.

Якщо джерело видає залежні повідомлення $x_i=x_1, \dots, x_m$, всі вони характеризуються умовними ймовірностями $p(x_i/x_1, \dots, x_m)$. І в цьому випадку кількість інформації обчислюється за формулою (1) з підстановкою до неї умовних ймовірностей повідомлень.

У системах, що працюють із двійковими кодами, використовується основа логарифму $b = 2$ і інформація вимірюється в двійкових одиницях, зазвичай замість двійкових одиниць використовується еквівалентна назва – біт, яка визначається як кількість інформації, що зменшує невизначеність у двічі (2 можливих варіанти події), тобто 1 біт це кількість інформації, що передається одиничним символом повідомлення, ймовірність передачі якого $p(x_i) = 0,5$:

$$I(x_i) = -\log_2(0,5) = -\log_2\left(\frac{1}{2}\right) = -\log_2 2^{-1} = 1 \text{ (біт)}.$$

Ентропія джерела повідомлень $H(x)$, або середня кількість інформації $\bar{I}(X)$, біт/повідомлення – це міра невизначеності або «інформаційної

насиченості» джерела повідомлень. Вона визначає середню кількість інформації, що міститься в одному символі повідомлення, генерованого джерелом.:

$$H(x) = \bar{I}(x) = \sum_{i=1}^N P(x_i) \cdot I(x_i) \quad (7.2)$$

де N – кількість символів в ансамблі повідомлень.

Ентропія максимальна, коли всі символи однаково ймовірні. При цьому невизначеність є найбільшою, тому що неможливо передбачити, який символ з'явиться далі.

Продуктивність джерела – це середня кількість інформації, яку генерує джерело за одиницю часу. Продуктивність можна розрахувати за формулою:

$$R = H(X) \cdot f$$

де R – продуктивність джерела (біт/с), $H(X)$ – ентропія джерела (біт на символ), f – швидкість генерування символів (символ/с).

Збитковість (надлишковість) джерела інформації це міра надлишкової інформації в повідомленнях. Якщо в мові або коді є надмірність, це означає, що певні частини повідомлення можна опустити без втрати його сутності:

$$S = \frac{H_{\max}(x) - H(x)}{H_{\max}(x)} = 1 - \frac{H(x)}{H_{\max}(x)} = 1 - \frac{H(x)}{\log_2(x)} \quad (7.3)$$

де $H(X)$ – ентропія джерела, H_{\max} – максимальна ентропія, яку можна досягти для даного набору символів.

Кількість інформації, переданої по дискретному каналу

Кількість інформації, переданої каналом, це різниця між ентропією на виході та ентропією шуму:

$$I(Y, X) = H(Y) - H(Y/X) = \log_2 m + (1 - p_{\text{пом}}) \cdot \log_2(1 - p_{\text{пом}}) + p_{\text{пом}} \cdot \log_2 \frac{p_{\text{пом}}}{m-1}, \quad (7.4)$$

$I(Y, X)$ – кількість інформації, переданої каналом;

$H(Y/X)$ – ентропія шуму (хибна інформація, створювана перешкодами);

$H(Y)$ – ентропія на виході каналу;

$X = \{x_i\}$ – алфавіт джерела повідомлень;

$Y = \{y_j\}$ – алфавіт одержувача повідомлень;

$p_{ном}$ – ймовірність помилки під час передачі будь-якого символу;

m – кількість символів.

Кількість інформації, переданої по каналу зв'язку, має такі основні властивості:

– $I(Y,X) \geq 0$, причому $I(Y,X) = 0$ тоді і тільки тоді, коли вхідні та вихідні повідомлення в каналі взаємно незалежні;

– $I(Y,X) \leq H(X)$, причому $I(Y,X) = H(X)$, тоді і тільки тоді, коли вхідна послідовність визначається однозначно по вихідній послідовності, наприклад, коли в каналі немає перешкод;

– $I(Y,X) = I(X,Y) = H(Y) - H(Y|X)$ випливає з того, що кількість інформації не зміниться, якщо вхідну та вихідну послідовність поміняти місцями.

Пропускна здатність дискретного каналу

Пропускною здатністю каналу, розрахованою на один вхідний символ, називається максимальна кількість інформації, яка може бути передана каналом, причому максимум шукається по всіх можливих джерелах X , що мають різні (довільні) ймовірнісні характеристики:

$$C' = \max I(Y,X) \text{ [біт/символ]}. \quad (7.5)$$

де $I(X;Y)$ — взаємна інформація між входом X і виходом Y ,

Взаємна інформація $I(X;Y)$ визначає кількість інформації, яку символи на виході каналу дають про символи на вході. Вона обчислюється за формулою:

$$I(X;Y) = H(Y) - H(Y|X) \quad (7.6)$$

де $H(Y)$ — ентропія вихідних символів, $H(Y|X)$ — умовна ентропія вихідних символів за умови відомих вхідних символів.

Взаємна інформація показує, наскільки результат на виході залежить від переданих символів на вході, і вона є основою для визначення максимальної швидкості передачі даних без втрат.

Часто зручніше користуватися пропускною здатністю каналу, розрахованою не на один вхідний символ, а на одиницю часу:

$$C = C'/T \text{ [біт/с]}. \quad (7.7)$$

Величину C називають **пропускною здатністю каналу** в одиницю часу або просто пропускною здатністю.

Пропускна здатність каналу має такі основні властивості:

– $C \geq 0$, $C = 0$ тоді і лише тоді, коли вхід та вихід каналу статистично незалежні;

– $C \leq \log_2 m / T$ для каналу без перешкод.

З визначення пропускної здатності випливає, що ця характеристика визначає властивості каналу, яким передається інформація від певного джерела. Очевидно, ніяке джерело не здатне передати по каналу кількість інформації більшу за пропускну здатність і дана характеристика описує потенційні можливості каналу з передачі інформації.

Теорема Шеннона є основоположною в теорії інформації та комунікації. Вона визначає межу максимальної швидкості передачі інформації по каналу зв'язку з обмеженням на кількість помилок. Основне твердження теореми формулює поняття пропускної здатності або пропускної спроможності каналу (channel capacity), яка визначає, скільки інформації може бути передано по каналу з мінімально можливою кількістю помилок. Теорема Шеннона для дискретного каналу формулюється так:

Існує таке значення C (пропускна здатність каналу), що, якщо швидкість передачі інформації R не перевищує C , то ймовірність помилки можна зробити як завгодно малою при належному кодуванні. Якщо ж $R > C$, то ймовірність помилки завжди буде значною незалежно від кодування.

Пропускна здатність симетричного дискретного каналу без пам'яті

Пропускна здатність дискретного каналу, яким передається m дискретних сигналів з урахуванням (2) обчислюється за формулою:

$$C = V_1 \left[\log_2 m + (1 - p_{\text{пом}}) \cdot \log_2 (1 - p_{\text{пом}}) + p_{\text{пом}} \cdot \log_2 \frac{p_{\text{пом}}}{m-1} \right], \quad (7.8)$$

де $V_1 = 1/T$ – швидкість модуляції, бод; T – тривалість сигналу; $p_{\text{пом}}$ – ймовірність помилки в каналі.

Зауважимо, що пропускна здатність дискретного каналу без перешкод (при $p_{\text{пом}} = 0$): $C_{\text{ДК}} = V_1 [\log_2 m]$.

Зокрема пропускна здатність двійкового каналу ($m = 2$):

$$C_{\text{ДК}} = V_1 [1 + (1 - p_{\text{пом}}) \log_2 (1 - p_{\text{пом}}) + p_{\text{пом}} \log_2 (p_{\text{пом}})]. \quad (7.9)$$

Залежність відношення C/V_1 від ймовірності помилки $p_{\text{пом}}$, розрахована за формулою (4), показано на рис. 1.

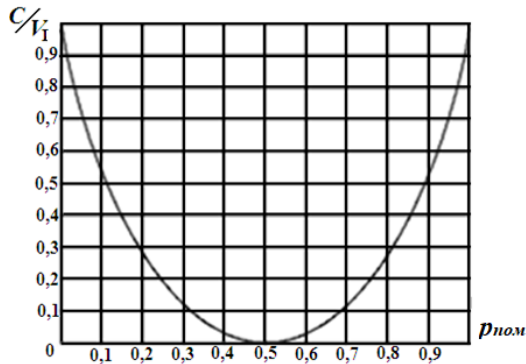


Рисунок 7.1 – Пропускна здатність дискретного каналу

Як випливає з графіка, при $p_{\text{пом}} = 0,5$ пропускна спроможність двійкового каналу дорівнює нулю ($C = 0$). Цей випадок називають обривом каналу. Дійсно, ймовірність помилки $p_{\text{пом}} = 0,5$ можна отримати і без передачі інформації по каналу зв'язку. А при $p_{\text{пом}} = 1$ пропускна здатність така сама, як і при $p_{\text{пом}} = 0$ (канал без перешкод). Це пояснюється тим, що при $p_{\text{пом}} = 1$ достатньо замінити нулі на одиниці та одиниці на нулі, щоб абсолютно правильно відновити переданий сигнал.

Приклад 1.

Визначимо пропускну здатність двійкового телеграфного каналу, якщо швидкість передачі в ньому 1000 біт/с і ймовірність помилки 10^{-3} і зробимо висновок про те, наскільки відрізняється пропускна здатність цього каналу від ідеального. Відповідно до формули (4), при заданих параметрах:

$$C_{\text{ДК}} = 1000 [1 + 0,001 \log_2 0,001 + (1 - 0,001) \log_2 (1 - 0,001)] = 989 \text{ [біт / с]}.$$

Для ідеального каналу при $p_{\text{пом}} = 0$ отримуємо $C_{\text{ДК}} = V_1 = 1000$ біт/с. Порівняння величин показує, що помилки в каналі призвели до зменшення пропускної спроможності на 11 біт/с (тобто втрати становили 1,1%).

Кількість інформації, переданої неперервним каналом

Замість послідовностей символів для дискретного каналу, в неперервному каналі здійснюється передача послідовності неперервних величин з дискретним або неперервним часом (у першому випадку ці послідовності можна представити у вигляді імпульсів різної величини, що з'являються в певні моменти часу, а в другому як неперервні функції часу).

Кількість інформації, що передається:

- $w(y)$ – густина розподілу ймовірності вихідних випадкових величин;
- $w(n)$ – щільність розподілу ймовірності перешкоди (адитивної);
- $h(Y)$ – диференціальна ентропія сигналу y ;

- $h(Y/X)$ – умовна диференціальна ентропія сигналу y за відомого сигналу x ;
Зазначимо такі властивості кількості інформації, що передається у неперервному каналі:
- $I(Y,X) \geq 0$, причому $I(Y,X) = 0$ тоді, і тоді, коли вхід і вихід каналу статистично незалежні, тобто. $w(y/x) = w(y)$;
- $I(Y,X) = I(X,Y)$ – властивість симетрії;
- $I(Y,X) = \infty$, якщо перешкоди у каналі відсутні, тобто $y = x$, $n = 0$

Диференціальна ентропія $h(Y)$ вже не є середньою кількістю інформації, що видається джерелом сигналу (для неперервного сигналу вона нескінченна). Аналогічно $h(Y/X)$ не є кількістю інформації, втраченої в каналі, оскільки ця величина теж нескінченна. Тому диференціальну ентропію слід розуміти лише формально як деяку допоміжну величину корисну при розрахунках.

Якщо перешкода адитивна $y = x + n$, то неважко показати, що

$$h\left(\frac{Y}{X}\right) = \int_{-\infty}^{+\infty} w(n) \cdot \log \frac{1}{w(n)} dn = h(N), \quad (7.10)$$

де $w(n)$ – густина розподілу ймовірності перешкоди, а $h(N)$ – диференціальна ентропія перешкоди.

Вираз визначення кількості інформації, переданої по неперервному гаусівському каналу (нормальний закон розподілу ймовірностей сигналу і перешкоди):

$$I(Y, X) = h(X) - h\left(\frac{X}{Y}\right) = \frac{1}{2} \log_2 \left(\frac{\sigma_c^2 + \sigma^2}{\sigma^2} \right). \quad (7.11)$$

де

$$\begin{aligned} h(N) &= \log_2 \left(\sqrt{2\pi\sigma^2} \right) + \frac{1}{2} \log_2 e = \log_2 \left(\sqrt{2\pi e \sigma^2} \right) \\ h(Y) &= - \int_{-\infty}^{+\infty} w(y) \cdot \log w(y) dy = \log_2 \left(\sqrt{2\pi e \sigma_y^2} \right) \end{aligned} \quad (7.12)$$

Отриманий вираз показує, що кількість інформації гаусівського каналу з дискретним часом визначається відношенням дисперсії сигналу σ_c^2 до дисперсії перешкоди σ^2 . Нерідко величину $\sigma_c^2/\sigma^2 = h_2$ називають відношенням сигнал/шум. Чим більше це відношення, тим вища пропускна здатність.

Пропускна здатність неперервного каналу

Нехай сигнал $y(t)$ на виході каналу є сумою корисного сигналу $x(t)$ і шуму $n(t)$, тобто. $y(t) = x(t) + n(t)$, причому $x(t)$ та $n(t)$ статистично незалежні, і канал має обмежену смугу пропускання шириною $\Delta F_{\text{НК}}$.

Пропускна здатність гаусівського каналу з дискретним часом, розрахована на одиницю часу, з урахуванням (7.10), може бути записана у вигляді:

$$C = V_1 \cdot I(Y, X) = \frac{V_1}{2} \log_2 \left(\frac{\sigma_c^2 + \sigma^2}{\sigma^2} \right) = \frac{V_1}{2} \log_2(1 + h^2) \quad (7.13)$$

Отриманий вираз показує, що пропускна здатність гаусівського каналу з дискретним часом визначається числом імпульсів, що передаються за секунду, і відношенням сигнал/шум (h).

З урахуванням взаємозв'язку швидкості передачі інформації та смуги частот неперервного каналу від (7.13) можна перейти до формули Шеннона, яка встановлює зв'язок пропускної здатності гаусівського каналу зі смугою пропускання неперервного каналу та відношенням потужності сигналу до перешкоди:

$$C = \Delta F_{\text{НК}} \log_2(1 + h^2) \quad (7.14)$$

Зауважимо, що з малому відношенні $h^2 \ll 1$, $C \approx \Delta F_{\text{НК}} \cdot 1,442 \cdot h^2$, за великого відношенні $h^2 \gg 1$, можна знехтувати одиницею і вважати, що $C = \Delta F_{\text{НК}} \cdot \log_2(h^2)$, тобто залежність пропускної здатності неперервного каналу від відношення сигнал/шум логарифмічна.

Графік відносини $C / \Delta F_{\text{НК}} = \log_2(1 + h^2)$ представлений рис. 7.2.

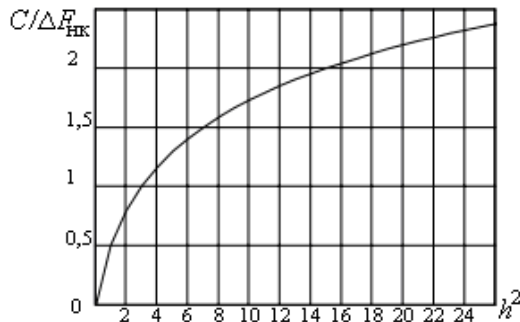


Рисунок 7.2 – Графік відносини $C / \Delta F_{\text{НК}}$

Пропускна здатність каналу як граничне значення швидкості безпомилкової передачі інформації є однією з основних характеристик будь-якого каналу.

Приклад 3. Визначимо пропускну здатність каналу тональної частоти, що має межі частот, що ефективно передаються 0,3...3,4 кГц, середню потужність сигналу на виході 56 мВт при середній потужності перешкоди 69000 пВт.

$$C_{\text{нк}} = 3,1 \cdot 10^3 \cdot \log_2 \left(1 + \frac{56 \cdot 10^{-6}}{69 \cdot 10^{-12}} \right) = 3,0 \cdot 10^4 \left[\frac{\text{біт}}{\text{с}} \right].$$

Таблиця 7.1

Характеристики типових каналів багатоканального зв'язку

Границі переданих частот, Гц	Пропускна спроможність, біт/с
300 ... 3400	$20,64 \cdot 10^3$
$12,3 \cdot 10^3 \dots 23,4 \cdot 10^3$	$73,91 \cdot 10^3$
$60,6 \cdot 10^3 \dots 107,7 \cdot 10^3$	$313,6 \cdot 10^3$
$312,3 \cdot 10^3 \dots 551,4 \cdot 10^3$	$1,59 \cdot 10^6$
$812,3 \cdot 10^3 \dots 2043,7 \cdot 10^3$	$8,2 \cdot 10^6$

Знаючи пропускну здатність каналу та інформаційні характеристики повідомлень (таблиця 7.1) можна визначити, які повідомлення (первинні сигнали) можна передавати по заданому каналу.

Наприклад, первинний сигнал телевізійного мовлення має $C = 208 \cdot 10^6$ біт/с і тому не може бути переданий по одному з типових неперервних або цифрових каналів без втрати якості. Отже, для передачі сигналу телевізійного мовлення потрібно створення спеціальних каналів з вищою пропускну здатністю чи зниження швидкості цифрового потоку.

Теорема Шеннона для неперервного каналу встановлює верхню межу швидкості передачі даних. Збільшення потужності сигналу до певної межі дозволяє підвищити пропускну здатність каналу, але вимагає збільшення енергоспоживання, що є важливим фактором у мобільних і супутникових системах зв'язку. На практиці цього максимуму не можна досягти повністю, але сучасні системи зв'язку використовують різні методи кодування, щоб максимально наблизитися до цієї межі.

Тема 8. Теорія кодування повідомлень

Узгодження джерела з каналом. Показники ефективності стиснення. Код Шеннона-Фано, Хаффмена, Завадостійке кодування. Коди з виявленням помилок. Коригувальні коди. Код Хеммінга. Циклічні коди. Коди Ріда-Соломона.

8.1 Завдання кодування повідомлень

При передачі повідомлення необхідно узгоджувати джерело з каналом шляхом визначення правила, яким кожному елементу повідомлення ставиться у відповідність деякий код, перетворений далі в сигнал. Розв'язання задачі вибору (відшукування) коду, оптимального за тим чи іншим критерієм, становить суть теорії кодування.

Кодуванням називається процес перетворення повідомлень у комбінації дискретних сигналів. Основними **завданнями**, які вирішуються кодуванням у процесі передачі повідомлень, є:

- узгодження джерела повідомлень із каналом за обсягами алфавітів;
- підвищення швидкості передачі інформації по каналу за рахунок усунення надмірності в послідовності повідомлень. Мета кодування при цьому полягає в такому перетворенні повідомлення, при якому надмірність кодової послідовності повинна стати меншою, ніж надмірність повідомлень джерела.;
- підвищення достовірності передачі в каналах з перешкодами, які вирішуються застосуванням стійких до перешкод кодів, які дозволяють виявляти або виправляти помилки. При цьому надмірність кодової послідовності вища, ніж надмірність джерела повідомлень.

Перші дві задачі вирішуються в кодері джерела повідомлень. Третє завдання вирішується у кодері каналу.

Процес кодування з метою зменшення надмірності джерела повідомлень зветься **узгодженням джерела з каналом** або стиснення джерела (економного кодування, ентропійного кодування).

Кількісно надмірність оцінюється коефіцієнтом надмірності:

$$\chi = \frac{H_{\max}(X) - H(X)}{H_{\max}(X)} = 1 - \frac{H(X)}{H_{\max}(X)}, \quad (8.1)$$

де $H(X)$ – ентропія джерела; $H_{\max}(X) = \log_2 t$ – максимальна ентропія джерела з алфавітом із t повідомлень.

Надмірність (9.1) дорівнює нулю тільки в тому випадку, коли елементи повідомлення з'являються на виході джерела з рівними ймовірностями $p(x_i) = 1/t$ ($i = 1, 2, 3, \dots, t$) незалежно один від одного $p(x_i, x_j) = p(x_i) \cdot p(x_j)$. Якщо

ж $H(X) < \log_2 m$, то виявляється можливим побудова кодів, що мають меншу надмірність, ніж джерело повідомлень.

Покажемо це на найпростішому прикладі.

Нехай джерело має алфавіт із чотирьох символів А, Б, В, Г із ймовірностями $p(A) = 0,5$; $p(B) = 0,25$; $p(B) = p(\Gamma) = 0,125$.

Ентропія такого джерела:

$$H(X) = -\sum_{i=1}^4 p(x_i) \cdot \log_2 p(x_i) =$$

$$= -0,5 \cdot \log 0,5 - 0,25 \cdot \log 0,25 - 0,125 \cdot \log 0,125 - 0,125 \cdot \log 0,125 = 1,75$$

При використанні рівномірного кодування, наприклад $A \rightarrow 00$, $B \rightarrow 01$, $V \rightarrow 10$, $\Gamma \rightarrow 11$. Тоді середня кількість двійкових символів у повідомленні, що припадають на один символ джерела, дорівнює двом. Оскільки це на 12,5% більше ентропії джерела, то код, що використовується, не є оптимальним.

Розглянемо тепер нерівномірний код: $A \rightarrow 0$, $B \rightarrow 10$, $V \rightarrow 110$, $\Gamma \rightarrow 111$. У цьому випадку середня кількість двійкових символів, що є одним символом джерела в повідомленні:

$$n_{cp} = \sum_{i=1}^4 p(x_i) \cdot n_i = 0,5 \cdot 1 + 0,25 \cdot 2 + 0,125 \cdot 3 + 0,125 \cdot 3 = 1,75.$$

Отже, середнє число двійкових знаків, що припадають на один символ джерела, дорівнює ентропії джерела, тобто. для зазначеного джерела нерівномірний код виявляється економічнішим, ніж рівномірний.

Важливо, що при кодуванні нерівномірним кодом повинна забезпечуватися можливість однозначного декодування символів повідомлення.

Наприклад, для розглянутого джерела, недоцільно застосовувати код: $A \rightarrow 0$, $B \rightarrow 1$, $V \rightarrow 10$, $\Gamma \rightarrow 11$, оскільки прийом послідовності 10 може означати передачу одного символу В, або двох символів Б і А. Неоднозначне також декодування символів 11. Для однозначного декодування нерівномірні коди повинні задовольняти умові префіксності: жодне коротке слово не повинне бути початком більш довгого слова. Нерівномірні коди, що задовольняють цю умову, називають префіксними.

Зауважимо, що сучасні методи кодування не дозволяють близько підійти до потенційної пропускнуєї спроможності каналу зв'язку за одночасної високої точності передачі. Однак грамотний вибір коду дозволяє, у багатьох

випадках, значно знизити ймовірність помилкового прийому при швидкості передачі $10 \div 50$ % пропускну́ї спроможності каналу.

Нерівномірні коди дозволяють в середньому зменшити кількість двійкових символів на інформаційне повідомлення. Однак їм притаманний істотний недолік: у разі виникнення помилки вона поширюється на всі наступні елементи повідомлення. Виникає помилка синхронізації, що веде до різкого погіршення достовірності прийому.

Цей недолік відсутній у рівномірних кодах. При кодуванні рівномірними кодами використовується одне й те число двійкових символів – блок; тому такі коди називають **блоковими**.

8.2 Показники ефективності стиснення

Поряд з коефіцієнтом надмірності (8.1) часто використовується коефіцієнт стиснення джерела:

$$K_{\text{сж.м}} = \frac{H_{\text{max}}(X)}{H(X)} = \frac{1}{1 - \chi}. \quad (8.2)$$

Коефіцієнт стиснення джерела показує, у скільки разів можна зменшити кількість двійкових символів для представлення одиничного символу джерела з ентропією $H(X)$ порівняно з випадком, коли всі повідомлення джерела передаються рівноймовірно.

Наприклад, для джерела, розглянутого у п. 8.1, коефіцієнт стиснення:

$$K_{\text{ст.м}} = \frac{\log_2 m}{H(X)} = \frac{\log_2 4}{1,75} = \frac{2}{1,75} = 1,14,$$

тобто швидкість передачі по каналу зв'язку при використанні економічного кодування може бути в 1,14 рази більша, ніж при рівномірному кодуванні.

Кодування джерела дискретних повідомлень методом Шеннона-Фано.

Кодування шляхом Шеннона-Фано розглянемо на прикладі. Нехай алфавіт джерела містить шість елементів $\{A, B, B, P, D, E\}$, що з ймовірностями $p(A) = 0,15$; $p(B) = 0,25$; $p(B) = 0,1$; $p(\Gamma) = 0,13$; $p(D) = 0,25$; $p(E) = 0,12$.

$$H(X) = -0,15 \cdot \log_2 0,15 - 0,25 \cdot \log_2 0,25 - 0,1 \cdot \log_2 0,1 - 0,13 \cdot \log_2 0,13 - 0,25 \cdot \log_2 0,25 - 0,12 \cdot \log_2 0,12 = 2,492.$$

Алгоритм побудови стискаючого коду Шеннона-Фано, полягає в наступному:

1. Усі m символів дискретного джерела розташовуються у порядку зменшення ймовірностей їх появи (таблиця 8.1).

2. Утворений стовпець символів поділяється на дві групи таким чином, щоб сумарні ймовірності кожної групи мало відрізнялися один від одного.
3. Верхня група кодується символом «1», а нижня – «0».
4. Кожна група поділяється на дві підгрупи з близькими сумарними можливостями; верхня підгрупа кодується символом «1», а нижня – «0».
5. Процес поділу та кодування триває доти, доки у кожній підгрупі не виявиться по одному символу повідомлення джерела.
6. Записується код кожного символу джерела; зчитування коду здійснюється зліва направо.

Таблиця 8.1

Побудова коду Шеннона-Фано

Елемент повідомлення	Імовірність елемента	Поділ повідомлення на групи і підгрупи	Код
Б	0,25		11
Д	0,25		10
А	0,15		011
Г	0,13		010
Е	0,12		001
В	0,1		000

При використанні найпростішого рівномірного коду для кодування шести елементів джерела алфавіту потрібно по три двійкові символи на кожну літеру повідомлення. Якщо ж використовується код Шеннона-Фано, то середня кількість символів на одну літеру:

$$n_{cp} = \sum_{i=1}^6 p(x_i) \cdot n_i = 0,25 \cdot 2 + 0,25 \cdot 2 + 0,15 \cdot 3 + 0,13 \cdot 3 + 0,12 \cdot 3 + 0,1 \cdot 3 = 2,5,$$

що менше, ніж при найпростішому рівномірному коді і трохи відрізняється від ентропії джерела.

Кодування джерела дискретних повідомлень методом Хаффмена.

Розглянемо ще один підхід до кодування, запропонований Хаффменом, з прикладу джерела повідомлень, заданого в таблиці 8.2.

Довжина кодової комбінації таких кодів залежить від ймовірності вибору відповідної літери алфавіту: найімовірнішим літерам відповідають короткі кодові комбінації, а менш ймовірним – довші.

8.3 Завадостійке кодування

В даний час підвищення достовірності передачі в каналах з перешкодами здійснюється за допомогою кодів, що дозволяють виявляти або виправляти помилки. Таке кодування називається **завадостійким**. При цьому надмірність кодової послідовності вища, ніж надмірність джерела повідомлень. Завдяки цьому і виявляється можливим виявлення та виправлення помилок прийому.

Побудова завадостійких кодів в основному пов'язана з додаванням до вихідної комбінації (m – символів) контрольних (k – символів). Закодована комбінація становитиме n – символів. Ці коди часто називають (n, m) – коди.

8.3.1 Коди з виявленням помилок

Код із перевіркою на парність.

Таблиця 8.3

Код із перевіркою на парність

m	k	n
11011	0	110110
11100	1	111001

Такий код утворюється шляхом додавання до комбінації, що передається і складається з m інформаційних символів, одного контрольного символу k (0 або 1), так, щоб загальна кількість одиниць в комбінації що передається була парною (таблиця 8.3). Якщо отримане слово з непарною кількістю одиниць, при передачі відбулася помилка.

Наприклад, для коду з $m = 5$ та ймовірністю помилки $p = 10^{-2}$ коефіцієнт виявлення складе $K_e = 0,9$. Тобто 90% помилок виявляємо, при цьому надмірність складатиме $L = 0,17$ (17%).

Код з постійними вагами.

Таблиця 8.4

Код із постійними вагами

$N = C_5^2 = 10$	$N = C_7^3 = 35$
11000	0000111
10010	1001001
00101	1010100

Цей код містить постійне число одиниць та нулів. Для прикладу наведено коди з двома одиницями із п'яти та трьома одиницями із семи.

Число кодових комбінацій складе:

$$N = C_n^k = \frac{n!}{k!(n-k)} \quad (8.4)$$

Цей код дозволяє виявляти будь-які поодинокі помилки та частину багаторазових помилок. Не виявляються цим кодом лише помилки усунення, коли одночасно одна одиниця перетворюється на нуль і один нуль перетворюється на одиницю, два нуля і дві одиниці змінюються на зворотні символи тощо.

Для коду C_7^3 при $p = 10^{-2}$ коефіцієнт виявлення складе $K_e = 0,985$ надмірність становить $L = 27\%$.

Кореляційний код (Код з подвоєнням).

Елементи цього коду замінюються двома символами, одиниця «1» перетворюється на 10, а нуль «0» на 01.

Замість комбінації 1010011 передається 10011001011010. Помилка виявляється в тому випадку, якщо парні елементи мають однакові символи 00 або 11 (замість 01 і 10).

Наприклад, при $m = 5$, $n = 10$ та ймовірності помилки $p = 10^{-2}$, $K_e = 0,995$. Але при цьому надмірність складатиме 50%.

Інверсний код.

Таблиця 8.5

Інверсний код		
m	k	n
11011	11011	1101111011
11100	00011	1110000011

До вихідної комбінації додається така сама комбінація по довжині, тобто кількість символів подвоюється. Якщо у вихідній комбінації парне число одиниць, то комбінація, що додається, повторює вихідну комбінацію, якщо непарне, то комбінація, що додається, є інверсною по відношенню до вихідної.

Приєм інверсного коду здійснюється у два етапи. На першому етапі підсумовуються одиниці у першій основній групі символів. Якщо число одиниць парне, контрольні символи приймаються без зміни, якщо непарне, то контрольні символи інвертуються. На другому етапі контрольні символи підсумовуються з інформаційними символами по модулю два. Нульова сума

свідчить про відсутність помилок. За ненульової суми, прийнята комбінація бракується. Покажемо підсумовування для прийнятих комбінацій без помилок (1,3) та з помилками (2,4):

1	11011	2	11111	3	11100	4	11000
	11011		00100		11100		11100
	-----		-----		-----		-----
	00000		11011		00000		00100

Цей код виявляє практично будь-які помилки, крім рідкісних помилок зміщення, які одночасно відбуваються як серед інформаційних символів, так і серед відповідних контрольних.

При $m = 5$, $n = 10$ та $p = 10^2$, $K_e = 1 \cdot 10^{-5}$.

8.3.2 Кориговальні коди

Кориговучими називаються коди, що дозволяють виявляти і виправляти помилки. Вони характеризуються кодовою (або хемінговою) відстанню, що позначається буквою d .

Кодова відстань – це мінімальна кількість елементів, у яких одна кодова комбінація відрізняється від іншої. Для визначення кодової відстані достатньо порівняти дві кодові комбінації за модулем 2.

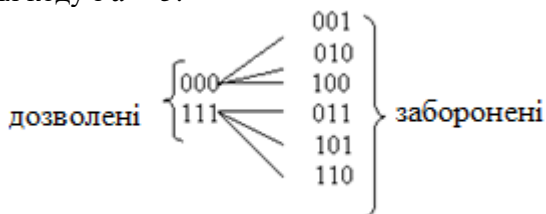
Так, склавши дві комбінації, визначимо, що відстань між ними $d = 7$.

$$\begin{array}{r} 10110101101 \\ \underline{11001010101} \\ 01111111000 \end{array}$$

Код з $n = 3$ і $d = 1$ для передачі використовуються всі вісім кодових комбінацій 000, 001, ..., 111. Такий код є не завадостійким, він не в змозі виявити помилку.

Якщо виберемо комбінації з кодовою відстанню $d = 2$, наприклад, 000, 110, 101, 011, такий код дозволить виявляти одноразові помилки. Назвемо ці комбінації дозволеними, призначеними для передачі інформації. Решта 001, 010, 100, 111 – заборонені. Будь-яка одиночна помилка призводить до того, що дозволена комбінація переходить в найближчу, заборонену комбінацію. Отримавши заборонену комбінацію, ми виявимо помилку.

Комбінації для коду з $d = 3$:



Такий код може виправити одну одиночну помилку або виявити дві помилки. Таким чином, збільшуючи кодову відстань можна збільшити перешкодостійкість коду. У загальному випадку кодова відстань визначається

за формулою $d = (S + r) + 1$, де S - число помилок, що виправляються, r - число виявлених помилок. Зазвичай $r > S$.

Більшість коригувальних кодів є лінійними кодами. Лінійні коди – це коди, у яких контрольні символи утворюються шляхом лінійної комбінації інформаційних символів.

Крім того, коригувальні коди є груповими кодами. Групові коди (G_n) – це коди, які мають одну основну операцію. При цьому має дотримуватися умова замкнутості (тобто, при складанні двох елементів групи виходить елемент, що належить цій групі). Число розрядів у групі не повинно збільшуватись. Цій умові задовольняє операція порозрядного додавання по модулю 2. У групі, крім того, має бути нульовий елемент.

Приклади кодових комбінацій:

- 1) 1101 1110 0111 1011 - не група, тому що немає нульового елемента;
- 2) 0000 1101 1110 0111 - не група, тому що не дотримується умова замкнутості ($1101 + 1110 = 0011$);
- 3) 000 001 010 011 100 101 110 111 – група;
- 4) 000001010111 - підгрупа.

Більшість коригуючих кодів утворюються шляхом додавання до вихідної m – комбінації k – контрольних символів. У результаті лінію передаються $n = m + k$ символів. При цьому коригувальні коди називаються (n, m) кодами.

Для побудови коду здатного виявляти та виправляти одиночну помилку необхідне число контрольних розрядів становитиме $n - m \geq \log(n + 1)$.

Якщо необхідно виправити дві помилки, кількість різних результатів буде становити C_n^2 . Тоді $n - m \geq \log(1 + C_n^1 + C_n^2)$, у цьому випадку виявляються однократні та двократні помилки. У загальному випадку число контрольних символів має бути не менше:

$$n - m \geq \log(1 + C_n^2 + C_n^2 + \dots + C_n^t) = \log \sum_{i=0}^t C_n^i \quad (8.5)$$

Ця формула називається **нерівністю Хеммінга**, або нижньою межею Хеммінга для числа контрольних символів.

Код Хеммінга.

Код Хеммінга, що є груповим (n, m) кодом, з мінімальною відстанню $d=3$ дозволяє виявляти та виправляти одноразові помилки. Побудова кодів Хеммінга базується на принципі перевірки на парність ваги W (числа одиничних символів) в інформаційній групі кодового блоку.

До кожного числа перевірочних символів $k = 3, 4, 5 \dots$ існує класичний код Хеммінга з маркуванням $(n, m) = (2k - 1, 2k - 1 - k)$, тобто $(7, 4)$, $(15, 11)$, $(31, 26)$...

Наприклад, розглянемо класичний код Хеммінга $(7, 4)$. У найпростішому варіанті для заданих чотирьох ($m = 4$) інформаційних символів (i_1, i_2, i_3, i_4) вважатимемо, що вони згруповані на початку кодового слова, хоча це не обов'язково. Доповнимо ці інформаційні символи трьома перевірочними символами ($k = 3$), задаючи їх такими рівностями перевірки на парність, що визначаються відповідними алгоритмами:

$$k_1 = i_1 \oplus i_2 \oplus i_3;$$

$$k_2 = i_2 \oplus i_3 \oplus i_4;$$

$$k_3 = i_1 \oplus i_2 \oplus i_4,$$

де знак \oplus означає додавання за модулем 2.

Відповідно до цього алгоритму визначення значень перевірочних символів k_i можливі 16 кодових слів $(7, 4)$ – коду Хеммінга.

$m = 4$				$k = 3$		
i_1	i_2	i_3	i_4	k_1	k_2	k_3
0	0	0	0	0	0	0
0	0	0	1	0	1	1
0	0	1	0	1	1	0
0	0	1	1	1	0	1
0	1	0	0	1	1	1
0	1	0	1	1	0	0
0	1	1	0	0	0	1
0	1	1	1	0	1	0
1	0	0	0	1	0	1
1	0	0	1	1	1	0
1	0	1	0	0	1	1
1	0	1	1	0	0	0
1	1	0	0	0	1	0
1	1	0	1	0	0	1
1	1	1	0	1	0	0
1	1	1	1	1	1	1

На вхід декодера надходить кодове слово $V = (i_1', i_2', i_3', i_4', k_1', k_2', k_3')$, апостроф означає, що будь-який символ слова може бути спотворений завадою каналу передачі.

У декодері у режимі виправлення помилок будується послідовність:

$$s_1 = k_1' \oplus i_1' \oplus i_2' \oplus i_3';$$

$$s_2 = k_2' \oplus i_2' \oplus i_3' \oplus i_4';$$

$$s_3 = k_3' \oplus i_1' \oplus i_2' \oplus i_4'.$$

Трисимвольна послідовність (s_1, s_2, s_3) називається синдромом S .

Синдром $S = (s_1, s_2, s_3)$ є поєднанням результатів перевірки на парність відповідних символів кодової групи і характеризує певну конфігурацію помилок (шумовий вектор).

Число можливих синдромів визначається виразом: $S = 2^k$.

При числі перевірочних символів $k = 3$ є вісім можливих синдромів ($2^3 = 8$). Нульовий синдром (000) вказує на те, що помилки при прийомі відсутні або не виявлені. Будь-якому ненульовому синдрому відповідає певна зміна помилок, яка і виправляється. Класичні коди Хеммінга мають число синдромів, що точно дорівнює їх необхідному числу, дозволяють виправити всі одноразові помилки в будь-якому інформативному і перевірочному символах і включають один нульовий синдром. Такі коди називаються щільноупакованими.

Усічені коди є нещільноупакованими, оскільки кількість синдромів у них перевищує необхідну. Так, у коді $(9,5)$ при чотирьох перевірочних символах число синдромів дорівнюватиме $2^4 = 16$, тоді як необхідно всього 10. Зайві 6 синдромів свідчать про неповну упаковку коду $(9,5)$. Для коду $(7,4)$, що розглядається в таблиці 8.6 представлені ненульові синдроми і відповідні конфігурації помилок.

Таблиця 8.6

Усічені коди

Синдром	001	010	011	100	101	110	111
Конфігурація помилок	00000 01	00000 10	00010 00	00001 00	10000 00	00100 00	01000 00
Помилка в символі	k_3	k_2	i_4	k_1	i_1	i_3	i_2

Таким чином, код $(7,4)$ дозволяє виправити усі поодинокі помилки. Проста перевірка показує, що кожна помилка має свій єдиний синдром. При цьому можливе створення такого цифрового коректора помилок (дешифратора синдрому), який за відповідним синдромом виправляє відповідний символ прийнятої кодової групи.

8.3.3 Циклічні коди

Циклічним кодом називається такий груповий код, який пов'язаний з додатковою умовою циклічності. Усі рядки утворюючої матриці такого коду можуть бути отримані циклічним зсувом однієї комбінації, яка називається утворюючою для даного коду.

Циклічні коди широко застосовуються при передачі даних у сучасних інформаційних системах завдяки ряду позитивних якостей, основними з яких є:

- висока ефективність, так як циклічні коди мають порівняно невелику надмірність, відрізняються простотою реалізації кодуючих і декодуючих пристроїв;
- висока завадостійкість - за рахунок здатності коду до виявлення та виправлення помилок.

Для оптимального коду кількість помилок, що виправляються, дорівнює $2^k - 1$, де k - число контрольних розрядів.

Цифрові коди, що застосовуються в даний час, містять n розрядів, з яких m є інформаційними, а $k = n - m$ - контрольними (перевірочними), розташованими в кінці кодової комбінації. Так як інформаційні та контрольні розряди займають строго певні місця та довжина кодової комбінації постійна, то циклічні коди відносяться до систематичних кодів.

Для опису циклічних кодів зазвичай користуються записом будь-якого n -розрядного двійкового числа як многочлена ступеня $(n-1)$. Наприклад, кодова комбінація 1011001 записується як многочлен $x^6 + x^4 + x^3 + 1$, тобто коефіцієнти многочлена не пишуться, а члени з коефіцієнтами 0 опускаються. Найвища ступінь числа з коефіцієнтом 1 називається ступенем полінома (многочлена). Так, у прикладі розглянуто многочлен 6 ступеня.

Таким чином, дії над кодовими числами можна звести до дій над многочленами. При цьому використовують теорію комутативних кілець. Комутативним кільцем називають множину, у якій особливим чином визначено операції додавання та множення.

У циклічному кодуванні всі математичні операції додавання проводяться з використанням додавання по $mod 2$. (додавання за модулем 2 полягає у додаванні двох двійкових чисел без урахування перенесення одиниці зі старшого розряду в наступний розряд)

Операцію множення символічно проводять за такими правилами:

1) Спочатку всі многочлени перемножуються за звичайними правилами, але з додаванням подібних членів за $mod 2$.

2) Якщо старший ступінь отриманого в результаті множення многочлена не перевищує $(n-1)$, цей многочлен є результатом символічного множення.

3) Якщо старший ступінь отриманого внаслідок множення многочлена більший $(n-1)$, то многочлен добутку ділиться на двочлен $x^n + 1$. І тут результатом символічного множення вважається залишок від ділення.

Приклад:

Маємо кодові комбінації 001101 та 101110 , де $n = 6$.

Ці комбінації відповідають многочленам: $x^3 + x^2 + 1$ і $x^5 + x^3 + x^2 + x$.

Допустимо, необхідно провести двічі циклічний зсув цих кодових комбінацій. В результаті отримаємо:

$$\begin{array}{l}
 001101 \Rightarrow 110100; \\
 \begin{array}{c} \longleftarrow \uparrow \\ \longleftarrow \uparrow \end{array} \\
 101110 \Rightarrow 111010. \\
 \begin{array}{c} \longleftarrow \uparrow \\ \longleftarrow \uparrow \end{array}
 \end{array}$$

Для проведення операції з многочленами необхідно їх символічно помножити на $x \times x = x^2$:

1) $(x^3 + x^2 + 1) \times x^2 = x^5 + x^4 + x^2$, ступінь отриманого многочлена не перевищує $(n-1) = 5$, цей многочлен приймається за результат множення і дійсно відповідає зсунутій кодовій комбінації 110100.

2) $(x^5 + x^3 + x^2 + x) \times x^2 = x^7 + x^5 + x^4 + x^3$, ступінь отриманого многочлена перевищує $(n-1)$, то для отримання результату символічного множення необхідно зробити ділення цього многочлена на двочлен $(x^n + 1)$:

$$\begin{array}{r}
 x^7 + x^5 + x^4 + x^3 \quad | \quad x^6 + 1 \\
 \hline
 x^7 + x \\
 \hline
 \text{-----} \\
 x^5 + x^4 + x^3 + x
 \end{array}$$

Остача від ділення $x^5 + x^4 + x^3 + x$ приймається як результат символічного множення, що відповідає циклічно зрушеній кодовій комбінації 111010.

В основі утворення циклічного коду лежить використання так званого утворюючого многочлена (полінома). Вибір утворюючого полінома визначає тип циклічного коду і характеризує його виявляючі і виправляючі здібності.

Степінь утворюючого полінома дорівнює k , тобто рівна числу контрольних символів. Будь-який многочлен циклічного коду повинен ділитися без остачі на утворюючий поліном.

У той же час жоден многочлен, що відповідає забороненій кодовій комбінації, не повинен ділитися без остачі на утворюючий поліном. Ця властивість дозволяє виявити помилку, а по вигляду остачі і вектор помилки, тобто виправляти помилки.

Утворення циклічного коду.

Для отримання циклічного коду многочлен $G(x)$ (що відповідає кодовим комбінаціям безнадлишкового m – розрядного коду), множать на x^k . Це відповідає приписуванню з боку молодших розрядів k нулів до кодових комбінацій.

Потім добуток $G(x) \times x^k$ ділиться на утворюючий многочлен $P(x)$. У загальному випадку ми отримуємо в результаті такого ділення $Q(x)$ такої ж

степені, як і $G(x)$ і остачу $R(x)$. остача $R(x)$ додається до $G(x) \times x^k$. Отримуємо многочлен $F(x) = G(x) \times x^k + R(x)$.

Оскільки в комбінаціях, що відповідають многочлену $G(x) \times x^k$, перші k молодших розрядів – нулі, а $R(x)$ – многочлен ступеня не вище $k-1$, то операція отримання многочлена відповідає приписуванню $R(x)$ до $G(x)$ із боку молодших розрядів.

Отриманий таким чином многочлен $F(x)$ буде ділитися на утворюючий многочлен $P(x)$ без залишку.

Таким чином, циклічний код можна отримати, якщо до кожної кодової комбінації безнадлишкового коду приписувати остачу від поділу многочлена, що відповідає цій кодовій комбінації на утворюючий многочлен.

Оскільки розпізнавальниками помилок є остачі від ділення многочленів циклічного коду на утворюючий многочлен, то коригуюча здатність коду буде тим вища, чим більше остач від ділення можна утворити. Найбільша кількість остач, рівна $2^k - 1$ (виключаючи нульову), може забезпечити простий многочлен, тобто такий многочлен, який ділиться лише сам на себе. Тому як утворюючий многочлен необхідно вибирати простий многочлен (або їх добуток).

Приклад утворення циклічного коду.

Нехай інформаційний код містить $m = 4$ розряди. Одна з $N = 2^m$ комбінацій цього коду: 1101 як многочлен запишеться так: $G(x) = x^3 + x^2 + 1$.

Якщо циклічний код виявляє і виправляє одну помилку, його мінімальна кодова відстань дорівнює: $d_{min} = S + r + 1 = 3$ (де r - число виявлених, а S - число помилок, що виправляються).

Виберемо з Таблиці 8.7 значення k (для $m = 4, k = 3$).

Виберемо з таблиці 8.8 поліном $P(x)$ для $k = 3$: $P(x) = x^3 + x + 1$ (т.е. 1011)

Помножимо $G(x)$ на x^k : $G(x) \times x^3 = (x^3 + x^2 + 1) \times x^3 = x^6 + x^5 + x^3$ (т.е. 1101 \times 1000 = 1101000)

Розділимо $G(x) \times x^k$ на поліном $P(x)$:

$$\begin{array}{r|l}
 x^6 + x^5 + x^3 & x^3 + x + 1 \\
 \hline
 x^6 + x^4 + x^3 & \\
 \hline
 x^5 + x^4 & \\
 x^5 + x^3 + x^2 & \\
 \hline
 x^4 + x^3 + x^2 & \\
 x^4 + x^2 + x & \\
 \hline
 x^3 + x & \\
 x + x + 1 & \\
 \hline
 0 + 0 + 1 &
 \end{array}
 \qquad
 \begin{array}{r|l}
 1101000 & 1011 \\
 1011 & \\
 \hline
 1100 & 1111 \\
 1011 & \\
 \hline
 1110 & \\
 1011 & \\
 \hline
 1010 & \\
 1011 & \\
 \hline
 001 &
 \end{array}$$

В результаті отримуємо: $G(x) \times x^3 / P(x) = (x^3 + x^2 + x + 1) + 001 / (x^3 + x + 1)$ або: 1111 + 001 / 1011

В відповідності до $n = m + k$: $(x^3 + x^2 + x + 1) = Q(x) \rightarrow 1111$

$$1/(x^3 + x + 1) = R(x) / P(x) \rightarrow 001 / 1011$$

$$R(x) = 001$$

Шуканий многочлен $F(x)$ рівний:

$$F(x) = Q(x) \times P(x) = G(x) \times x^k + R(x) = x^6 + x^5 + x^3 + 1 \rightarrow 1101001$$

Для виявлення і виправлення помилок прийнята комбінація ділиться на утворюючий многочлен $P(x)$. Якщо залишок $R(x) = 0$, то комбінація прийнята без помилок. Наявність залишку свідчить про те, що комбінацію прийнято спотвореною.

Таблиця 8.7

Залежність між n , m і k

n	3	5	6	7	9...15	17...31	33...63	65...127
m	1	2	3	4	5...11	12...26	27...57	28...120
k	2	3	3	3	4	5	6	7

Таблиця 8.8

Незведені поліноми $P(x)$

Незведені многочлени	Їх еквіваленти
$q(x) = x + 1$	11
$q(x^2) = x^2 + x + 1$	111
$q(x^3) = x^3 + x + 1$	1011
$q(x^3) = x^3 + x^2 + 1$	1101
$q(x^4) = x^4 + x + 1$	10011
$q(x^4) = x^4 + x^3 + 1$	11001
$q(x^4) = x^4 + x^3 + x^2 + x + 1$	11111
$q(x^5) = x^5 + x^2 + 1$	100101
$q(x^5) = x^5 + x^3 + 1$	101001
$q(x^5) = x^5 + x^3 + x^2 + x + 1$	101111
$q(x^5) = x^5 + x^4 + x^2 + x + 1$	110111
$q(x^5) = x^5 + x^4 + x^3 + x + 1$	111011
$q(x^5) = x^5 + x^4 + x^3 + x^2 + 1$	111101
$q(x^6) = x^6 + x + 1$	1000011
$q(x^7) = x^8$	10001001
$q(x^8) = x^8 + x^4 + x^3 + x^2 + 1$	100011101
$q(x^9) = x^9 + x^4 + 1$	1000010001
$q(x^{10}) = x^{10} + x^3 + 1$	10000001001

Якщо число перевірочних символів менше за кількість інформаційних символів, то простіше комбінації циклічного коду отримати за допомогою матриці.

Утворююча матриця складається з двох підматриць – основної та додаткової.

Основна матриця – це поодинокі транспортована матриця Im . Вона містить m стовпців та m рядків (за кількістю інформаційних розрядів коду).

Додаткова матриця – матриця залишків. Вона утворюється шляхом поділу на утворюючий поліном $P(x)$ многочлена у вигляді одиниці з рядом нулів та виписуванням всіх проміжних залишків. Ця матриця містить k стовпців та m рядків.

$$C_{n,m} = \left\| \begin{array}{cccc|ccc} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{array} \right\|$$

$$C_{n,m} = \tilde{I}_m \cdot C_{(n-m),m}$$

За допомогою утворюючої матриці можна отримати будь-яку з $N = 2^m - 1$ дозволених комбінацій циклічного коду. Кожен із рядків утворюючої матриці – це вже комбінація циклічного коду. Інші $N-m$ комбінацій можна отримати додаванням по модулю 2 рядків матриці.

Приклад.

$$\oplus \begin{array}{r} 0001011 \\ \underline{0010110} \\ 0011101 \end{array}$$

$$\oplus \begin{array}{r} 0001011 \\ \underline{0100111} \\ 0101100 \end{array}$$

Кодери та декодери циклічних кодів в основному виконують операції множення та поділу многочленів.

Операція ділення на утворюючий поліном $P(x)$ здійснюється також за допомогою регістра зсуву, але в цьому випадку регістр має зворотні зв'язки, включені через суматори по модулю 2. Розподіл многочленів – це операція додавання по модулю 2 дільника з розрядами діленого, починаючи зі старшого розряду.

Для алгоритму кодування використовується k -розрядний регістр зсуву зі зворотними зв'язками через суматори за модулем 2. Число суматорів дорівнює числу членів утворюючого полінома, відмінних від нуля, мінус одиниця. Структурна схема кодувального пристрою наведена нижче на рис. 8.1.

Схема працює в такий спосіб. У початковому стані ключ П2 замкнутий, ключ П1 знаходиться у положенні «1». Інформаційна кодова комбінація $a(x)$, що має m розрядів (символів) подається на вхід через «1» П1 і одночасно в регістр зсуву Р1. За k тактів роботи регістра Р1 у ньому формується залишок $r(x)$, який є контрольними розрядами (символами) циклічного коду. Після цього ключ П2 розмикається, ключ П1 перекидається в положення «2» і контрольні символи за k тактів роботи регістру Р1 виводяться з нього, слідуючи за інформаційними символами.

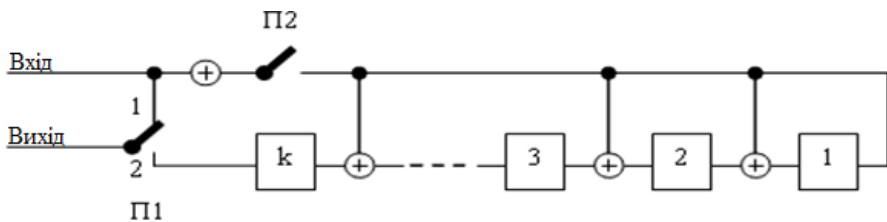


Рисунок 8.1 – Структурна схема кодуєчого пристрою

Необхідно пам'ятати, що інформаційна комбінація $a(x)$ надходить на вхід кодера у зворотній послідовності, тобто починаючи із старших розрядів. Комбінація циклічного коду на виході кодера має таку саму послідовність.

Схема декодування (утворюється за допомогою ділення на утворюючий многочлен) зображена на рис.8.2.



Рисунок 8.2 – Схема декодування циклічного коду.

АП – аналізатор помилок.

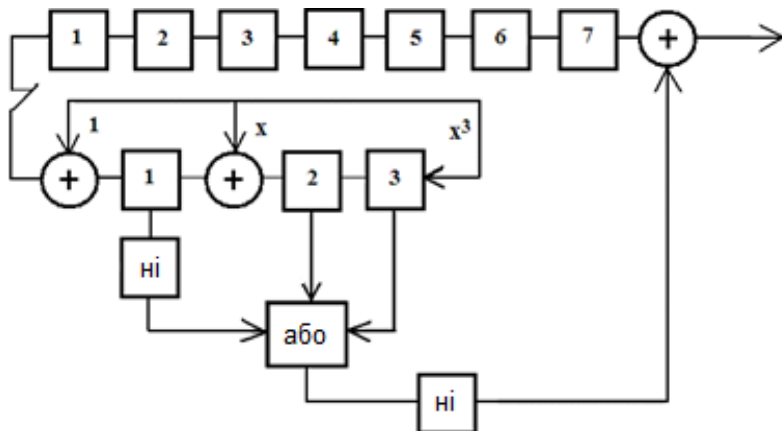


Рисунок 8.3 – Приклад схеми декодування методом ділення на поліном $P(x) = x^3+x+1$.

Вихідна комбінація подається в буферний регістр і одночасно через ключ в декодуєний регістр. Якщо з приходом останнього символу зафіксовано нульовий залишок, то помилок немає, якщо не нульовий, то існує помилка. Прийнята комбінація подається через вихідний суматор і спотворений сигнал виправляється аналізатором помилок (АП).

Коди Ріда-Соломона.

Коди Ріда-Соломона є субнабором циклічних кодів і є лінійними блоковими кодами. Код Ріда-Соломона специфікуються як $RS(n, m)$ s -бітних символів.

Це означає, що кодувальник сприймає m інформаційних символів по s біт кожний і додає символи парності для формування n символів кодового слова. Є $n - m$ символів парності s біт кожний. Декодер Ріда-Соломона може коригувати до $k/2$ символів, які містять помилки в кодовому слові, де $k = n - m$.

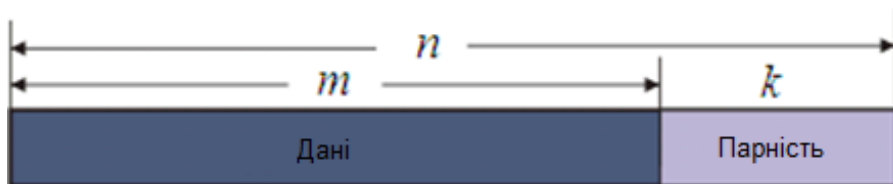


Рисунок 8.4 – Структура кодового слова $R-S$

Приклад. Популярним кодом Ріда-Соломона є $RS(255, 223)$ із 8-бітними символами. Кожне кодове слово містить 255 байт, з яких 223 є інформаційними та 32 байтами парності. Для цього коду: $n = 255$, $m = 223$, $s = 8$, $k = 32$, $k/2 = 16$.

Декодер може виправити будь-які 16 символів з помилками в кодовому слові: тобто помилки можуть бути виправлені, якщо кількість спотворених байт не перевищує 16.

Коди Ріда-Соломона базуються на спеціальному розділі математики – поля Галуа (GF) або кінцевих полях. Арифметичні дії (+, -, x, / і т.д.) над елементами кінцевого поля дають результат, який є елементом цього поля. Ці арифметичні операції для реалізації вимагають спеціального обладнання або спеціалізованого програмного забезпечення.

Загальна форма утворюючого полінома має вигляд: $P(x) = (x - a^i)(x - a^{i+1}) \dots (x - a^{i+k})$, кодове слово формується за допомогою операції: $F(x) = P(x) \cdot i(x)$, де $P(x)$ утворюючий поліном, $i(x)$ є інформаційним блоком, $F(x)$ – кодове слово, що називається простим елементом поля.

Тема 9. Теорія оптимального прийому сигналів

Оптимальний прийом сигналів. Ідеальний приймач. Критерій Байєса. Бінарна задача перевірки простих гіпотез. Помилки першого і другого роду. Критерій максимуму правдоподібності. Когерентний прийом. Некогерентний прийом. Оптимальний прийом неперервних повідомлень. Оптимальне оцінювання сигналу.

9.1 Методи оптимального прийому сигналів

У процесі передачі повідомлень у системах зв'язку здійснюються різні перетворення, основні з яких показано на спрощеній структурній схемі дискретної системи зв'язку (рис. 9.1).

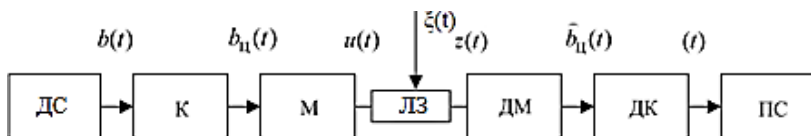


Рисунок 9.1 – Спрощена структурна схема дискретної системи зв'язку

Джерело сигналу ДС включає джерело повідомлень і перетворювач повідомлення $a(t)$ в первинний сигнал $b(t)$. Первинний сигнал піддається кодуванню (економному і/або стійкому до перешкод) в кодері К, після чого цифровий сигнал $b_u(t)$ надходить в модулятор М (передавач), що видає сигнал $u(t)$, пристосований за своїми характеристиками для передачі по лінії зв'язку ЛЗ. У лінії зв'язку відбувається спотворення сигналу та його взаємодія з перешкодою $\xi(t)$ (у найпростішому випадку адитивна), в результаті чого на вхід демодулятора ДМ (приймача) надходить коливання $z(t)$, що спостерігається. Демодулятор виконує функцію, обернену до модуляції, тому на його виході повинен в ідеальному випадку з'явитись сигнал $b_u(t)$. Проте внаслідок впливу перешкод результат демодуляції $b'_u(t)$, відрізняється в загальному випадку від сигналу $b_u(t)$ тому результат декодування також не збігається з первинним сигналом $b(t)$.

Для полегшення сприйняття надалі розглядається ідеалізований канал зв'язку без пам'яті, в якому відсутні спотворення спостережуваного сигналу, тоді:

$$z(t) = \sum_{k=-\infty}^{+\infty} b_u(t) \cdot s(t - k\tau) + \xi(t), \quad (9.1)$$

де $s(t)$ – посліжка тривалості τ , $\xi(t)$ – перешкода.

Завдання демодулятора полягає в тому, щоб по спостережуваному коливанню $z(t)$, прийняти рішення про переданий сигнал $b_u(t)$, таке, щоб

забезпечити максимальну достовірність. Правило (алгоритм) прийняття рішення – це закон перетворення $z(t)$ на $b'_u(t)$. Оскільки перешкода є випадковою, завдання побудови оптимального (найкращого) демодулятора є статистичним завданням і вирішується на основі методів теорії ймовірностей та математичної статистики (теорії статистичних рішень).

Матеріалом для прийняття рішення в демодуляторі служить в даному випадку реалізація коливання $z(t)$ на інтервалі тривалості T . Якби перешкода була відсутня, то ця реалізація збігалася б з елементарним сигналом (посилкою). Реалізація перешкоди при взаємодії з посилкою, зміщує її значення. Якщо зсув буде значним, демодулятор може помилятися. Помилка є випадковою подією, тому якість рішення можна характеризувати ймовірністю помилки.

Оптимальний прийом сигналів – це метод прийому, спрямований на максимальне розпізнавання та збереження інформації в умовах завад (шуму, інтерференції). Він дозволяє ефективно відокремити корисний сигнал від випадкових перешкод і спотворень, що виникають у каналі передачі. Для цього застосовуються методи обробки сигналів, що дозволяють мінімізувати ймовірність помилок при детекції (розпізнаванні) сигналу. Наприклад, для цифрових сигналів це методи зменшення ймовірності помилки прийняття бітів, а для аналогових – покращення співвідношення сигнал/шум (SNR).

Основні методи оптимального прийому сигналів

1. Кореляційне приймання.

Принцип роботи: кореляційний приймач використовує кореляцію (подібність) між прийнятим сигналом та еталонним сигналом для визначення його наявності. Він порівнює вхідний сигнал із заздалегідь відомим шаблоном (еталонним сигналом) і визначає схожість, що дозволяє виявити корисний сигнал навіть у присутності шуму.

Застосування: кореляційний приймач ефективний при детекції сигналів з відомою формою (наприклад, імпульсних сигналів або сигналів з кодовою структурою).

Переваги: забезпечує високу чутливість до слабких сигналів і є оптимальним для умов, де основним фактором завад є білий шум.

2. Фільтрація (Фільтр узгодження)

Суть фільтра узгодження: узгоджений фільтр налаштований на конкретну форму сигналу для того, щоб максимально збільшити співвідношення сигнал/шум (SNR) на виході приймача. В ідеалі він «узгоджений» з прийнятим сигналом для максимізації вихідного сигналу на момент детекції.

Застосування: використовується в системах зв'язку, де приймається відомий сигнал у присутності білого шуму. Узгоджені фільтри є оптимальними для задач, де потрібно виявити слабкі сигнали серед потужних завад.

Переваги: забезпечує максимальну ефективність у задачах виділення сигналу з шуму; широко використовується у цифрових комунікаціях, радіозв'язку, радіолокації.

3. Метод порогової детекції

Принцип роботи: метод порогової детекції порівнює потужність сигналу з певним порогом. Якщо сигнал перевищує цей поріг, він вважається корисним і передається на подальшу обробку. Поріг може бути фіксованим або динамічно змінюватися залежно від умов каналу.

Застосування: використовується для виявлення сигналів з низькою потужністю, коли потрібно швидко визначити наявність корисного сигналу.

Переваги: дозволяє налаштовувати чутливість приймача; особливо корисний для систем зі змінним рівнем шуму та сигналів.

4. Адаптивний прийом

Адаптивний прийом сигналів: приймач динамічно змінює свої параметри в залежності від стану каналу, що дозволяє автоматично компенсувати зміни каналу (наприклад, при фейдингу або мультиплікативних завадах).

Адаптивні фільтри: у разі, якщо характеристики каналу змінюються, застосовуються адаптивні фільтри, які підлаштовуються під зміни для мінімізації спотворень.

Застосування: у мобільних мережах, Wi-Fi системах, радіозв'язку з рухомими об'єктами, де умови передачі сигналів можуть швидко змінюватися.

Переваги: підвищення надійності зв'язку в умовах, де статичні методи прийому неефективні.

5. Метод вирівнювання

Суть вирівнювання: спрямований на усунення спотворень сигналу, викликаних багатошляховою інтерференцією та дисперсією. Вирівнювачі обчислюють зворотний профіль каналу, коригуючи сигнали так, щоб максимально наблизити їх до початкової форми.

Застосування: у стільникових мережах, супутниковому зв'язку, особливо в умовах багатошляхового розповсюдження сигналу.

Переваги: значно підвищує якість прийнятого сигналу за умови фейдингу або інших каналних спотворень.

9.2 Оптимальний приймач

Завдання синтезу оптимального приймача (демодулятора) ставиться так: знайти оптимальний алгоритм обробки та оптимальне правило, що забезпечують максимальну ймовірність безпомилкового (правильного) рішення.

Максимум цієї ймовірності називається потенційною стійкістю до перешкод, а приймач, що реалізує цей максимум, – **ідеальним приймачем**.

Алгоритм роботи приймача полягає у розбитті гільбертового простору реалізацій вхідного коливання на області, так що рішення приймається відповідно до того, якій області належить прийнята реалізація. Кількість областей дорівнює кількості різних кодових символів цієї системи зв'язку. Помилка виникає в тому випадку, якщо внаслідок впливу перешкоди реалізація потрапляє до «чужої» області. Оптимальний приймач найкраще розбиває простір реалізацій, так що середня ймовірність помилки мінімальна серед усіх можливих розбиття.

Кожна область відповідає припущенню (гіпотезі) про те, що переданий був один із можливих сигналів.

Приклад 1. Припустимо, що результатом обробки в двійковій системі зв'язку з амплітудною телеграфією є значення y , що відповідає закінченню інтервалу спостереження. Якщо в колюванні $z(t)$ присутній лише шум, з гаусовим розподілом та нульовим математичним очікуванням, то густина розподілу величини y має вигляд:

$$w_0(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^2}{2\sigma^2}} \quad (9.2)$$

Якщо крім шуму на вхід приймача надходить сигнал, то результат обробки має ненульове (для визначеності – додатне) середнє значення a , і густина розподілу величини y має вигляд:

$$w_1(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-a)^2}{2\sigma^2}} \quad (9.3)$$

Гіпотези, що відповідають виразам (17.2) та (17.3), є простими. Якщо середньоквадратичне відхилення σ невідоме, гіпотези є складними.

При цьому можливі помилки: може бути прийнято рішення D_j на користь j -ї гіпотези, тоді як справедливою є i гіпотеза. Така ситуація характеризується умовною ймовірністю помилки $p_{ij} = P\{D_j/H_i\}$. Різні помилки можуть завдавати різної шкоди, тому вводиться чисельна характеристика Π_{ij} , яка називається втратою, або ризиком.

$$R = \sum_{i=1}^K \sum_{j=1}^K p_i p_{ij} \Pi_{ij} \quad (9.3)$$

Кожна (i -а) гіпотеза характеризується деякою ймовірністю p_i здійснення, яка називається апіорною ймовірністю, підсумовуючи, можна ввести усереднену характеристику (критерій) якості прийняття рішення, що називається середнім ризиком.

Середній ризик є математичним очікуванням втрат, пов'язаних з прийняттям рішення.

Якщо апіорні ймовірності гіпотез точно відомі, а втрати призначені обґрунтовано, то приймач, що забезпечує найменший середній ризик, буде найвигіднішим. Критерій **мінімуму середнього ризику** називають також **критерієм Байєса**.

Іноді втрати, пов'язані з різними помилками, приймають рівними один одному $P_{ij} = P$; $P_{ii} = 0$; $i = 1, \dots, K$, тоді оптимальний байєсовий приймач забезпечує мінімальну середню ймовірність помилки (критерій ідеального спостерігача) і називається ідеальним приймачем Котельникова:

$$P_{\text{пом}} = \sum_{i=1}^K \sum_{\substack{j=1 \\ i \neq j}}^K P_i P_{ij} . \quad (9.4)$$

Якщо також прийняти рівні апіорні ймовірності гіпотез $p_i = 1/K$; $i = 1, \dots, K$, то критерій Байєса зводиться до критерію **мінімуму сумарної умовної ймовірності помилки**:

$$P_{\text{пом ум}} = \sum_{i=1}^K \sum_{\substack{j=1 \\ i \neq j}}^K P_{ij} . \quad (9.5)$$

Бінарна задача перевірки простих гіпотез

Найбільш просто завдання побудови оптимального приймача вирішується для випадку амплітудної телеграфії з пасивною паузою, що відповідає ухваленню рішення про те, що передавався символ 0 (сигналу немає) або символ 1 (сигнал є).

Передбачається, що перешкода в каналі є гаусівським шумом з нульовим середнім і відомою дисперсією, який взаємодіє з сигналом адитивно (додається). Результатом обробки спостережуваного коливання є випадкова величина y , яка може мати різний розподіл залежно від того, чи є сигнал у коливанні, а саме: розподіл при гіпотезі H_0 – «сигналу немає» – є гаусівським з нульовим середнім, а розподіл при гіпотезі H_1 – «сигнал є» – відрізняється зсувом на величину a , що залежить від способу обробки (наприклад, якщо обробка зводиться до отримання відліку в момент, коли несуче коливання

досягає максимуму, величина a являє собою його амплітуду). Значення a передбачається відомим.

Таким чином, гіпотези, що перевіряються, описуються двома умовними щільностями розподілу ймовірностей $w(y|H_0)$ і $w(y|H_1)$, зображеними на рис. 9.2.

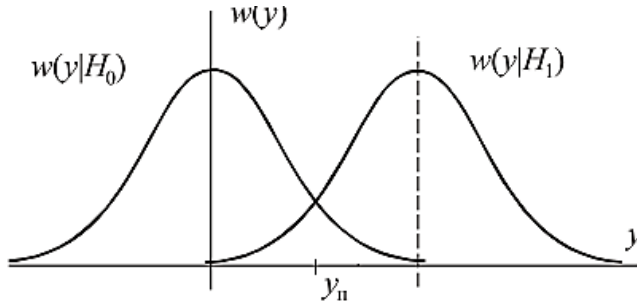


Рисунок 9.2 – Умовні щільності розподілу ймовірностей величини y за простих гіпотез

Приймач у такому разі має порівняти y з деяким фіксованим значенням (порогом) y_n і якщо y більший за поріг, прийняти рішення про наявність сигналу, в іншому випадку – про його відсутність, що можна коротко записати в наступній символічній формі:

$$\begin{aligned} y \geq y_n &\rightarrow \langle 1 \rangle; \\ y \leq y_n &\rightarrow \langle 0 \rangle. \end{aligned}$$

Яким би не був поріг y_n , очевидно, є деяка ненульова ймовірність p_{01} прийняти рішення про наявність сигналу за його фактичної відсутності. Ця ймовірність називається умовною ймовірністю **помилки першого роду** («хибної тривоги») і визначається виразом:

$$p_{01} = \int_{y_n}^{+\infty} w(y | H_0) dy. \quad (9.6)$$

Аналогічно, існує ненульова ймовірність прийняти рішення про відсутність сигналу, у той час як насправді він є (умовна ймовірність помилки другого роду, або пропуск сигналу):

$$p_{10} = \int_{-\infty}^{y_n} w(y | H_1) dy. \quad (9.7)$$

Аналіз рисунка 9.2 показує, що сума зазначених умовних ймовірностей мінімальна, якщо поріг y_n знаходиться як абсциса точки перетину умовних щільностей $w(y|H_0)$ і $w(y|H_1)$.

Очевидно, при такому виборі порога приймач є оптимальним за критерієм мінімуму сумарної умовної ймовірності помилки (9.5) і прийняття рішення ґрунтується на порівнянні значень функцій $w(y|H_0)$ і $w(y|H_1)$, при значенні y :

$$w(y|H_0) < w(y|H_1) \rightarrow \langle 1 \rangle$$

$$w(y|H_0) \geq w(y|H_1) \rightarrow \langle 0 \rangle$$

Це правило ухвалення рішення можна переписати також у формі:

$$\frac{w(y|H_1)}{w(y|H_0)} > 1 \rightarrow \langle 1 \rangle; \quad \frac{w(y|H_1)}{w(y|H_0)} \leq 1 \rightarrow \langle 0 \rangle .$$

Рішення приймається на користь тієї гіпотези, яка є більш правдоподібною при даному значенні y , тому відношення $\frac{w(y|H_1)}{w(y|H_0)}$ називається відношенням правдоподібності і позначається $\Lambda(y)$. Правило (9.6) називають **правилом максимальної правдоподібності**. Зауважимо, що критерій (9.5) часто називають **критерієм максимуму правдоподібності**.

Критерій ідеального спостерігача передбачає врахування апріорних ймовірностей гіпотез, і оптимальний у сенсі цього критерію приймач забезпечує мінімум середньої ймовірності помилки, тобто. найменшу суму безумовних ймовірностей помилок першого та другого роду. Інакше висловлюючись, порівнянню підлягають функції $w(y|H_0)$ і $w(y|H_1)$, помножені відповідні апріорні ймовірності. Правило прийняття рішення у такому приймачі можна записати у формі:

$$\frac{p_1 w(y|H_1)}{p_0 w(y|H_0)} > 1 \rightarrow \langle 1 \rangle; \quad \frac{p_1 w(y|H_1)}{p_0 w(y|H_0)} \leq 1 \rightarrow \langle 0 \rangle$$

Використовуючи поняття відношення правдоподібності, можна записати правило як:

$$\Lambda(y) > \frac{p_0}{p_1} \rightarrow \langle 1 \rangle; \quad \Lambda(y) \leq \frac{p_0}{p_1} \rightarrow \langle 0 \rangle ,$$

при цьому відношення правдоподібності порівнюється з граничним значенням, що залежить від апріорних ймовірностей.

Нарешті, у разі байесовського критерію рішення приймається за правилом:

$$\frac{\Pi_{10} p_1 w(y|H_1)}{\Pi_{01} p_0 w(y|H_0)} > 1 \rightarrow \langle 1 \rangle; \quad \frac{\Pi_{10} p_1 w(y|H_1)}{\Pi_{01} p_0 w(y|H_0)} \leq 1 \rightarrow \langle 0 \rangle ,$$

або

$$\Lambda(y) > \frac{p_0 \Pi_{01}}{p_1 \Pi_{10}} \rightarrow \langle 1 \rangle; \quad \Lambda(y) \leq \frac{p_0 \Pi_{01}}{p_1 \Pi_{10}} \rightarrow \langle 0 \rangle.$$

Отже, у всіх випадках оптимальний приймач (демодулятор, або вирішальний пристрій) «влаштований однаково»: для спостережуваного значення y , що залежить від прийнятої реалізації $z(t)$, обчислюється значення відношення правдоподібності, яке порівнюється з порогом; поріг дорівнює: $\frac{p_0 \Pi_{01}}{p_1 \Pi_{10}}$ для оптимального приймача за критерієм мінімуму середнього ризику; p_0/p_1 для ідеального приймача Котельникова та 1 для приймача максимальної правдоподібності.

Слід зазначити, що іноді зручніше обчислювати не відношення правдоподібності, а його логарифм.

9.3 Прийом відомого сигналу (когерентний прийом)

Розглянемо прийняття рішення у системі зв'язку за таких умов: синхронізація є точною і форма сигналу на інтервалі спостереження достеменно відома, невідома лише сама наявність чи відсутність сигналу в коливанні. (Ця ситуація найбільш близька до реальності в кабельних лініях зв'язку, де умови поширення сигналів відомі та практично незмінні.) така ситуація називається **когерентний прийом**. Він передбачає, що приймач має точну інформацію про фазу несучої частоти сигналу, що дає змогу приймати сигнали зі збереженням фазової інформації. Це особливо важливо для сигналів із фазовою модуляцією, де фаза несе основну частину інформації. Когерентний прийом забезпечує максимальну ефективність передачі, оскільки дозволяє розпізнавати амплітуду і фазу сигналу одночасно. Вважатимемо, що на інтервалі спостереження незалежно від сигналу присутній гаусовий шум з нульовим середнім і спектральною щільністю потужності $N_0/2$, постійною у смузі частот $-F < f < F$ («квазібілий» шум). Правило виявлення сигналу відомої форми на фоні гаусівського квазібілого шуму, оптимальне в сенсі критерію максимуму правдоподібності, засноване на порівнянні з нульовим порогом величини:

$$\int_0^T y(t)s(t)dt - \frac{E}{2}, \quad (9.8)$$

де $E = \int_0^T s^2(t)dt$ – енергія сигналу.

Перший доданок у виразі (17.8) називається кореляційним інтегралом, так як збігається за формою з виразом взаємно кореляційної функції сигналу при нульовому зсуві. Енергія сигналу відома, і для виявлення можна

порівнювати значення кореляційного інтегралу (випадкове через випадковість реалізації $z(t)$) з порогом, рівним $E/2$.

Приклад 2. У провідних системах зв'язку з амплітудною телеграфією можуть застосовуватися посліжки у вигляді прямокутного відеоімпульсу. Припустимо, що сигнал, що відповідає символу «1», є прямокутним відеоімпульсом з амплітудою a і тривалістю T . Тоді кореляційний інтеграл має вигляд

$$\int_0^T z(t)s(t)dt = a \int_0^T z(t)dt ,$$

поріг дорівнює $E/2 = a^2T/2$, тоді вирішальне правило має вигляд:

$$\int_0^T z(t)dt > \frac{aT}{2} \rightarrow \text{«1»}, \int_0^T z(t)dt \leq \frac{aT}{2} \rightarrow \text{«0»} .$$

Структурна схема приймача показано на рисунку 17.4. Постійна часу інтегруючого кола повинна бути набагато більшою за тривалість посліжки T . У цьому випадку початкову ділянку експоненти $a(1 - e^{-t/(RC)})$, що відображає заряд ємності, можна апроксимувати прямою лінією з тангенсом кута нахилу $a/(RC)$, рівним похідній експоненти у нулі. Тоді за час T напруга на вході вирішального пристрою, обумовлена сигналом, складе $aT/(RC)$, а значення порогу має дорівнювати $aT/(2RC)$.

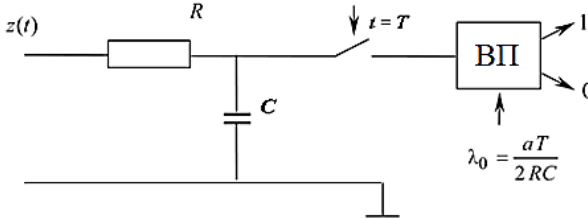


Рисунок 9.3 – Структурна схема приймача прямокутного відеоімпульсу

Когерентний прийом широко використовується в цифрових системах зв'язку, таких як 4G і 5G, де важлива висока точність передачі даних.

9.4 Некогерентний прийом

На практиці іноді не вдається забезпечити умови для когерентного прийому сигналів, так як один або кілька параметрів сигналу виявляються невідомими. Така ситуація типова, наприклад, для систем супутникового зв'язку, радіозв'язку з рухомими об'єктами тощо, оскільки відстань між передавачем і приймачем змінюється випадковим чином. Це призводить, зокрема, до того, що змінюється початкова фаза коливання. Якщо зміна

відбувається настільки повільно, що сусідні посилки мають практично однакову початкову фазу, її можна оцінити і оцінку використовувати замість точного значення при організації прийому. Такий прийом називають квазікогерентним.

Якщо ж початкова фаза змінюється (флюктує) швидко або пристрій оцінювання виявляється занадто складним, тоді розглядається завдання прийому сигналу з випадковою початковою фазою або **некогерентного прийому**.

Сигнал при некогерентному прийомі відомий з точністю до початкової фази, тому позначимо $s(t, \varphi)$ і запишемо:

$$s(t, \varphi) = \operatorname{Re} \left\{ \dot{s}(t) \cdot e^{-j\varphi} \right\}.$$

Кореляційний інтеграл можна записати в вигляді

$$\int_0^T z(t) s(t, \varphi) dt = V \cos(\psi + \varphi),$$

тоді відношення правдоподібності буде:

$$\ln \Lambda = \frac{2}{N_0} V \cos(\psi + \varphi) - \frac{1}{N_0} E, \quad \Lambda = e^{\frac{2}{N_0} V \cos(\psi + \varphi)} e^{-\frac{1}{N_0} E}. \quad (9.9)$$

Вважаючи, що початкова фаза сигналу є випадковою величиною, що має рівномірний в інтервалі $(0, 2\pi)$ розподіл, виконаємо усереднення відношення правдоподібності по ансамблю:

$$\bar{\Lambda} = e^{-\frac{E}{N_0}} \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{\frac{2V}{N_0} \cos(\psi + \varphi)} d\varphi. \quad (9.10)$$

Врахуємо відоме співвідношення

$$\frac{1}{2\pi} \int_0^{2\pi} e^{a \cos(\psi + \varphi)} d\varphi = I_0(a), \quad (9.11)$$

де $I_0(a)$ – модифікована функція Бесселя нульового порядку, тоді

$$\bar{\Lambda} = e^{-\frac{E}{N_0}} \cdot I_0 \left(\frac{2V}{N_0} \right). \quad (9.12)$$

Правило некогерентного прийому сигналу з випадковою рівномірною початковою фазою на тлі гаусівського шуму має бути засноване на порівнянні величини $\bar{\Lambda}$ з деяким порогом, а правило розрізнення двох сигналів – на порівнянні двох відношень правдоподібності між собою.

Припустимо, що розглядається прийом двох сигналів $s_1(t)$ і $s_0(t)$. Порівняння усереднених відношень правдоподібності можна замінити порівнянням їхніх логарифмів:

$$\ln I_0 \left(\frac{2V_1}{N_0} \right) - \frac{E_1}{N_0} \underset{0}{\geq} \ln I_0 \left(\frac{2V_0}{N_0} \right) - \frac{E_0}{N_0}, \quad (9.13)$$

або порівнянням з порогом різниці логарифмів:

$$\ln I_0 \left(\frac{2V_1}{N_0} \right) - \ln I_0 \left(\frac{2V_0}{N_0} \right) \underset{0}{\geq} \frac{E_1 - E_0}{N_0}. \quad (9.14)$$

Алгоритм сильно спрощується, якщо енергії сигналів рівні, в цьому випадку в силу монотонності функції I_0 можна порівнювати між собою величини:

$$V_1 \underset{0}{\geq} V_0. \quad (9.15)$$

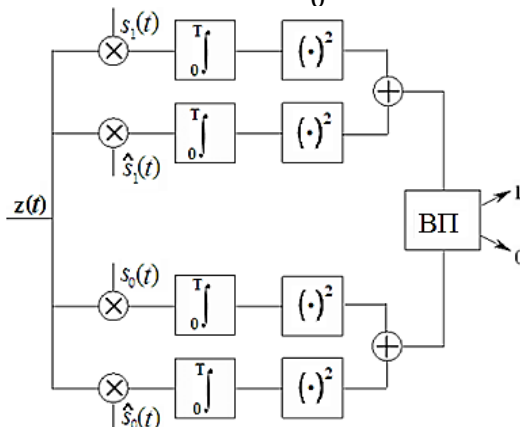


Рисунок 9.4 – Структура некогерентного приймача двох сигналів з рівними енергіями

9.5 Оптимальний прийом неперервних повідомлень

Неперервні повідомлення (наприклад, мова, музика тощо) можуть передаватися по каналу зв'язку безпосередньо (наприклад, місцевою проводовою радімережею, телефонним кабелем) або за допомогою модуляції. У першому випадку сигнал $s(t)$, що передається каналом, може збігатися з повідомленням (первинним сигналом) $b(t)$ або бути пов'язаний з ним простою пропорційною залежністю, у другому – сигнал $s[t, b(t)]$, що передається, є функцією повідомлення, в загальному випадку нелінійною (рис. 9.5).

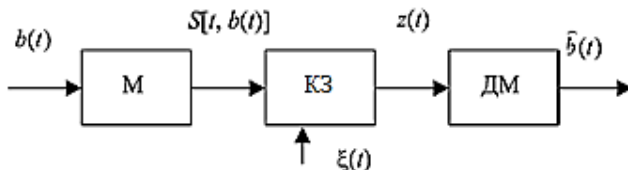


Рисунок 9.5 – Структура системи передачі неперервних повідомлень

Коливання на вході демодулятора $z(t) = s[t, b(t)] + \xi(t)$ являє собою в найпростішому випадку суму сигналу, що передається, і шуму $\xi(t)$.

Завдання демодулятора полягає у знаходженні оцінки первинного сигналу (повідомлення), найкращої у сенсі обраного критерію близькості. Як критерій часто використовують середній квадрат помилки:

$$\overline{\varepsilon^2} = [\hat{b}(t) - b(t)]^2, \quad (9.16)$$

У системах телеметрії використовується критерій максимальної помилки, у радіомовленні – збільшення вихідного відношення сигнал/шум проти вхідного, критерій розбірливості мовних повідомлень тощо.

9.6 Оптимальне оцінювання сигналу

Оцінювання сигналу як функції часу – досить складне завдання. У багатьох випадках його можна звести до більш простого завдання оцінювання одного або декількох параметрів сигналу.

Найпростішим завданням, пов'язаним з оцінюванням параметрів сигналу, є оцінка параметра, що є постійним або настільки повільно змінюється в часі, що на інтервалі спостереження його можна вважати постійним.

Розглянемо завдання оцінювання єдиного параметра λ , який розглядається як випадкова величина, що має апріорний розподіл із щільністю $w(\lambda)$.

Реалізація цієї випадкової величини є значення, постійне на інтервалі $(0, T)$ спостереження коливання $z(t) = s[t, \lambda] + \xi(t)$.

Правило оцінювання – це алгоритм обробки коливання, результатом виконання якого є значення оцінки параметра λ .

Для оцінювання одного й того ж параметра може бути безліч алгоритмів, які видають різні оцінки. Для порівняння алгоритмів оцінювання між собою та вибору найкращого використовують такі показники: незміщеність; спроможність; ефективність.

1) **Незміщеність.** Оцінка називається незміщеною, якщо виконується умова $\overline{\tilde{\lambda}} - \lambda = 0$, що означає, що при будь-якому значенні параметра умовне математичне очікування оцінки дорівнює цьому значенню. Інакше кажучи, незміщеність означає відсутність систематичної помилки оцінювання. В іншому випадку оцінка називається зміщеною. Слід зазначити, що зміщені оцінки також знаходять застосування, якщо зміщення досить мале або прагне нуля зі збільшенням часу спостереження або потужності сигналу.

2) **Спроможність.** Оцінка називається спроможною, якщо при необмеженому зростанні часу спостереження оцінка сходиться за ймовірністю до значення параметра:

$$\lim_{T \rightarrow \infty} P\{\tilde{\lambda} - \lambda \geq \Delta\} = 0 \text{ при будь-якому } \Delta > 0, \quad (9.17)$$

(тут $P\{A\}$ означає ймовірність події A).

Зміщена оцінка може бути спроможною, якщо її зміщення прагне нуля при $T \rightarrow \infty$. Для спроможної оцінки, очевидно, дисперсія помилки прагне нуля.

$$\lim_{T \rightarrow \infty} \overline{|\tilde{\lambda} - \lambda|^2} = 0. \quad (9.18)$$

3) **Ефективність.** Незміщена оцінка називається ефективною, якщо серед усіх оцінок, отриманих при заданому часі спостереження усілякими алгоритмами оцінювання, вона забезпечує найменшу дисперсію помилки:

$$\overline{|\tilde{\lambda} - \lambda|^2} = \min. \quad (9.19)$$

Класичний підхід до оцінювання параметрів сигналів ґрунтується на формулі Байеса для апостеріорної щільності розподілу ймовірностей параметра, що оцінюється.

$$w(\lambda | z) = \frac{w(\lambda)w(z | \lambda)}{w(z)}, \quad (9.20)$$

де $w(\lambda)$ – апіорна ЩРІ параметра λ ; $w(\lambda|z)$ – умовна ЩРІ спостережуваного процесу при заданому значенні λ , що розглядається як функція від λ при даному z (функція правдоподібності); $w(z)$ – при фіксованій реалізації z стала величина.

Вираз (9.20) показує, що знаючи апіорну щільність $w(\lambda)$ і спостерігаючи реалізацію процесу z , можна отримати уточнене уявлення про значення параметра λ . На рисунку 9.6 показані приклади апіорної та апостеріорної ЩРІ параметра (справжнє значення параметра позначено λ_0).

Апостеріорний розподіл загострюється порівняно з апіорним розподілом, тому що реалізація z містить додаткову інформацію про параметр, що зменшує вихідну невизначеність апіорної ЩРІ.

Апостеріорний розподіл містить всю інформацію про параметр, яку можна отримати з спостережуваної реалізації та апіорних даних. Тому правило оцінювання має використовувати апостеріорну ЩРІ, а спосіб її використання залежить від обраного критерію якості оцінки.

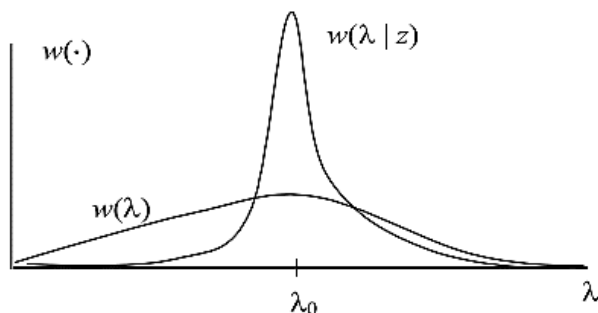


Рисунок 9.6 – Апіорна та апостеріорна ЩРІ оцінюваного параметра

Помилки оцінювання параметра призводять до різних наслідків, тому для їхнього обліку вводиться функція втрат $L(\tilde{\lambda} - \lambda)$, що залежить від різниці оцінки та істинного значення параметра.

Усереднюючи функцію втрат за апостеріорним розподілом параметра, отримуємо кількісну характеристику, яка називається апостеріорним (умовним) ризиком.

$$r(\tilde{\lambda}, z) = \int_{(\lambda)} L(\tilde{\lambda} - \lambda) w(\lambda | z) d\lambda, \quad (9.21)$$

що описує втрати, пов'язані з отриманням оцінки при спостереженні реалізації z . Усереднення апостеріорного ризику (9.21) за всіх реалізаціях приводить до середнього ризику.

$$R(\tilde{\lambda}) = \int_{(z)} w(z) \left[\int_{(\lambda)} L(\tilde{\lambda} - \lambda) w(\lambda | z) d\lambda \right] dz. \quad (9.22)$$

Правило оцінювання, якому відповідає найменший середній ризик, називається байєсовим, а відповідна оцінка – байєсовою, або оцінкою за критерієм мінімуму середнього ризику. Правило, оптимальне у сенсі мінімуму середнього ризику, виходить із умови мінімізації умовного ризику (9.22).

Часто використовують квадратичну функцію втрат $L(\tilde{\lambda} - \lambda) = (\tilde{\lambda} - \lambda)^2$, тоді

$$r(\tilde{\lambda}, z) = \int_{(\lambda)} L(\tilde{\lambda} - \lambda)^2 w(\lambda | z) d\lambda = \overline{(\tilde{\lambda} - \lambda)^2}, \quad (9.23)$$

Таким чином, апостеріорний ризик дорівнює середньому квадрату помилки (а якщо оцінка незміщена, то дисперсії помилки). Байєсова оцінка в цьому випадку стає оцінкою мінімуму середньоквадратичної помилки. Для знаходження правила розкриємо дужки у виразі (9.23):

$$r(\tilde{\lambda}, z) = \overline{(\tilde{\lambda} - \lambda)^2} = \tilde{\lambda}^2 - 2\tilde{\lambda} \int_{(\lambda)} \lambda w(\lambda | z) d\lambda + \int_{(\lambda)} \lambda^2 w(\lambda | z) d\lambda. \quad (9.24)$$

Диференціюючи отриманий вираз по λ і прирівнюючи результат до нуля, отримуємо правило:

$$\tilde{\lambda} = \int_{(\lambda)} \lambda w(\lambda | z) d\lambda. \quad (9.25)$$

Таким чином, оцінка, оптимальна в сенсі мінімуму середньоквадратичної помилки, дорівнює апостеріорному середньому значення параметра.

Крім квадратичної, на практиці часто використовується проста функція втрат:

$$L(\tilde{\lambda} - \lambda) = \text{const} - \delta(\tilde{\lambda} - \lambda), \quad (9.26)$$

$$\int_{(\lambda)} [\text{const} - \delta(\tilde{\lambda} - \lambda)] \cdot \lambda w(\lambda | z) d\lambda = \text{const} \cdot w(\lambda | z) \Big|_{\lambda=\tilde{\lambda}}$$

Очевидно, цей вираз досягає мінімуму, якщо як оцінку $\tilde{\lambda}$ використовувати значення параметра, що відповідає максимуму апостеріорної ЩРІ $w(\lambda|z)$. Така оцінка називається МАІ-оцінкою (оцінкою максимуму апостеріорної імовірності).

В багатьох задачах априорна ЩРІ параметра невідома, тоді приймають її рівною константі і максимізують функцію правдоподібності $w(\lambda|z)$. Отримані в такий спосіб оцінки називаються оцінками максимальної правдоподібності.

Приклад 2. Нехай спостерігається коливання $z(t) = \gamma s(t) + \zeta(t)$, де $s(t)$ – точно відомий сигнал; γ – амплітудний множник, що підлягає оцінюванню; $\zeta(t)$ – гаусівський шум з нульовим середнім та спектральною щільністю потужності $N_0/2$ постійний у смузі частот $-F < f < F$, («квазібілий» шум). Знайдемо правило оцінювання параметра, оптимальне за критерієм максимальної правдоподібності.

Візьмемо n відліків коливання, що спостерігається на інтервалі спостереження T з кроком $\Delta t = 1/2F = T/n$, при цьому відліки шуму є некорельованими. Спільна щільність розподілу ймовірності взятих відліків дорівнює

$$w(z_1, \dots, z_n | \gamma) = \frac{1}{(\sqrt{2\pi\sigma})^n} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^n (z_k - \gamma \cdot s_k)^2},$$

де $\sigma^2 = N_0 F = N_0 / (2\Delta t)$. Спрямовуючи Δt до нуля ($n \rightarrow \infty$), запишемо функцію правдоподібності

$$w(z | \gamma) = C \cdot \exp \left\{ -\frac{1}{N_0} \int_0^T [z(t) \cdot \gamma s(t)]^2 dt \right\},$$

де C – константа, несуттєва для завдання оцінювання.

Для знаходження правила оцінювання слід продиференціювати функцію правдоподібності або, що простіше, її логарифм та прирівняти результат до нуля. Рівняння правдоподібності $\frac{d[\ln w(z | \gamma)]}{d\gamma} = 0$ що виходить

при цьому, для даного випадку має вигляд $\int_0^T [z(t) - \gamma s(t)] \cdot s(t) dt = 0$, звідки

$$\gamma \int_0^T s^2(t) dt = \int_0^T z(t) \cdot s(t) dt.$$

Рішенням цього рівняння щодо γ є оцінка $\tilde{\gamma}$, яка визначається виразом

$$\tilde{\gamma} = \frac{1}{E} \int_0^T z(t) \cdot s(t) dt, \quad (9.27)$$

де $E = \int_0^T s^2(t) dt$ – енергія сигналу, відома за умовою завдання.

Якість отриманої МП-оцінки можна оцінити, підставивши в (9.27) вираз для $z(t)$:

$$\tilde{\gamma} = \frac{\int_0^T [\gamma s(t) + \xi(t)] \cdot s(t) dt}{E} = \frac{\gamma E}{E} + \frac{\int_0^T \xi(t) \cdot s(t) dt}{E} = \gamma + \frac{1}{E} \int_0^T \xi(t) \cdot s(t) dt. \quad (9.28)$$

Другий доданок є помилкою оцінювання, причому дисперсія інтеграла дорівнює $N_0 E/2$, тому дисперсія помилки дорівнює $N_0/2E$.

Таким чином, оцінка тим точніша, чим більша енергія сигналу (для гармонійного сигналу $s(t)$ збільшення енергії еквівалентне збільшенню тривалості інтервалу спостереження) і чим менша спектральна щільність потужності перешкоди.

З виразу (9.28) видно, що оцінка незміщена, оскільки $\xi(t)$ має нульове математичне очікування. Враховуючи незміщеність та прагнення дисперсії до нуля зі збільшенням інтервалу спостереження, можна зробити висновок, що оцінка є спроможною. Крім того, можна показати, що оцінка також є ефективною.

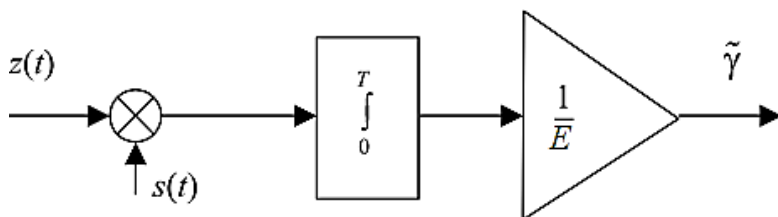


Рисунок 9.7 – Структура пристрою оцінювання амплітуди сигналу

Тема 10. Перешкодостійкість систем передачі повідомлень

Перешкодостійкість систем передачі дискретних і неперервних повідомлень. Потенційна перешкодостійкість когерентного та некогерентного прийому. Поріг завадостійкості.

10.1 Перешкодостійкість систем передачі дискретних повідомлень

Завдання, що стоять перед технікою зв'язку, зводяться до забезпечення двох основних характеристик:

- 1) ефективності зв'язку;
- 2) завадостійкості зв'язку.

Ефективність зв'язку: передати найбільшу кількість інформації найбільш економним способом.

Швидкість передачі по каналу зв'язку вимірюється кількістю інформації, що передається в одиницю часу. Максимальна швидкість передачі інформації, яку може забезпечити канал зв'язку з даними характеристиками, називається його пропускнуою здатністю.

Перешкодостійкість зв'язку: здатність системи зберігати свої функції незмінними або такими, що змінюються в допустимих межах при дії перешкод.

Кількісно завадостійкість оцінюється різними показниками, що використовують ймовірнісний опис сигналів та перешкод. Наприклад, застосовуються такі показники, як відношення сигнал/шум на вході та виході приймального пристрою, імовірність правильного виявлення сигналу, імовірність помилки, тощо.

В теорії завадостійкості розрізняють дві основні завдання: аналізу та синтезу.

Завдання аналізу присвячено розрахунку показників завадостійкості існуючих (розроблених) систем.

У цьому випадку, вважаючи відомими імовірнісний опис сигналу і перешкоди на вході, визначають імовірнісні характеристики вихідного процесу, а по ньому – показники стійкості до перешкод. Це завдання, за своєю суттю, зводиться до аналізу проходження випадкового процесу через лінійні та нелінійні кола, з яких складається система.

Завдання синтезу присвячена визначенню структурної схеми системи або, у більш простому варіанті, структурної схеми радіоприймального пристрою, яка мала б найкращі, або оптимальні показники перешкодостійкості при заданому призначенні пристрою і при відомому ймовірнісному описі сигналу і перешкоди на вході. Завдання синтезу називаються також завданнями оптимального радіоприймання, які поділяють на чотири часткові задачі: виявлення сигналу, розрізнення сигналів, оцінки параметрів сигналу, фільтрації сигналу або повідомлень.

В задачі виявлення потрібно найкраще за заданим критерієм оптимальності на підставі спостереження процесу відповісти на питання, чи містить спостережуваний процес сигнал разом з перешкодою або він є тільки перешкодою.

В задачі розрізнення потрібно за заданим критерієм оптимальності найкращим чином відповісти на питання, який саме сигнал разом з перешкодою присутній в спостережуваному процесі.

В задачі оцінки параметрів потрібно найкраще за заданим критерієм оцінити невідомі параметри сигналу, при цьому вважається, що у процесі разом із перешкодою існує сигнал з одним або декількома невідомими параметрами (параметр є випадковою величиною постійною на інтервалі спостереження).

До задачі оцінки параметрів тісно примикає завдання розрізнення сигналу, коли вважається, що разом з перешкодою в спостережуваному процесі можуть існувати один або два сигнали, невідомі параметри яких незначно відрізняються між собою. Однак скільки цих сигналів – один чи два – наперед невідомо. Потрібно, збільшуючи різницю між параметрами сигналу, визначити ту найменшу різницю, при якій настає впевнене розрізнення сигналів.

В задачі оптимальної фільтрації потрібно за заданим критерієм оптимальності в кожен момент часу дати найкращу оцінку змінюваному параметру. Вважається, що у спостережуваному процесі разом із перешкодою існує сигнал із змінним у часі, відповідно до випадкового закону модуляції, параметром, тобто параметр є випадковою функцією часу.

10.2 Потенційна завадостійкість когерентного прийому

Потенційною завадостійкістю називається максимум ймовірності правильного рішення, досяжний за заданих умов прийому сигналів і впливу перешкод (шумів).

Визначимо потенційну завадостійкість прийому двох сигналів, $s_0(t)$ і $s_1(t)$ відомої форми на тлі білого гауссового шуму при рівних апіорних ймовірностях сигналів.

Алгоритм прийняття рішення у приймачі, що реалізує критерій максимуму правдоподібності, коротко запишемо у вигляді

$$\int_0^T z(t)s_1(t)dt - \frac{E_1}{2} \stackrel{1}{\cong} \int_0^T z(t)s_0(t)dt - \frac{E_0}{2}. \quad (10.1)$$

Цей вираз можна привести до вигляду:

$$\int_0^T z(t)[s_1(t) - s_0(t)]dt \leq \frac{1}{0} \frac{(E_1 - E_0)}{2}. \quad (10.2)$$

Помилки прийому полягають у тому, що при передачі першого сигналу приймається рішення про прийом другого і навпаки. Оскільки гауссовий розподіл симетричний і апіорні ймовірності рівні.

Знайдемо умовну імовірність помилки, тобто ймовірність події, що полягає у прийнятті рішення про наявність сигналу $s_0(t)$ за умови, що в коливанні, що спостерігається, присутній сигнал $s_1(t)$. Ця подія відповідає виконанню нерівності

$$\int_0^T [s_1(t) + \xi(t)] \cdot [s_1(t) - s_0(t)]dt < \frac{(E_1 - E_0)}{2},$$

яку можна переписати у вигляді

$$\int_0^T s_1^2(t)dt + \int_0^T \xi(t) \cdot [s_1(t) - s_0(t)]dt - \int_0^T s_1(t) \cdot s_0(t)dt < \frac{1}{2} \int_0^T s_1^2(t)dt - \frac{1}{2} \int_0^T s_0^2(t)dt.$$

Провівши перетворення, отримаємо:

$$\int_0^T \xi(t) \cdot [s_1(t) - s_0(t)]dt < -\frac{1}{2} \int_0^T [s_1(t) - s_0(t)]^2 dt. \quad (10.3)$$

Ліва частина нерівності є випадковою величиною (це інтеграл за часом від випадкового процесу $\xi(t)$ з вагою, що дорівнює різниці сигналів $s_\Delta(t) = [s_1(t) - s_0(t)]$), що має нормальний розподіл (оскільки процес $\xi(t)$ гауссовий) з нульовим середнім; позначимо її v , її дисперсія (середній квадрат), дорівнює:

$$D_v = \overline{v^2} = \int_0^T \int_0^T \overline{\xi(t_1) \cdot \xi(t_2)} \cdot s_\Delta(t_1) \cdot s_\Delta(t_2) dt_1 dt_2 = \frac{N_0 \cdot E_\Delta}{2}. \quad (10.4)$$

Ймовірність виконання нерівності (10.3) – це, очевидно, ймовірність того, що нормальна випадкова величина з нульовим середнім та дисперсією $N_0 E_\Delta / 2$ набуває значення менше, ніж $E_\Delta / 2$. Ця ймовірність дорівнює:

$$p_{10} = \int_{-\infty}^{-E_\Delta/2} \frac{1}{\sqrt{2\pi D_v}} e^{-\frac{v^2}{2D_v}} dv = \int_{\alpha}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, \quad (10.5)$$

де $t = -\frac{v}{\sqrt{D_v}}$ – центрована нормальна випадкова величина з одиничною

дисперсією, а $\alpha = \frac{E_\Delta}{2\sqrt{D_v}}$ – додатнє число.

$$\text{Очевидно, } p_{10} \text{ залежить лише від } \alpha = \frac{E_\Delta}{2\sqrt{D_v}} = \sqrt{\frac{E_\Delta}{2N_0}},$$

тому можна ввести функцію $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt = 1 - \Phi(x)$, де

$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$ – інтеграл імовірності, і записати:

$$p_{10} = Q\left(\sqrt{\frac{E_\Delta}{2N_0}}\right), \quad (p_{10} = p_{01}). \quad (10.6)$$

Таким чином, умовна ймовірність помилки, що дорівнює середній ймовірності помилки при когерентному прийомі сигналів на фоні білого шуму, визначається енергією різницевого сигналу $s\Delta(t)$ та спектральної щільністю потужності шуму N_0 .

Розглянемо потенційну стійкість до перешкод двійкового когерентного приймача максимальної правдоподібності для різних способів модуляції, вважаючи, що енергія сигналу E фіксована.

Представляє практичний інтерес порівняльна оцінка потенційної заводостійкості сигналів дискретної модуляції: амплітудної (ДАМ), частотної (ДЧМ), фазової (ДФМ).

При ДАМ символу «1» відповідає сигнал $S_1(t)$, а символу «0» відповідає сигнал $S_2(t) = 0$ («пасивна пауза»), отже норма різницевого сигналу дорівнює:

$$\|\Delta\vec{S}\|_{\text{ДАМ}} = \|\vec{S}_1 - \vec{S}_2\| = \|\vec{S}_1\| = \sqrt{\int_0^T S_1^2(t) dt} = \sqrt{E_S}. \quad (10.7)$$

При ДЧМ та ДФМ символам «1» і «0» відповідають сигнали $S_1(t)$ і $S_2(t)$, енергія яких однакова. При цьому для норми різницевого сигналу маємо:

$$\|\Delta\vec{S}\| = \sqrt{\int_0^T [S_1(t) - S_2(t)]^2 dt} = \sqrt{2E_S(1 - b_{12})}; \quad b_{12} = \frac{\int_0^T [S_1(t) - S_2(t)] dt}{E_S}, \quad (10.8)$$

де b_{12} – нормований коефіцієнт кореляції ($-1 \leq b_{12} \leq 1$).

Для ортогональних сигналів ДЧМ $b_{12} = 0$, $\|\Delta\vec{S}\|_{\text{ДЧМ}} = \sqrt{2}\|\vec{S}\| = \sqrt{2E_S}$,

а для протилежних сигналів ДФМ $b_{12} = -1$, $\|\Delta\vec{S}\|_{\text{ДФМ}} = \sqrt{2}\|\vec{S}\| = \sqrt{4E_S}$.

Сигнали з $b_{12} = 1$ є однаковими, тобто їх неможливо розрізнити. Їх $P_{\text{ном}} = 0,5$ – це еквівалент обриву каналу зв'язку.

Тобто енергія різницевого сигналу $E_{\Delta S} = \|\vec{S}_1 - \vec{S}_2\|$ для ДЧМ вдвічі, а ДФМ – вчетверо більше проти ДАМ. Відповідно зростає і завадостійкість.

1. Амплітудна телеграфія з пасивною паузою.

У цьому випадку $s_0(t) = 0$ та енергія різницевого сигналу дорівнює E (норма дорівнює $E/2$), рис. 17.7.а. Отже, потенційна стійкість до перешкод визначається середньою ймовірністю помилки.

$$P_{\text{ОШ}}^{\text{АТ-ПП}} = Q\left(\sqrt{\frac{E}{2N_0}}\right). \quad (10.9)$$

2. Частотна телеграфія з ортогональними сигналами.

Два сигнали є радіоімпульсами однакової форми з різними несучими частотами, так що сигнали взаємно ортогональні (рис. 17.7.б). Енергія різницевого сигналу дорівнює $2E$, а середня ймовірність помилки:

$$P_{\text{ОШ}}^{\text{ЧТ}} = Q\left(\sqrt{\frac{E}{N_0}}\right) \quad (10.10)$$

Підвищення потенційної стійкості до перешкод при переході від АТ-ПП до частотної телеграфії представляється природним, так як у другому випадку вдвічі зростає середня потужність передавача. Середня можливість помилки може бути знижена без збільшення потужності передавача, якщо перейти до взаємно обернених сигналів.

3. Фазова телеграфія з маніпуляцією фази на 180° .

У разі фазової телеграфії із взаємно оберненими сигналами (рис. 17.7.в) енергія різницевого сигналу становить $4E$, середня ймовірність помилки дорівнює:

$$P_{\text{ОШ}}^{\text{ФТ}} = Q\left(\sqrt{\frac{2E}{N_0}}\right) \quad (10.11)$$

і подальше підвищення потенційної завадостійкості за рахунок вибору сигналів при заданій енергії, очевидно, неможливе.

Зауважимо, що якщо використовуються три сигнали однакової енергії, то для досягнення максимальної стійкості до перешкод вони повинні мати взаємний фазовий зсув 120° , тобто відповідні сигналам точки повинні розташовуватися на колі радіуса $E^{1/2}$ у вершинах рівностороннього трикутника (рис. 17.7.г). Якщо сигналів чотири, то оптимальним є їхнє розміщення у вершинах правильного тетраедра, вписаного у сферу радіуса $E^{1/2}$.

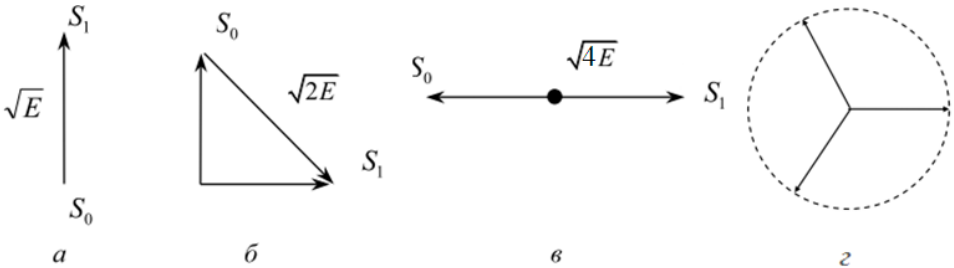


Рисунок 10.1 – До перешкодостійкості прийому двох і трьох сигналів

У загальному випадку оптимальний вибір системи з n сигналів відповідає їх розташуванню у вершинах правильного $(n - 1)$ -мірного симплексу, вписаного в $(n - 1)$ -мірну сферу (відрізок, трикутник і тетраедр є одновимірним, двовимірним та тривимірним симплексами).

10.3 Потенційна перешкодостійкість некогерентного прийому

Визначимо потенційну стійкість до перешкод некогерентного прийому на прикладі системи з пасивною паузою при рівних апіорних ймовірностях посилок $s_1(t) = A \cos(\omega t + \varphi)$, $s_0(t) = 0$, $p_1 = p_0 = 0,5$.

Середня ймовірність помилки дорівнює

$$P_{ном} = 0,5 \cdot p_{01} + 0,5 \cdot p_{10} = 0,5 \cdot \int_0^{V_n} w_1(V | H_1) dV + 0,5 \cdot \int_{V_n}^{\infty} w_0(V | H_0) dV .$$

Тут $w_1(V|H_1)$ і $w_0(V|H_0)$ – умовні щільності розподілу ймовірності огинаючої кореляційного інтеграла за умови гіпотез про передачу сигналів $s_1(t)$ і $s_0(t)$ відповідно, V_n – поріг (рис. 10.2).

При гіпотезі H_0 значення огинаючої обумовлене лише шумом, тоді квадратурні складові є незалежними нормальними випадковими величинами з нульовими середніми та дисперсіями $N_0E/2$.

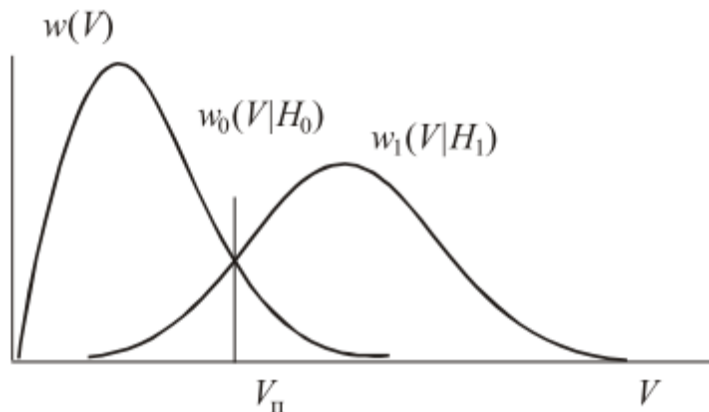


Рисунок 10.2 – Вибір порогу при некогерентному прийомі

Умовна щільність розподілу ймовірностей огинаючої має релеївський вигляд:

$$w_0(V | H_0) = \frac{2V}{EN_0} e^{-\frac{V^2}{EN_0}} . \quad (10.12)$$

Якщо коливання містить сигнал $s_1(t)$, то огинаюча має узагальнений релеївський розподіл (розподіл Релея – Райса):

$$w_1(V | H_1) = \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) . \quad (10.13)$$

Середня імовірність помилки рівна:

$$P_{ном} = \frac{1}{2} \int_0^{V_П} \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) dV + \frac{1}{2} \int_{V_П}^{\infty} \frac{2V}{EN_0} e^{-\frac{V^2}{EN_0}} dV , \quad (10.14)$$

тоді

$$P_{ном} = \frac{1}{2} \int_0^{V_П} \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) dV + \frac{1}{2} e^{-\frac{V_П^2}{EN_0}} . \quad (10.15)$$

Оптимальне значення порогу, при якому досягається потенційна стійкість до перешкод некогерентного прийому, є рішенням рівняння $dp_{\text{пом}}/dV_{\text{п}} = 0$. Точно вирішити отримане рівняння не вдається. Тому оптимальний поріг визначається наближеними виразами:

$$V_{\text{п опт}} = \begin{cases} \frac{E}{2} & \text{— при великих значеннях } \frac{\text{сигнал}}{\text{шум}}, \\ \sqrt{EN_0} & \text{— при малих значеннях } \frac{\text{сигнал}}{\text{шум}}. \end{cases} \quad (10.16)$$

Підставляючи (10.15) поріг $E/2$, отримаємо середню ймовірність помилки при великих відношеннях сигнал/шум (SNR):

$$p_{\text{пом}} = \frac{1}{2} \int_0^{E/2} \frac{2V}{EN_0} e^{-\frac{V^2+E^2}{EN_0}} \cdot I_0\left(\frac{2V}{N_0}\right) dV + \frac{1}{2} e^{-\frac{E}{4N_0}} \quad (10.16)$$

При великих SNR (≥ 10) першим доданком можна знехтувати, тоді:

$$p_{\text{пом}} = \frac{1}{2} e^{-\frac{E}{4N_0}}. \quad (10.17)$$

Аналогічно можна проаналізувати завадостійкість прийому двох ортогональних частотно-маніпульованих сигналів, де середня ймовірність помилки:

$$p_{\text{пом}} = \frac{1}{2} e^{-\frac{E}{2N_0}}. \quad (10.18)$$

Сигнали з фазовою маніпуляцією при випадковій початковій фазі кожної послілки, очевидно, застосовувати при некогерентному прийомі не можна. Однак при повільних змінах фази можна використовувати відносну фазову маніпуляцію, при якій початкова фаза наступної послілки збігається з початковою фазою попередньої послілки під час передачі символу «0» і відрізняється від неї на 180° – при передачі символу «1». При цьому середня ймовірність помилки:

$$P_{\text{ном}} = \frac{1}{2} e^{-\frac{E}{N_0}}. \quad (10.19)$$

10.4 Потенційна завадостійкість передачі неперервних повідомлень

Нехай передається неперервне повідомлення $a(t)$, причому $|a(t)| \leq 1$, а його потужність $P_a = \overline{a^2(t)} = 1$; у загальному випадку $a(t)$ можна подати у вигляді розкладання $a(t) = \sum_1^n \lambda_i \cdot \psi_i(t)$ по одиничних ортогональних

складових з коефіцієнтами λ_i . Тоді сигнал $S(t; a) = S(t; \lambda_1, \lambda_2, \dots, \lambda_n)$ отримуватиме в процесі модуляції збільшення відповідно до модуляційного вектора $D_i = \frac{dS(t)}{d\lambda_i}$, що визначає властивості модулятора.

З іншого боку, прийнятий сигнал $S^*(t)$ за рахунок перешкоди $\xi(t)$, що діє в каналі, буде відрізнятися від переданого $S(t)$. Як наслідок цього, коефіцієнти $\lambda_i = \lambda_i + \Delta \lambda_i$ також відрізнятимуться від переданих λ_i . В результаті матимемо:

$$a^*(t) = \sum \lambda_i^* \psi_i(t) = \sum (\lambda_i + \Delta \lambda_i^*) \psi_i(t) = a(t) + \sum \Delta \lambda_i^* \psi_i(t), \quad (10.20)$$

а для похибки передачі отримаємо $\varepsilon(t) = a^*(t) - a(t) = \sum \Delta \lambda_i^* \psi_i(t)$.

Оптимальний приймач Котельникова вимірює відстань $R = \overline{[S^*(t) - S(t)]^2}$.

Мінімальному значенню R відповідають збільшення $\Delta \lambda_i^*$, що визначаються з умови:

$$\frac{dR}{d\lambda_i} = -2 \overline{[S^*(t) - S(t)]} \frac{dS}{d\lambda_i} = 0, \text{ тобто } \Delta \lambda_i^* = \frac{\overline{\xi(t) \cdot D_i(t)}}{D_i^2(t)} \quad (10.21)$$

На основі (10.21) отримуємо вираз для середньої потужності шуму на виході приймача $P_\varepsilon = \overline{(\Delta \lambda_i^*)^2}$. Крім того, при слабких перешкодах можна вважати, що спектральну потужність шуму на виході можна уявити у вигляді:

$$G_\varepsilon(\omega) = \frac{G_\xi(\omega)}{D_i^2(t)}, \quad (10.22)$$

а P_ε – відповідно

$$P_{\varepsilon} = \int_0^{\omega_b} G_{\varepsilon}(\omega) d\omega . \quad (10.23)$$

Складемо відношення середніх потужностей повідомлення та шуму на виході приймача:

$$h_2^2 = \frac{P_a}{P_{\varepsilon}} = \frac{1}{\int_0^{\omega_b} G_{\varepsilon}(\omega) d\omega} . \quad (10.24)$$

Вирази (10.23) і (10.24) використовуються для порівняльної оцінки потенційної завадостійкості різних видів модуляції. В результаті аналізу отримані наступні вирази $G_{\varepsilon}(\omega)$ и h_2^2 .

Амплітудна модуляція $S(t) = A[1 + ma(t)]\cos \omega_0 t$:

$$G_{\varepsilon}(\omega) = \frac{2G_0}{m^2 A^2}; \quad h_2^2 = \frac{m^2 A^2}{2G_0 \omega_b} . \quad (10.25)$$

Фазова модуляція $S(t) = A\cos[\omega_0 t + \Delta\varphi_m a(t)]$:

$$G_{\varepsilon}(\omega) = \frac{2G_0}{\Delta\varphi_m^2 A^2}; \quad h_2^2 = \frac{\Delta\varphi_m^2 A^2}{2G_0 \omega_b} . \quad (10.26)$$

Частотна модуляція

$$S(t) = A\cos\left[\omega_0 t + \Delta\omega_m \int_0^t a(t) dt\right]; \quad a(t) = \cos \Omega t; \quad \beta_m = \frac{\Delta\omega_m}{\nu}$$

$$G_{\varepsilon}(\omega) = \frac{2G_0}{\Delta\omega_m^2 A^2}; \quad h_2^2 = \frac{\Delta\beta_m^2 A^2}{2G_0 \omega_b} . \quad (10.27)$$

З (10.25-10.27) випливає, що при АМ ($m = 1$) перешкодостійкість можна збільшити тільки за рахунок збільшення потужності сигналу; при ФМ та ЧС перешкодостійкість можна збільшити також шляхом збільшення

індексів модуляції $\Delta\varphi_m$ і β_ω , що досягається за рахунок розширення спектра сигналу.

Всі ширококугові системи модуляції забезпечують високу перешкодостійкість h_2^2 за умови, що відношення сигнал-перешкода на вході h_1^2 більше деякого порогового значення (рис. 18.5) При $h_1^2 < h_{1^2\text{пор}}$ ширококугові системи втрачають свої переваги порогу завадостійкості (різко знижується завадостійкість); явище, що виникає при цьому, називають явищем **порогу завадостійкості**.

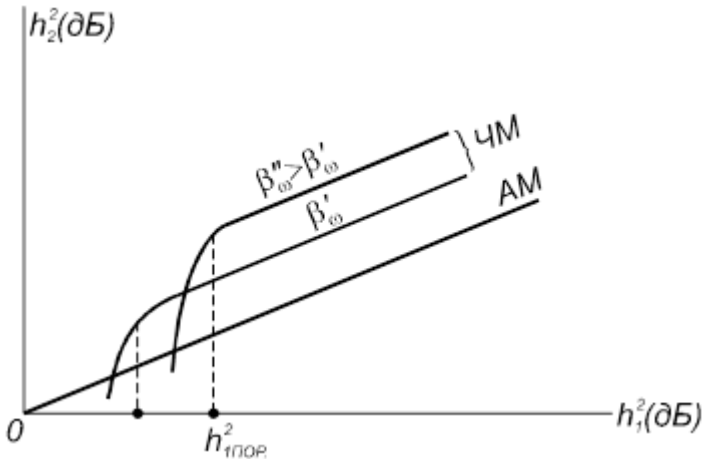


Рисунок 10.3 – Графіки $G_s(\omega)$ для АМ, ФМ и ЧМ в відповідності з 10.25-10.27

Тема 11. Оптимальна фільтрація випадкового сигналу.

Оптимальна фільтрація випадкового сигналу. Фільтр Колмогорова-Вінера. Фільтр Калмана. Узгоджена фільтрація.

11.1 Основні підходи фільтрації

Більш складною і загальною, ніж завдання оцінювання постійного параметра, є завдання оцінювання повідомлення, що змінюється (первинного сигналу) на основі спостережуваної реалізації. Таке оцінювання прийнято називати фільтрацією.

Повідомлення розглядається як реалізація випадкового процесу, множина різноманітних повідомлень – як ансамбль реалізацій з деяким імовірнісним розподілом. Повідомлення (первинний сигнал) модулює несуче коливання, тому сигнал на виході каналу зв'язку також випадковий. Таким чином, ставиться завдання по випадковому коливанню, що спостерігається, оцінити інше випадкове коливання (первинний сигнал, або закон модуляції), пов'язане зі спостережуваним у загальному випадку нелінійною залежністю (завдання нелінійної фільтрації, або демодуляції). Це завдання називають **фільтрацією**, і воно використовується для того, щоб відновити або передбачити значення змінного сигналу на основі обмеженої кількості спостережень. Це завдання може бути дуже складним.

Найпростіший випадок – лінійна фільтрація – це процес обробки сигналів, що використовується для видалення небажаних частотних компонентів або для виділення певного частотного діапазону. Фільтри, які використовуються для цих цілей, класифікуються за частотою, наприклад:

- Низькочастотні фільтри (LPF): Пропускають низькі частоти та відфільтровують високі.
- Високочастотні фільтри (HPF): Пропускають високі частоти та відфільтровують низькі.
- Смугові фільтри (BPF): Пропускають тільки певний діапазон частот.
- Смуго-загороджувальні фільтри (BSF): Блокують певний діапазон частот.

Лінійні фільтри мають дві основні властивості:

Лінійність: Відгук на лінійну комбінацію вхідних сигналів дорівнює лінійній комбінації відгуків на ці сигнали.

Інваріантність від часу: Характеристики фільтра не змінюються з часом.

Основні параметри фільтрів:

Амплітудно-частотна характеристика (АЧХ) – описує, як фільтр впливає на амплітуди сигналів різних частот.

Фазо-частотна характеристика (ФЧХ) – показує зміну фази сигналів під дією фільтра.

Перехідна характеристика – описує, як фільтр реагує на одиничний імпульс.

Фільтри можна описувати через їхню передавальну функцію, яка визначає, як фільтр змінює амплітуду та фазу вхідного сигналу на кожній частоті.

Цифрові фільтри є класом лінійних дискретних систем, що здійснюють обробку сигналу в часовій області і змінюють спектральний склад сигналу. Зміни спектрального складу полягають у подавленні небажаних з точки зору прикладного завдання спектральних компонент вхідного сигналу. Наприклад, в сучасних системах звукового оточення для подавлення шумів використовують процедуру цифрової попередньої фільтрації або в акустичних системах класу hi-fi вводять вирівнюючий фільтр для компенсації нелінійності частотної характеристики гучномовців. Цифрові фільтри можуть бути реалізовані різними шляхами – комп'ютерною програмою, програмованим апаратним сигнальним процесором або інтегральною схемою.

Фільтрація може виконуватися різними методами залежно від вимог до точності, обчислювальних ресурсів і природи сигналу. Основними підходами є:

1. Класична фільтрація, що передбачає використання фільтрів, які мають фіксовані параметри. До класичних методів відносять: фільтр Баттерворта; фільтр Чебишева; фільтр Бесселя.

2. Оптимальна фільтрація, зокрема фільтрація Вінера і фільтрація Калмана, спрямована на мінімізацію середньоквадратичної помилки між оцінюваним і дійсним значеннями сигналу.

У цьому підрозділі розглядається найпростіший випадок оптимальної лінійної фільтрації.

Нехай сигнал на вході лінійного фільтра є сумою переданого сигналу $S(t)$ і перешкоди $\zeta(t)$: $x(t) = S(t) + \zeta(t)$.

Потрібно знайти таку передатну характеристику $K(\omega)$, яка мінімізує середній квадрат похибки:

$$\overline{\varepsilon^2(t)} = \overline{[\hat{S}(t) - S(t)]^2}, \quad (11.1)$$

де $\hat{S}(t)$ – оцінка сигналу на виході фільтра. Тут, вважаємо, що час запізнення сигналу $\hat{S}(t)$ у фільтрі $\tau_z = 0$, а середнє значення береться за ансамблями сигналів $S(t)$ та перешкод $\zeta(t)$. Вважаємо також, що $S(t)$ та $\zeta(t)$ стаціонарні взаємно-некорельовані процеси з відомими спектрами щільності потужності $G_s(\omega)$ і $G_\zeta(\omega)$. Для відгуку сигналу на виході фільтра запишемо $y(t) = L[S(t) + \zeta(t)]$, де $L[-]$ – оператор лінійного перетворення.

Середньоквадратичну похибку можна представити двома складовими:

$$\overline{\varepsilon^2(t)} = \overline{\varepsilon_S^2(t)} + \overline{\varepsilon_\xi^2(t)}, \quad (11.2)$$

де $\overline{\varepsilon^2(t)} = \overline{\{L[S(t) - S_x(t - \tau_z)]\}^2}$ – складова, обумовлена спотвореннями сигналу у фільтрі (вважаємо ФЧХ лінійною, тобто $\varphi(\omega) = -\omega\tau_z$); $\overline{\varepsilon^2(t)} = \overline{\{L[\xi(t)]\}^2}$ – складова, обумовлена дією шуму.

Зважаючи на те, що

$$\varepsilon_S(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(\omega) \left[1 - \dot{K}(\omega) e^{-j\omega\tau_z} \right] \cdot e^{j\omega} d\omega,$$

для спектральної густини потужності $G_{\varepsilon_S}(\omega)$ маємо

$$G_{\varepsilon_S}(\omega) = G_S(\omega) [1 - K(\omega)]^2.$$

У свою чергу, спектральна щільність потужності складової шуму дорівнює $G_{\varepsilon_\xi}(\omega) = G_\xi(\omega) \cdot K^2(\omega)$ а для $G_\varepsilon(\omega)$ сумарної похибки отримаємо:

$$G_\varepsilon(\omega) = G_{\varepsilon_S}(\omega) + G_{\varepsilon_\xi}(\omega) = G_S(\omega) [1 - K(\omega)]^2 + G_\xi(\omega) K^2(\omega). \quad (11.3)$$

Знайдемо тепер вираз передатної функції фільтра $K(\omega)$ – при якому сумарна похибка стає мінімальною:

$$\frac{dG_\varepsilon(\omega)}{dK(\omega)} = -2G_S(\omega) \cdot [1 - K(\omega)] + 2G_\xi(\omega) \cdot K(\omega) = 0,$$

Звідки

$$K_{\text{оп}}(\omega) = \frac{G_S(\omega)}{G_S(\omega) + G_\xi(\omega)}, \quad (11.4)$$

або

$$\dot{K}_{\text{оп}}(\omega) = \frac{G_S(\omega)}{G_S(\omega) + G_\xi(\omega)} e^{-j\omega\tau_z}. \quad (11.5)$$

Фільтр (11.5), що забезпечує мінімальну середньоквадратичну похибку випадкового сигналу на тлі білого шуму називається оптимальним фільтром **Колмогорова-Вінера**. Цей фільтр оцінює сигнал, який зазнає адитивного білого або кольорового шуму, з мінімізацією середньоквадратичної похибки між відновленим і початковим сигналом. Завдяки своїй здатності адаптуватися до умов із постійними спектральними характеристиками, фільтр Вінера часто використовується для обробки зображень, відновлення спотворених сигналів у зв'язку та видалення адитивного шуму. Слід

значити, що фільтр Колмогорова-Вінера з коефіцієнтом передачі (11.5) фізично не реалізуємий.

Конструктивні результати у вирішенні задачі синтезу фільтра, оптимального за критерієм мінімуму СКП отримав Рудольф Еміль Калман, який як апріорні дані про повідомлення використовував рекурентну модель: стан – спостереження. Передбачається, що повідомлення $a(t)$ породжується лінійним стохастичним диференціальним рівнянням стану

$$\frac{da}{dt} = -\alpha \cdot a(t) + n_1(t), \quad (11.6)$$

де $n_1(t)$ – нормальний білий шум (що породжує процес). Фізично це означає, що повідомлення $a(t)$ моделюється як результат проходження білого шуму через інтегруючий ланцюг з постійною часу $\alpha = 1/RC$.

Розглянемо випадок лінійної фільтрації, коли процес, що спостерігається на вході фільтра заданий рівнянням спостереження

$$x(t) = S(t, a(t)) + \zeta(t) = a(t)f(t) + \zeta(t),$$

де повідомлення $a(t)$ – задано рівнянням (18.12); $S(t, a(t))$ – сигнал; $f(t)$ – носій; $\zeta(t)$ – білий гаусівський шум.

Беручи до уваги основну умову лінійної оптимальної фільтрації, яка полягає в тому, щоб помилка повинна бути не корельована з вхідним сигналом $x(t)$, рекурентне співвідношення має вигляд:

$$\frac{da}{dt} = -\alpha \cdot \hat{a}(t) + \frac{2k}{G_0} f(t)[x(t) - \hat{a}(t) \cdot f(t)] \quad (11.7)$$

$$\frac{dk}{dt} = \frac{N_1}{2} - 2\alpha k - \frac{2}{N_0} f^2 k^2 \quad (11.8)$$

Рівняння (11.7) та (11.8) прийнято називати рівняннями фільтра Калмана. Рівняння (11.7) визначає алгоритм формування оцінки, отже, і структурну схему фільтра, а (11.8) – помилку фільтрації $k = \overline{\varepsilon^2(t)}$.

Фільтр Калмана – це алгоритм, що забезпечує оптимальну оцінку змінних стану системи, яка підлягає дії випадкових шумів. Його основна задача – знаходження оцінки стану динамічної системи в умовах неповної та зашумленої інформації. Фільтр Калмана особливо корисний у системах зв'язку, навігації, відстеження об'єктів, і зменшення шуму в сигналах. Зокрема, він ефективний для оцінки каналу, що дозволяє враховувати зміни середовища передачі та компенсувати їх в процесі прийому даних.

Структурна схема, що моделює рівняння (11.7), наведена на малюнку 11.1. Цей фільтр є динамічною системою зі змінним коефіцієнтом підсилення,

величина якого залежить від точності поточних оцінок (дисперсії похибки фільтрації) і рівня шумів спостереження. Інакше висловлюючись, фільтр Калмана можна розглядати як лінійну слідкуючу систему, ефективна смуга пропускання якої, змінюється у процесі роботи.

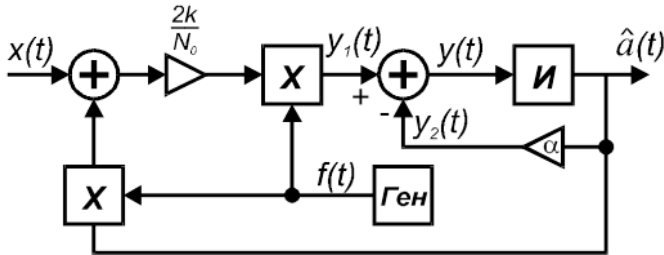


Рисунок 11.1 – Структурна схема, що моделює рівняння 11.7

11.2 Узгоджена фільтрація

Фактично кореляційний приймач є активним фільтром і виконує операцію скалярного добутку

$$(\bar{x}\bar{s}_i) = \int_0^T x(t)s_i(t)dt . \quad (11.9)$$

Цю операцію можна реалізувати за допомогою пасивного лінійного фільтра з постійними параметрами. Якщо на вхід фільтра подати сигнал, то напруга на виході фільтра

$$y(t) = \int_0^T x(t-\tau)g(\tau)d\tau , \quad (11.10)$$

де $g(t)$ – імпульсна реакція фільтра, що є дзеркальним відображенням $s(t)$ щодо $t = T$.

Виберемо $g(t)$ такий, щоб у момент $t = T$ отримати на виході значення $y(T)$, що збігається зі скалярним добутком (11.9).

Легко бачити, що це буде виконано, якщо $g(t) = s_i(T-t)$. Дійсно, при цьому

$$y(T) = \int_0^T x(T-\tau)s_i(T-\tau)d\tau = \int_0^T x(t)s_i(t)dt = (\bar{x}\bar{s}) . \quad (11.11)$$

Такий фільтр називають узгодженим (УФ) із сигналом $s_i(t)$. Інакше кажучи, фільтр є узгодженим із сигналом $s(t)$, якщо його імпульсна реакція

$$g(t) = a \cdot s(T - t), \quad (11.12)$$

де a – постійна.

Передатна функція УФ з імпульсною реакцією (11.12) визначається перетворенням Фур'є

$$\dot{K}_{-\Phi}(\omega) = \int_{-\infty}^{\infty} g(t)e^{-j\omega t} dt = a\dot{s}^*(\omega)e^{-j\omega T}, \quad (11.13)$$

де $\dot{S}^*(\omega) = \dot{S}(-\omega)$ – Функція комплексного пов'язана зі спектральною щільністю сигналу $s(t)$.

Отже, з точністю до коефіцієнта a АЧХ узгодженого фільтра визначається амплітудним спектром сигналу $s(t)$. Сенс узгодження проявляється в тому, що УФ добре пропускає ті частоти, які дають великий внесок у енергію сигналу. ФЧХ УФ (не враховуючи доданку $-\omega T$) зворотна за знаком ФЧХ сигналу $s(t)$. Завдяки цьому при $t = T$ всі складові спектру сигналу складаються у фазі і дають максимальний відгук.

Зазначимо одну важливу властивість УФ, яка іноді розглядається як його визначення. Будемо подавати суму детермінованого сигналу та білого шуму на вхід різних лінійних кіл з постійними параметрами та вимірювати в момент $t = T$ відношення пікової потужності сигнальної складової до середньої потужності шуму на виході ланцюга. Виявляється, що це відношення максимальне для УФ і рівне:

$$h_{2\max}^2 = \frac{2E_s}{G_0}, \quad (11.14)$$

де E_s – енергія сигналу; G_0 – спектральна густина білого шуму.

Отже, УФ є єдиним лінійним фільтром, який забезпечує отримання максимального можливого відношення сигналу до перешкоди на виході.

Цікаво порівняти із відношенням h_1^2 середніх потужностей сигналу P_s і перешкоди P_ξ на вході фільтра:

$$h_1^2 = \frac{P_s}{P_\xi} = \frac{2T}{2T} \frac{P_s}{G_0 F_s} = \frac{2E_s / G_0}{2F_s T} = \frac{h_{2\max}^2}{2F_s T}$$

Звідки

$$h_{2\max}^2 = 2F_s T h_1^2 = n h_1^2, \quad (11.15)$$

де $n = 2F_s T$ – база сигналу.

Отже, покращення відношення сигналу до перешкоди, що дає УФ, тим більше, що більша база сигналу n , тобто що складніша форма сигналу.

Узгоджений фільтр для сигналу довільної форми може бути реалізований (наближено) на основі лінії затримки з відведеннями (рис. 11.2).

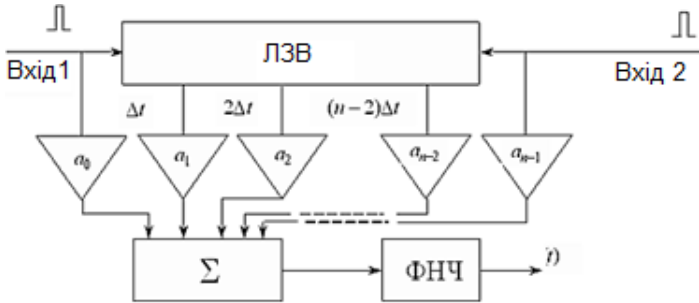


Рисунок 11.2 – Узгоджений фільтр на основі лінії затримки з відводами

При подачі на вхід 1 лінії затримки з відводами (ЛЗВ) короткого імпульсу, на вхід ФНЧ надходять (з інтервалом Δt , обумовленим конструкцією лінії затримки) такі ж імпульси з амплітудами, що визначаються коефіцієнтами посилення $a_0, a_1, a_2, \dots, a_{n-1}$. Тоді на виході ФНЧ формується, зокрема якщо ФНЧ є ідеальним, відгук, що являє собою кінцеву суму ряду Котельникова, що апроксимує сигнал $s(t)$ необхідного вигляду.

Неважко бачити, що якщо короткий імпульс подати на вхід 2, відгук буде дзеркальною копією сигналу $s[(n-1)\Delta t - t]$. Коефіцієнти $a_0, a_1, a_2, \dots, a_{n-1}$ являють собою відліки сигналу $s(t)$ з кроком, що визначається верхньою частотою F_c спектра сигналу.

Слід мати на увазі, що такий спосіб реалізації узгодженого фільтра є хоч і універсальним, але наперед наближеним, так як будь-який сигнал кінцевої тривалості має нескінченну спектральну щільність, а ідеальний ФНЧ нереалізований. Проте такий фільтр застосовується на практиці: наприклад, для узгодженої фільтрації сигналів з лінійною частотною модуляцією використовують як лінії затримки з відведеннями інтегральні пристрої на поверхневих акустичних хвилях (ПАХ).

Форма сигналу на виході такого фільтра відрізняється від форми вхідного сигналу. Але призначення узгодженого фільтра полягає у обчисленні кореляційного інтеграла для найбільш надійного ухвалення рішення про наявність або відсутність сигналу на вході приймача. Іншими словами, узгоджений фільтр повинен забезпечувати максимальне відношення сигнал/шум в момент часу t_0 та досягати високої чутливості прийому в системах із низьким рівнем сигналу.

Тема 12. Основи адаптивного придушення перешкод

Адаптивні пристрої придушення перешкод. Придушення стаціонарних перешкод. Адаптивний режсекторний фільтр. Адаптивний високочастотний фільтр.

Мета даного розділу – показати можливість підвищення відношення сигнал-перешкода, а також інші переваги методів придушення перешкод при використанні адаптивних фільтрів у порівнянні зі звичайними методами фільтрації перешкод.

12.1 Адаптивні пристрої придушення перешкод

Основна схема придушення перешкод показана на рисунку 12.1. Сигнал передається каналом зв'язку на приймальний пристрій, який приймає суміш сигналу і некорельованої з ним перешкоди n_0 . Суміш сигналу та перешкоди $s + n_0$ є вхідним сигналом пристрою придушення. Інший приймальний пристрій приймає перешкоду n_1 , некорельовану з сигналом, але деяким невідомим чином корельовану з перешкодою n_0 . У ньому формується еталонний сигнал для пристрою придушення. В результаті фільтрації перешкоди n_1 формується сигнал y який приблизно являє собою копію n_0 . Цей сигнал віднімається з вхідного сигналу $s + n_0$, щоб сформувати вихідний сигнал системи $s + n_0 - y$.

У загальному випадку при відомих характеристиках каналів, за якими перешкода надходить на обидва приймальні пристрої, можна синтезувати фільтр з постійними параметрами, що перетворюють n_1 в $y = n_0$. Тоді вихідний сигнал фільтра можна відняти з вхідного сигналу, і на виході системи залишиться один сигнал.

Однак застосування фільтра з постійними параметрами не забезпечує гнучкості, оскільки вважають, що характеристики трактів передачі або невідомі або відомі тільки приблизно і можуть змінюватися. Більш того, навіть якщо фільтр з постійними параметрами забезпечує гнучкість, його параметри потрібно налаштовувати з точністю, яка важко реалізується, і найменша помилка може призвести до збільшення потужності перешкоди на виході системи.

У наведеній на рисунку 12.1 системі еталонний сигнал обробляється адаптивним фільтром, який автоматично перебудовує свою власну імпульсну характеристику по одному із середньоквадратичних алгоритмів (наприклад, за методом найменших квадратів), що функціонує за сигналом помилки, яка залежить, крім усього іншого, від вихідного сигналу. Таким чином, при правильному алгоритмі фільтр може працювати в змінних умовах і перебудовуватися для мінімізації сигналу помилки.

Можна припустити, що для синтезу фільтра або його адаптації з метою формування сигналу, що компенсує перешкоду, необхідні деякі апріорні

відомості про сигнал s або перешкоди n_0 і n_1 . Однак, як буде показано далі, цього не потрібно або майже не потрібно.

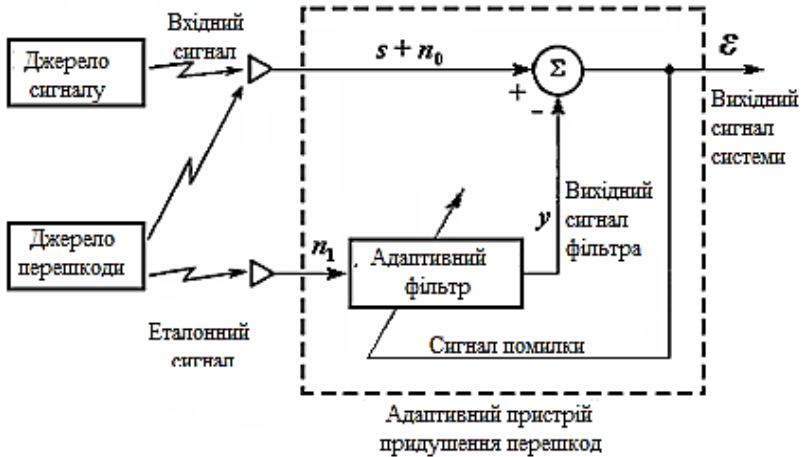


Рисунок 12.1 – Ілюстрація процесу адаптивного придушення перешкод

Нехай s , n_0 , n_1 і y – стаціонарні випадкові процеси з нульовими середніми значеннями, і s некорельований з n_0 і n_1 , а n_0 і n_1 – корельовані. Вихідний сигнал:

$$\varepsilon = s + n_0 - y. \quad (12.1)$$

Зведемо обидві частини рівності квадрат:

$$\varepsilon^2 = s^2 + (n_0 - y)^2 + 2s(n_0 - y). \quad (12.2)$$

Для обох частин (12.2) знайдемо математичне очікування і, оскільки s не корелюється з n_0 і y , отримаємо

$$E[\varepsilon^2] = E[s^2] + E[(n_0 - y)^2] + 2E[s(n_0 - y)] = E[s^2] + E[(n_0 - y)^2]. \quad (12.3)$$

Потужність сигналу $E[s^2]$ не змінюється при перебудові фільтра у процесі мінімізації $E[s^2]$. Відповідно до цього мінімальна потужність вихідного сигналу

$$E_{\min}[\varepsilon^2] = E[s^2] + E_{\min}[(n_0 - y)^2]. \quad (12.4)$$

Якщо фільтр побудований так, що $E[s^2]$ мінімальне, то мінімальне також і $E[(n_0 - y)^2]$. У цьому випадку вихідний сигнал фільтра є найкращою

середньоквадратичною оцінкою перешкоди n_0 . Більше того, при мінімальному значенні $E[(n_0 - y)^2]$ мінімальне значення має також і $E[(\varepsilon - s)^2]$, оскільки з (12.1)

$$(\varepsilon - s) = (n_0 - y). \quad (12.5)$$

Таким чином, перебудова або адаптація фільтра для мінімізації загальної потужності вихідного сигналу аналогічна тому, що при заданих структурі адаптивного фільтра і еталонному вхідному сигналі вихідний сигнал ε змінюється так, що він є найкращим в середньоквадратичному сенсі наближенням сигналу s .

У загальному випадку вихідний сигнал дорівнює сумі сигналу s і деякої перешкоди. Відповідно до (12.1) перешкода на виході дорівнює $n_0 - y$. Оскільки при мінімізації $E[\varepsilon^2]$ здійснюється мінімізація $E[(n_0 - y)^2]$, то мінімізація загальної потужності вихідного сигналу призводить до мінімізації потужності перешкоди на виході і, оскільки сигнал на виході залишається постійним, до максимізації вихідного відношення сигнал-перешкода (SNR).

З (12.3) та (12.4) видно, що найменша можлива потужність вихідного сигналу $E_{\min}[\varepsilon^2] = E[s^2]$. Якщо цього можна досягти, то $E[(n_0 - y)^2] = 0$. Отже, $y = n_0$ і $\varepsilon = s$, і мінімізація потужності вихідного сигналу призводить до того, що сигнал зовсім не спотворений перешкодою.

Крім того, якщо еталонний сигнал зовсім не корельований з вхідним сигналом, фільтр відключається і не збільшує перешкоду на виході. У цьому випадку вихідний сигнал фільтра некорельований з вхідним сигналом і потужність вихідного сигналу

$$E[\varepsilon^2] = E[(s + n_0)^2] + 2E[-y(s + n_0)] + E[y^2] = -E[(s + n_0)^2] + E[y^2]. \quad (12.6)$$

Для мінімізації потужності вихідного сигналу необхідно, щоб було мінімальним $E[y^2]$, що досягається за рівності всіх вагових коефіцієнтів нулю (в такому випадку $E[y^2] = 0$).

Ці міркування легко поширити на випадок, коли вхідний та еталонний сигнали містять, крім n_0 і n_1 складові адитивного випадкового шуму, некорельовані між собою і з s , n_0 і n_1 . Крім того, ці міркування справедливі для випадку, коли n_0 і n_1 – детерміновані, а не випадкові сигнали.

12.2 Придушення стаціонарних перешкод

Як зазначалося раніше, при придушенні перешкод фільтри з постійними параметрами в більшості випадків не застосовуються, тому що в загальному випадку автокореляційна та взаємкореляційна функція вхідного та еталонного сигналів невідомі і часто змінюються в часі. Тому необхідно спочатку навчати адаптивні фільтри за навчальними статистиками, а потім

здійснювати стеження за ними, у випадку повільної їх зміни. Однак для стаціонарних вхідних сигналів встановлений режим роботи фільтрів, що повільно адаптуються, наближається до режиму роботи вінерівських фільтрів, тому зручним математичним апаратом для аналізу статистичних завдань придушення перешкод є вінерівська теорія фільтрації.

На рисунку 12.2. представлена дещо докладніша схема системи, показаної на малюнку 12.1. Тут показаний один із способів отримання вхідного сигналу та перешкод на входах системи. Вхідний сигнал є сумою сигналу s_k , і двох перешкод $-n_k$ і m_{0k} , а еталонний сигнал – суму двох інших перешкод m_{1k} і n_k , що надійшов через тракт з імпульсною характеристикою h_k і передатною функцією $H(z)$. Обидві перешкоди $-n_k$ на вході системи та n_k , що пройшла через тракт з h_k , мають спільне джерело і є корельованими між собою та не корельованими із сигналом s_k . Передбачається також, що їх енергетичні спектри обмежені на всіх частотах. Перешкоди m_{0k} і m_{1k} некорельовані між собою, з s_k , n_k на вході системи та n_k , що пройшла через тракт, з h_k . Для подальшого аналізу вважатимемо, що всі тракти поширення перешкод еквівалентні лінійним фільтрам з постійними параметрами.

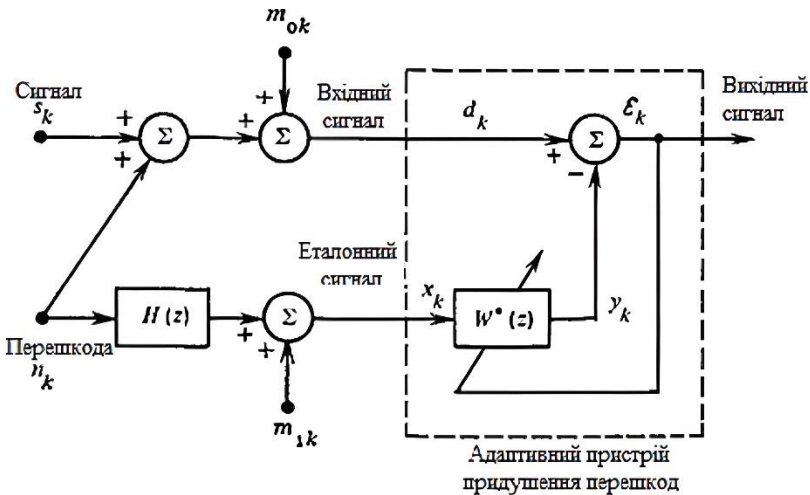


Рисунок 12.2 – Одноканальний адаптивний пристрій придушення перешкод з корельованою та некорельованою перешкодами на вході пристрою та на еталонному вході

Схема придушення на рис. 12.2. включає адаптивний фільтр, вхідний сигнал x_k (еталонний сигнал пристрою придушення) якого дорівнює сумі m_{1k} і n_k , що пройшли через тракт з h_k , а корисний відгук (вхідний сигнал пристрою придушення) $d_k = s_k + m_{0k} + n_k$. Сигналом помилки ϵ_k є вихідний сигнал пристрою придушення. Якщо вважати, що адаптивний процес завершено і

знайдено оптимальні в сенсі мінімуму СКВ вагові коефіцієнти (оптимальне рішення), то адаптивний фільтр еквівалентний вінерівському фільтру з передатною функцією

$$W^*(z) = \Phi_{xd}(z)/\Phi_{xx}(z), \quad (12.7)$$

де $\Phi_{xx}(z)$ – спектр вхідного сигналу фільтра; $\Phi_{xd}(z)$ – взаємний енергетичний спектр вхідного сигналу фільтра та корисного відгуку.

Оптимальна передатна функція адаптивного фільтра $W^*(z)$ є вінерівським рішенням (12.7), для якого тепер можна отримати наступне узагальнення. Спектр вхідного сигналу фільтра $\Phi_{xx}(z)$ можна виразити через спектри його двох некорельованих складових. Спектр перешкоди m_1 – $\Phi_{m_1m_1}(z)$, а спектр перешкоди n , що пройшла тракт із $H(z)$, – $\Phi_{nn}(z) |H(z)|^2$. Звідси спектр вхідного сигналу фільтра

$$\Phi_{xx}(z) = \Phi_{m_1m_1}(z) + \Phi_{nn}(z) |H(z)|^2. \quad (12.8)$$

Взаємний енергетичний спектр вхідного сигналу фільтра та корисного відгуку залежить тільки від корельованих складових вхідного та еталонного сигналів та має вигляд

$$\Phi_{xd}(z) = \Phi_{nn}(z) H(z^{-1}), \quad (12.9)$$

Підставляючи (12.9) в (12.7), отримуємо вінерівську передатну функцію

$$W^*(z) = \frac{\Phi_{nn}(z)H(z^{-1})}{\Phi_{m_1m_1}(z) + \Phi_{nn}(z)|H(z)|^2}. \quad (12.10)$$

Значимо, що $W^*(z)$ залежить від спектра сигналу $\Phi_{ss}(z)$ і спектра некорельованого шуму $\Phi_{m_0m_0}(z)$ на вході системи.

При нульовій адитивній перешкоді m_1 на еталонному вході маємо цікавий окремий випадок, коли $\Phi_{m_1m_1}(z) = 0$, і оптимальна передатна функція (12.10) набуває вигляду

$$W(z)^* = 1/H(z). \quad (12.11)$$

Інтуїтивно ясно, що цей результат є правильним, так як адаптивний фільтр призводить до обнулення n_k на виході пристрою придушення. Некорельована перешкода m_{0k} не пригнічується і разом із сигналом з'являється на виході системи.

Функціонування пристрою придушення перешкод з одним входом можна розглянути більш загальному вигляді з погляду співвідношення двох відношень щільності потужності сигналу до щільності потужності перешкод – на виході системи $\rho_{\text{вих}}(z)$ і її вході $\rho_{\text{вх}}(z)$. Вважаючи, що спектр сигналу невід'ємний на всіх частотах, після відповідних перетворень маємо

$$\frac{\rho_{\text{вих}}(z)}{\rho_{\text{вх}}(z)} = \frac{\Phi_{nn}(z) + \Phi_{m_0m_0}(z)}{\Phi_{\text{вих шум}}}. \quad (12.12)$$

З рис. 12.2 випливає, що енергетичний спектр перешкоди на виході пристрою придушення дорівнює сумі трьох складових, одна з яких відповідає прямому проходженню на вихід m_{0k} , інша – m_{1k} , проходження через ланку з передатною функцією $W^*(z)$, а третя – n_k , проходження через ланку з передатною функцією $1-H(z) W^*(z)$. Звідси енергетичний спектр перешкоди на виході

$$\Omega_{\text{вих шум}}(z) = \Phi_{m_0m_0}(z) + \Phi_{m_1m_1}(z) \cdot |W^*(z)|^2 + \Phi_{nn}(z) \cdot |1 - H(z) \cdot W^*(z)|^2 \quad (12.13)$$

Для зручності позначимо відношення спектрів корельованих та некорельованих перешкод на вході та еталонному вході відповідно через

$$A(z) \equiv \frac{\Phi_{m_0m_0}(z)}{\Phi_{nn}(z)} \quad (12.14)$$

і

$$B(z) \equiv \frac{\Phi_{m_1m_1}(z)}{\Phi_{nn}(z) |H(z)|^2} \quad (12.15)$$

У цих позначках передатну функцію (12.10) можна переписати у вигляді

$$W^*(z) = 1/H(z)[B(z) + 1] \quad (12.16)$$

Тоді енергетичний спектр перешкоди на виході

$$\Phi_{\text{вих шум}}(z) = \Phi_{m0m0}(z) + \frac{\Phi_{m1m1}(z)}{|H(z)|^2 \cdot |B(z)+1|^2} + \Phi_{mn}(z) \cdot \left| 1 - \frac{1}{B(z)+1} \right|^2 = \Phi_{mn}(z)A(z) + \Phi_{mn}(z) \frac{B(z)}{B(z)+1} \quad (12.17)$$

При цьому відношення (12.12)

$$\frac{\rho_{\text{вих}}(z)}{\rho_{\text{вх}}(z)} = \frac{\Phi_{mn}(z) \cdot [1 + A(z)]}{\Phi_{\text{вих шум}}(z)} = \frac{1 + A(z)}{A(z) + \frac{B(z)}{B(z)+1}} = \frac{[A(z)+1] \cdot [B(z)+1]}{A(z) + A(z) \cdot B(z) + B(z)} \quad (19.18)$$

Цей вираз є загальним уявленням ідеальної характеристики придушення перешкод у даному випадку. Він дозволяє оцінити очікуваний рівень придушення перешкоди при ідеальній системі пригнічення, що містить фільтр Віннера з двосторонньою імпульсною характеристикою. У такій системі сигнал проходить на вихід неспотвореним. На противагу цьому класичні схеми Вінера, Калмана та адаптивних фільтрів у процесі придушення перешкод вносять деякі спотворення в сигнал.

З (19.18) видно, що придушення перешкод обмежуються відношенням $A(z)$ і $B(z)$. Зокрема, при малому $A(z) - \frac{\rho_{\text{вих}}(z)}{\rho_{\text{вх}}(z)} \approx \frac{1+B(z)}{B(z)}$;

$$\text{при малому } B(z) - \frac{\rho_{\text{вих}}(z)}{\rho_{\text{вх}}(z)} \approx \frac{1+A(z)}{A(z)};$$

$$\text{при малих } A(z) \text{ и } B(z) - \frac{\rho_{\text{вих}}(z)}{\rho_{\text{вх}}(z)} \approx \frac{1}{A(z)+B(z)}.$$

З цих співвідношень випливає, що повне придушення можливе, коли $A(z)$ та $B(z)$ дорівнюють нулю.

12.3 Адаптивний режекторний фільтр

У деяких випадках вхідний сигнал являє собою суму складової сигналу та адитивної синусоїдальної перешкоди. Зазвичай, для придушення такої перешкоди використовується режекторний фільтр. На рис. 12.3. наведено схему пристрою придушення одночастотної перешкоди з двома адаптивними ваговими коефіцієнтами. На вхід пристрою може подаватися сигнал будь-якого виду – випадковий, періодичний, детермінований, імпульсний, або будь-яка комбінація цих сигналів. На еталонному вході діє чистий

синусоїдальний сигнал $C \cdot \cos(\Omega_0 t + \varphi)$. Відліки вхідних сигналів беруться з інтервалами T секунд. Тут x_{1k} – відліки еталонного сигналу, а x_{2k} – відліки цього сигналу, зсунутого фазою на 90° .

Переваги такого режекторного фільтра полягають у тому, що він дозволяє регулювати смугу частот, формувати нулі і здійснювати адаптивне стеження за точним значенням частоти і фази перешкоди.

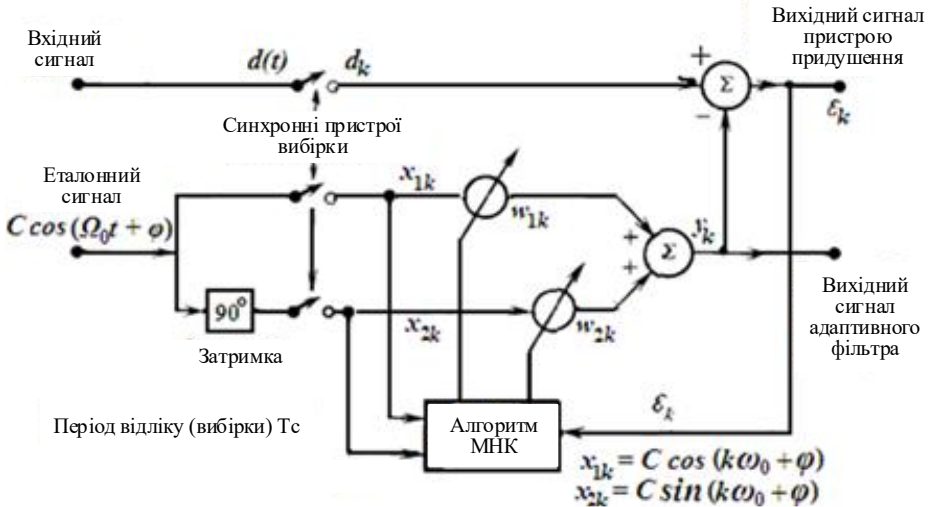


Рисунок 12.3 – Одночастотний адаптивний режекторний фільтр

Розглядаючи проходження сигналу від входу до виходу системи малюнку 12.3. можна знайти лінійну передатну функцію пристрою придушення перешкод. При замкнутій петлі зворотного зв'язку передавальна функція пристрою придушення перешкод

$$H(z) = \frac{z^2 - 2z \cdot \cos \omega_0 + 1}{z^2 - 2(1 - \mu C^2)z \cdot \cos \omega_0 + 1 - 2\mu C^2} \quad (12.19)$$

З рівності (12.19) випливає, що на частоті ω_0 еталонного сигналу пристрій придушення одночастотної перешкоди має властивості режекторного фільтра. Нулі передавальної функції розташовані на z -площині в точках

$$z = e^{\pm j\omega_0} \quad (19.20)$$

тобто точно на колі одиничного радіусу під кутами $\pm\omega_0$ радіан. Полюси передавальної функції розташовані у точках

$$z = (1 - \mu C^2) \cdot \cos \omega_0 \pm j[(1 - 2\mu C^2) - (1 - 2\mu C^2)^2 \cdot \cos^2 \omega_0]^{1/2}, \quad (12.21)$$

тобто всередині кола одиничного радіусу на відстані від початку координат $(1 - 2\mu C^2)^{1/2}$, приблизно рівному $1 - 2\mu C^2$, і під кутами

$$\pm \cos^{-1}[(1 - \mu C^2)(1 - 2\mu C^2)^{-1/2} \cos \omega_0]. \quad (12.22)$$

При повільній адаптації (тобто при невеликих значеннях μC^2) ці кути визначаються множителем, яким приблизно дорівнює 1.

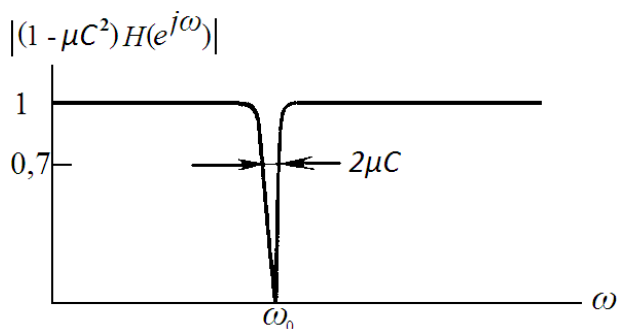


Рисунок 12.4 – Передатна функція адаптивного одночастотного пристрою придушення перешкод

Основний висновок у тому, що у практичних випадках кути полюсів і нулів майже рівні.

На рис. 12.4 показано розташування полюсів, нулів та точок половинної потужності передавальної функції. Оскільки нулі лежать на колі одиничного радіусу, глибина режекції в децибелах для передавальної функції на частоті ω_0 дорівнює нескінченності.

Форма провалу АЧХ визначається відстанню, яка приблизно дорівнює μC^2 . Довжина дуги кола одиничного радіусу, укладеного між точками половинної потужності, відповідає смузі режекції фільтра і дорівнює

$$BW = 2\mu C^2 \text{ рад} = \mu C^2 / \pi T \text{ Гц}. \quad (12.23)$$

Форма АЧХ у смузі режекції зазвичай визначається добротністю Q , що є відношенням центральної частоти до ширини смуги режекції:

$$Q = \frac{\omega_0}{2\mu C^2}. \quad (12.24)$$

Таким чином, пристрій придушення одночастотної перешкоди при

еталонному синусоїдальному сигналі еквівалентний стійкому режекторному фільтру. У загальному випадку глибина режекції адаптивного пристрою вище, оскільки в результаті адаптивного процесу навіть при повільній зміні частоти еталонного сигналу підтримується правильне співвідношення фаз для пригнічення.

12.4 Адаптивний високочастотний фільтр

Використання в адаптивному фільтрі вагового коефіцієнта для придушення низькочастотного дрейфу на вході системи є окремим випадком режекції сигналу на нульовій частоті. На рисунку 12.5. наведено схему тільки з одним ваговим коефіцієнтом, оскільки не потрібне підстроювання фази сигналу. Еталонним сигналом може бути постійний сигнал одиничної амплітуди.

Передаточну функцію тракту проходження сигналу від входу до виходу придушення пристрою можна отримати наступним чином. З рисунку 12.5 видно, що $y_k = w_k$, тому алгоритм найменших квадратів:

$$y_{k+1} = y_k + 2\mu\epsilon_k = y_k + 2\mu(d_k - y_k). \quad (12.25)$$

Знаходимо z -перетворення від (12.25) і отримуємо для режиму, що встановився.

$$Y(z) = \frac{2\mu \cdot D(z)}{z - (1 - 2\mu)} \quad (12.26)$$

Підставимо тепер в (12.26) $Y(z) = D(z) - E(z)$ і отримаємо

$$H(z) = \frac{E(z)}{D(z)} = \frac{z - 1}{z - (1 - 2\mu)}. \quad (19.27)$$

З рівності (12.27) випливає, що фільтр з ваговим коефіцієнтом зміщення являє собою високочастотний фільтр з нулем на колі одиничного радіусу на нульовій частоті та полюсом на дійсній осі на відстані 2μ вліво від нуля. Зазначимо, що це відповідає одночастотному режекторному фільтру (12.19) з $\omega_0 = 0$ і $C = 1$. Частота, на якій потужність сигналу, що режектується, зменшується вдвічі, дорівнює $\omega = 2\mu$ радіан.

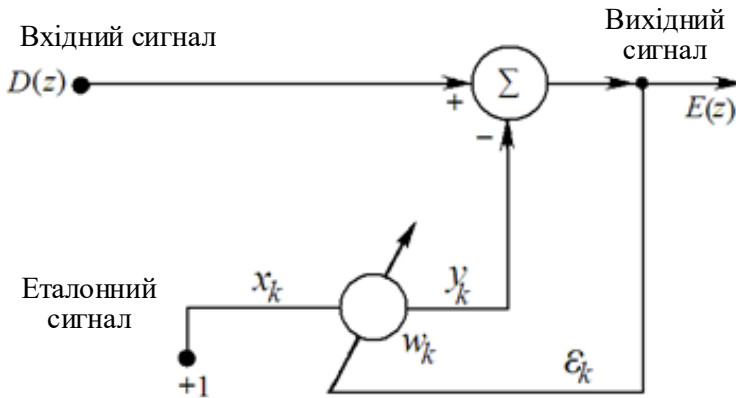


Рисунок 12.5 – Схема адаптивного високочастотного фільтра отриманого шляхом встановлення на еталонному вході постійного одиничного сигналу

В адаптивному високочастотному фільтрі виключається не тільки постійне зміщення, але і дрейф, що повільно змінюється у входньому сигналі. Більше того, хоча це тут і не показано, в цьому випадку одночасно здійснюється пригнічення випадкової низькочастотної перешкоди при постійному еталонному сигналі.

12.5 Пригнічення періодичної перешкоди за допомогою адаптивного прогнозуючого пристрою

У деяких випадках широкосмуговий сигнал спотворюється періодичною перешкодою, і при цьому немає зовнішнього еталонного сигналу, що не містить сам сигнал. Як приклад можна назвати відтворення мовлення чи музики за наявності шумів магнітофонної плівки чи диска, і навіть прийом сейсмічних сигналів у шумах автомобільного двигуна для силових установок.

На перший погляд може здатися, що для зменшення або виключення такої перешкоди не можна застосовувати адаптивне придушення. Однак при отриманні еталонного сигналу безпосередньо з входного через задану затримку Δ , як показано на рисунку 12.6, у багатьох випадках можна легко придушити періодичну перешкоду. Вибрана затримка повинна бути достатньою для того, щоб складові широкосмугового сигналу на еталонному вході не були корельовані з цим сигналом на вході. Через періодичність складові перешкоди залишаються корельованими.

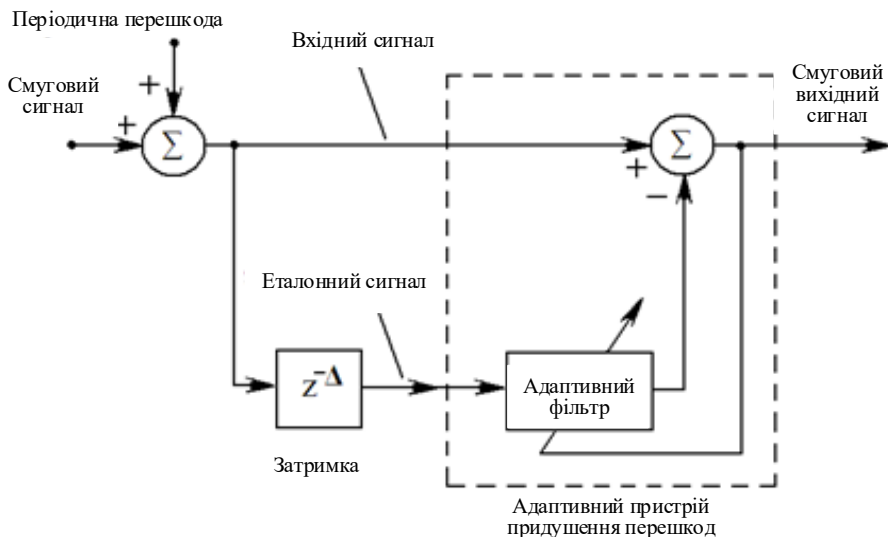


Рисунок 12.6 – Придушення періодичної перешкоди у системі без зовнішнього еталонного сигналу

Таким чином система містить адаптивний прогнозуючий пристрій. Передбачувана складова вхідного сигналу виключається і на виході системи залишається непередбачувана складова.

Адаптивне придушення перешкод є спосіб оптимальної фільтрації, який можна застосовувати завжди, коли є відповідний еталонний вхідний сигнал. Принциповими перевагами цього є його адаптивність, низький рівень перешкод на виході і малі спотворення сигналу, що вносяться. Адаптивність дозволяє обробляти вхідні сигнали з невідомими властивостями та деяких випадках нестационарні сигнали. Цей спосіб призводить до стійкої системи, яка автоматично відключається, якщо не досягається покращення відношення сигнал-шум. У загальному випадку рівень перешкод на виході і ступінь спотворення сигналу нижче ніж у оптимальних фільтрів, що використовуються в звичайних схемах.

Тема 13. Багатоканальний зв'язок та розподіл інформації

Багатоканальний зв'язок. Ущільнення каналів. Часове, частотне, кодове мультиплексування. Синхронізація у СПІ з багатостанційним доступом. Плезіохронна цифрова ієрархія. Синхронна цифрова ієрархія. Режим асинхронної передачі.

Найбільш дорогою частиною системи передачі є лінії зв'язку – кабельні, хвилеводні, світловодні, радіорелейні тощо. Природно постає завдання спільного використання цього устаткування різними користувачами (абонентами), тобто багатоканального зв'язку чи ущільнення. Тим самим підвищується ефективність використання ресурсів лінії.

Багатоканальний зв'язок можливий лише тоді, коли пропускна спроможність спільно використовуюваного обладнання більша за сумарну інформаційну продуктивність усіх джерел. При цьому ресурси лінії зв'язку мають бути певним чином розподілені між користувачами.

На рисунку 13.1. зображено спрощену схему передавальної та приймальної частин багатоканальної системи передачі.

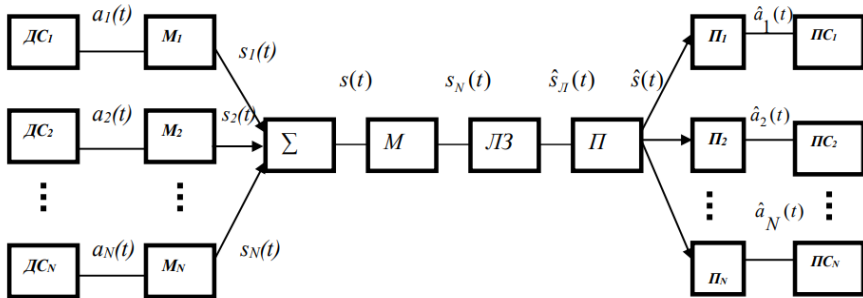


Рисунок 13.1 – Спрощена схема багатоканальної СПІ

На схемі позначено: $ДС_1, \dots, ДС_N$ – джерела інформації (сигналів); реалізації повідомлень кожного джерела ($ДС_i$) $a_1(t), a_2(t), \dots, a_n(t)$ за допомогою індивідуальних передавачів (модуляторів) M_1, M_2, \dots, M_N перетворюються у відповідні каналні сигнали $s_1(t), s_2(t), \dots, s_n(t)$. Сукупність каналних сигналів на виході підсумовуючого обладнання Σ утворює груповий сигнал $s(t)$. Нарешті, у груповому передавачі M сигнал $s(t)$ перетворюється в лінійний сигнал $s_n(t)$, який і надходить у напрямку систему (лінію зв'язку ЛЗ). Припустимо, що лінія пропускає сигнал без перекручувань і не вносить шумів. Тоді на прийомному кінці лінії зв'язку лінійний сигнал $s_n(t)$ за допомогою групового приймача Π може бути знову перетворений у груповий сигнал $\hat{s}(t)$. Канальними або індивідуальними приймачами $\Pi_1, \Pi_2, \dots, \Pi_N$ із групового сигналу $\hat{s}(t)$ виділяються відповідні каналні сигнали $\hat{s}_1(t) \dots \hat{s}_N(t)$ і потім перетворюються в призначені одержувачам повідомлення $\hat{a}(t)$.

Отже, операція ущільнення каналів – це звичайне підсумовування каналних сигналів:

$$S(t) = \sum_{j=1}^n s_j(t). \quad (13.1)$$

Генератор піднесучих видає n періодичних сигналів (гармонічних, послідовностей імпульсів тощо), що мають різну форму в різних каналах. Вид піднесучих i , отже, можливі способи їх модуляції визначаються методом ущільнення каналів. Модульовані піднесучі $S_1(t), \dots, S_n(t)$ називаються каналними сигналами.

Для поділу каналів, як правило, застосовуються також лінійні пристрої. Вважають, що в приймачі відбувається поділ каналних сигналів без виникнення міжканальних перешкод, якщо виконується наступна умова: за будь-яких значень повідомлень, що передаються в каналах, напруга на j -му виході лінійного пристрою поділу каналів дорівнює нулю тоді і тільки тоді, коли $S_j(t) = 0$.

Відомо, що для виконання цієї умови необхідно і достатньо, щоб система n функцій, що описують каналні сигнали, була лінійно незалежною за будь-яких значень повідомлень, що передаються.

Якщо виконується лише ця умова, пристрій розділення каналів може бути досить складним. Тому зазвичай висувають ще одну, додаткову вимогу: каналні сигнали повинні бути попарно ортогональні.

$$\int_{-\infty}^{+\infty} s_j(t) \cdot s_k(t) dt = \begin{cases} E_j, & j = k, \\ 0, & j \neq k. \end{cases} \quad (13.2)$$

де E_j – енергія j -го каналного сигналу.

Наведемо приклади типових СПІ, що використовують ортогональні каналні сигнали.

1. Канальні сигнали не перекриваються у часі (рис. 13.2). Умова ортогональності (13.2) тут виконується, оскільки за будь-якого t підінтегральний вираз дорівнює нулю при $j \neq k$. Це метод часового розподілу каналів.

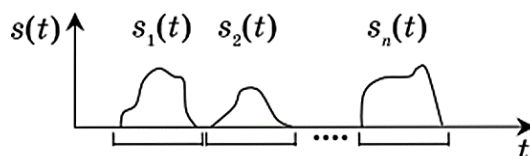


Рисунок 13.2 – Канальні сигнали.

Фактично лінія зв'язку почергово підключається на короткий час до кожного з джерел.

Спочатку передається сигнал (імпульс) 1-го каналу, потім другого каналу і т.д. до останнього каналу за номером N , після чого знову включається 1-й канал і процес періодично повторюється. Тому важливими елементами такої СПІ є мультиплексор (пристрій ущільнення каналів) і демультиплексор (пристрій поділу каналів), які працюють синхронно (рис. 13.3).

Тому невід'ємною складовою будь-якої СПІ з часовим розподілом є система синхронізації приймального пристрою.

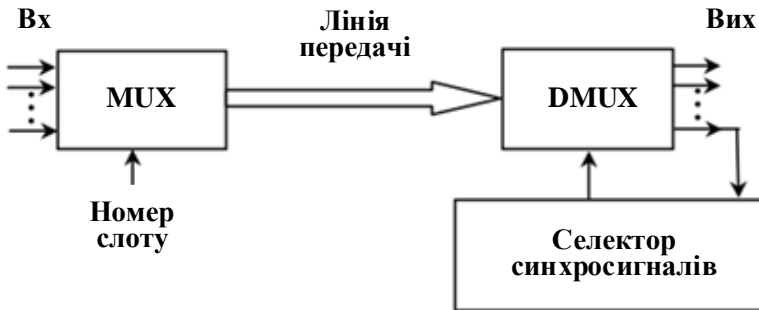


Рисунок 13.3 – Принципові елементи СПІ з TDMA

Важливо відзначити, що тривалість каналного інтервалу (слота) мала (інакше у користувачів зникне ефект «одночасності» передачі), тому за час одного підключення до лінії вдається передати лише частину повідомлення і неодноразово повторювати цей процес, всі перелічені елементи джерел сигналів повинні бути стиснуті у часі.

2. Спектри каналів не перекриваються на осі частот. Спектр групового сигналу (сума спектрів каналних сигналів) показано рисунку 13.4. Умова ортогональності (13.2) тут виконується для спектрів, отже, і функцій часу. Це метод частотного поділу каналів.

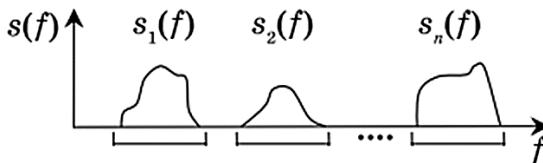


Рисунок 13.4 – Спектри каналних сигналів, що не перекриваються на осі частот

Кожній парі «джерело-одержувач» виділено свою смугу частот. Тому принциповими елементами такої СПІ є блок n генераторів гармонійних

коливань різних частот піднесучих в передавачі і набір смугових фільтрів для поділу каналних сигналів у приймачі.

3. Всі каналні сигнали передаються одночасно (перекриваються в часі) в одній загальній смузі частот (їхні спектри також перекриваються), але взаємна ортогональність сигналів забезпечена тим, що вони мають різну форму. Тоді пристрій поділу каналів – це набір n паралельно включених кореляційних приймачів або узгоджених фільтрів, на входи яких подається груповий сигнал. У кожному приймачі вагова функція $w(t)$ збігається формою з відповідним каналним сигналом, тому він реагує лише на сигнал даного каналу. Останню вимогу вдається виконати лише у цифровій СПП.

Такий метод називається багатоканальною передачею з поділом за формою сигналів. Як набір каналних сигналів можна використовувати будь-яку відому систему ортогональних функцій (функції Ерміта, Лагерра, Чебишева тощо), але ці функції мають досить складну форму. Тому для практичного використання запропонований більш простий варіант методу, коли як каналні сигнали застосовуються імпульси з внутрішньоімпульсної фазовою модуляцією, причому закон модуляції кожного каналного сигналу задається індивідуальною двійковою кодовою комбінацією.

Розглянемо детальніше кожен тип.

13.1 Частотний поділ каналів

Використовуються також терміни: частотне мультиплексування, багатостанційний доступ із частотним поділом каналів, Frequency Division Multiple Access (FDMA).

Повідомлення з виходів джерел (рис. 13.1) надходять на каналні модулятори, де відбувається модуляція гармонійних коливань з різними частотами, які називаються піднесучими. Частоти піднесучих коливань, повинні відрізнятися настільки, щоб спектри модульованих сигналів не накладалися один на одного, щоб уникнути взаємних перешкод. Після модуляції сигнали займають обмежені смуги частот, які можуть відрізнятися шириною від спектрів вихідних, або збігатися з ними. Важливо, щоб смуги частот, займані різними сигналами, не тільки не перекривалися, а й відстояли один від одного на ширину деякого захисного інтервалу, що полегшує їхнє поділ за допомогою реальних фільтрів, що мають кінцеву крутизну АЧХ в перехідній смузі.

Індивідуальні каналні (модульовані) сигнали підсумовуються і надходять на груповий передавач, де відбувається модуляція несучого коливання груповим сигналом, після чого модульований лінійний сигнал передається в лінію зв'язку. Груповий приймач здійснює демодуляцію лінійного сигналу, після чого кожен каналний приймач виділяє за допомогою смугового фільтра «свій» каналний сигнал, демодулює його та виділяє повідомлення.

Як видно, частотний поділ каналів заснований на розподілі одного з ресурсів – смуги пропускання групового каналу між окремими каналами.

Частотному поділу каналів властиві такі недоліки.

По-перше, через неідеальність смугових фільтрів необхідні захисні інтервали, які сумарно становлять близько 20% смуги пропускання групового каналу зв'язку. Наприклад, у багатоканальних телефонних системах передачі мовних сигналів встановлена смуга частот 3100 Гц (вважається, що для передачі мови задля забезпечення розбірливості із збереженням індивідуальних голосових ознак достатній діапазон від 300 до 3400 Гц), а ширина захисного інтервалу становить 900 Гц; таким чином, при об'єднанні N телефонних каналів загальна ширина смуги частот групового каналу становить $4N$ кГц.

По-друге, пред'являються дуже жорсткі вимоги до лінійності каналу (нелінійність призводить до появи кратних і комбінаційних складових, а оскільки спектри каналних сигналів мають ширину значно більшу за захисний інтервал, ці складові потрапляють у «чужі» канали і розділити їх шляхом фільтрації або яким-небудь іншим способом неможливо.

13.2 Часовий поділ каналів

Часове ущільнення (використовуються терміни: часове мультиплексування, багатостанційний доступ з часовим поділом каналів, Time Division Multiple Access (TDMA).) засноване на розподілі часового ресурсу групового каналу між різними індивідуальними каналами. Кожен користувач, який передає інформацію, отримує канал у своє користування багаторазово на короткий час. Очевидно, таким чином можна передавати лише відліки сигналу, взяті з кроком, що дорівнює періоду проходження тактових інтервалів. Таким чином, в його основі лежить використання теореми відліків, і можна передавати лише сигнали з фінітним спектром.

Для формування каналних сигналів використовуються різні види імпульсної модуляції (АІМ, ЧІМ, ШІМ). Груповий сигнал може передаватися безпосередньо по лінії або модулювати гармонійну несучу.

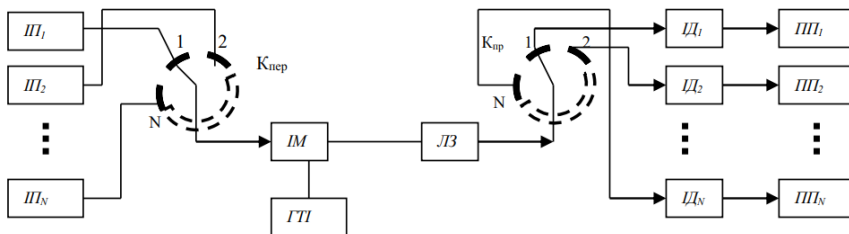


Рисунок 13.5 – Спрощена структурна схема СПЗ з ЧРК

Джерела первинних інформаційних повідомлень ПП з'єднані з комутатором передавача $K_{\text{пер}}$. Комутатор по черзі підключає джерела імпульсного модулятора ІМ, який модулює сигнал-переносник – періодичну послідовність імпульсів. В результаті цього виходить груповий сигнал, який надходить у лінію зв'язку ЛЗ, яка може включати гармонійний модулятор несучої, лінію зв'язку і загальний демодулятор.

Після загальної демодуляції груповий сигнал розділяється комутатором приймача $K_{\text{пр}}$ на каналні сигнали $s_1(t), \dots, s_N(t)$, які після демодуляції імпульсних демодуляторів ІД надходять до приймачів повідомлень ПП.

Слід зазначити, що в системах зв'язку з ЧРК щоб уникнути міжканальних перешкод необхідна синхронізація приймальної та передавальної станцій, тому в лінійний сигнал, крім каналних сигналів, додається періодична послідовність синхроімпульсів, які повинні досить сильно відрізнитися від каналних імпульсів, щоб їх можна було легко виділити.

Для зниження рівня взаємних перешкод доводиться вводити часові захисні інтервали між сусідніми імпульсами, що призводить або до зменшення числа каналів, або до укорочення каналних імпульсів, що веде до розширення необхідної смуги частот групового каналу.

Наприклад, при передачі мовних сигналів (ширина спектру 3100 Гц), мінімальна частота дискретизації повинна дорівнювати 6200 Гц. Однак насправді частоту дискретизації приймають рівною 8 кГц, що вимагає для передачі каналної імпульсної послідовності смуги частот близько 4 кГц.

В результаті загальна смуга частот групового каналу практично збігається зі смугою, яка потрібна при частотному мультиплексуванні. Якщо зважити на необхідність додаткової передачі синхроімпульсів, то порівняння виходить не на користь TDMA.

Проте роль часового поділу зростає у зв'язку з поширенням цифрових систем зв'язку.

13.3 Кодовий поділ каналів

Використовуються терміни: кодове мультиплексування, багатостанційний доступ з кодовим поділом каналів, Code Division Multiple Access (CDMA).

Цей метод застосовується лише у двійкових цифрових СПІ з ФМ. В одноканальній СПІ з ФМ, де для передачі кожного біта від джерела використовується радіоімпульс прямокутної форми, початкова фаза якого визначається значенням цього біта, наприклад, 0° при передачі одиниці і 180° при передачі нуля. Тривалість цього інформаційного імпульсу $\tau_i = 1/V$, де V – швидкість передачі, біт/с. Ширина спектру такого сигналу приблизно дорівнює $\Delta f \approx 2/\tau_i = 2V$. При малій швидкості передачі сигнал займає вузьку смугу частот.

Для організації кодового поділу каналів багатоканальної СПІ кожен каналний інформаційний імпульс (рис. 13.6.а) піддається додатковій внутрішньоімпульсній фазовій модуляції (рис. 13.6.б).

Для цього імпульс розбивається на частини V рівної довжини (на V більш коротких імпульсів-чипів).

Для даного каналу раз і назавжди задана V -розрядна двійкова послідовність, що визначає закон внутрішньоімпульсної фазової модуляції, причому елементи цієї послідовності зручно позначити як $+1$ і -1 замість традиційних 1 і 0 . Далі кожен елемент сигналу множиться на відповідний елемент послідовності, тобто залишається без змін, якщо множник дорівнює $+1$ та інвертується у разі -1 .

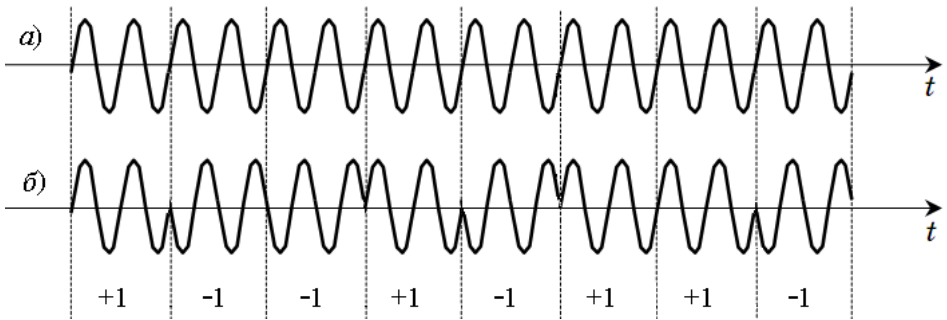


Рисунок 13.6 – Перетворення інформаційного імпульсу на сигнал із внутрішньоімпульсною ФМ:

а – до перетворення; б – після перетворення

У результаті ширина діапазону каналного сигналу збільшується в V раз порівняно з традиційним способом передачі без внутрішньоімпульсної модуляції, оскільки тепер сигнал фактично складається з більш коротких імпульсів. До речі, метод внутрішньоімпульсної ФМ є найпоширенішим методом штучного розширення спектра сигналу і часто застосовується в інших областях радіоелектроніки, наприклад радіолокації.

Всі каналні сигнали передаються одночасно в одній і тій же широкій смузі частот. Кожному каналному сигналу присвоюється своя двійкова послідовність, що визначає закон внутрішньоімпульсної ФМ, так щоб сигнали всіх каналів були взаємно ортогональні. Два каналних ФМ сигнали ортогональні, якщо в двійкових послідовностях, що визначають закони їх внутрішньоімпульсної ФМ, число елементів, що збігаються, дорівнює числу розбіжностей. Наприклад, якщо одному з каналів присвоєно послідовність, показана на рис. 13.6, то іншим каналам можна присвоїти, припустимо, такі послідовності:

$+1, -1, +1, -1, -1, +1, -1, +1, \dots$

+1, -1, -1, +1, +1, -1, -1, +1, ...

У системах CDMA операції виділення потрібного каналного сигналу із сумарного групового сигналу та демодуляції можна поєднати. Для цього можна використовувати кореляційний приймач (або узгоджений фільтр, що його замінює) і в якості опорного сигналу використовувати цілком імпульс з внутрішньоімпульсною ФМ. Таким чином, у системах CDMA пристрій поділу каналів та каналні демодулятори – це набір, що складається з n кореляційних приймачів.

Максимальна кількість ортогональних послідовностей дорівнює B , отже, саме така є кількість каналів, що працюють без взаємних перешкод.

Для практичного застосування підбирають послідовності, що мають кращі спектральні властивості (широкий спектр). Хороші результати дає використання відрізків псевдовипадкової двійкової послідовності (ПВП).

Оскільки в системах CDMA не передбачається взаємна синхронізація каналів, тобто модулюючі послідовності в різних каналах довільно зсунуті в часі, це може призвести до порушення їхньої ортогональності і спричинити міжканальні перешкоди.

Метод CDMA використовується поки лише у цивільних та військових системах радіозв'язку з рухомими об'єктами. Слід згадати стандарт IS-95, відповідно до якого побудовані стільникові системи CDMA і супутникова система зв'язку з рухомими об'єктами Globalstar ($B = 64$, ширина групового спектру і кожного з каналних сигналів дорівнює 1,25 МГц).

13.4 Синхронізація у СПІ з багатостанційним доступом

Розглянуті вище системи багатоканальному зв'язку із ортогональними сигналами вимагають для нормальної роботи тієї або іншої синхронізації:

- точного збігу спектра сигналу зі смугою пропускання каналу при частотному поділі;
- точного збігу часових інтервалів передачі сигналів окремих каналів при часовому поділі;
- точного визначення моментів початку і кінця тактового інтервалу в системах із поділом сигналів за формою активними фільтрами;
- точної установки моменту відліку в системі із узгодженими фільтрами.

Параметри цифрових СПІ з часовим розподілом каналів стандартизовані. Перший з найбільш важливих стандартів має назву **плезіохронної цифрової ієрархії** (ПЦІ) – Plesyochronous Digital Hierarchy (PDH) – і визначає спосіб мультиплексування стандартних цифрових телефонних сигналів для їх передачі мідним кабелем.

Так стандартний цифровий телефонний сигнал (DSO) виходить в результаті АЦП стандартного аналогового телефонного сигналу і характеризується швидкістю 64 кбіт/с.

Перший ступінь ущільнення PDH (апаратура ІКМ-30) розрахований на організацію 32 каналів методом часового побайтного мультиплексування, в результаті на виході формується цифровий потік Е1 зі швидкістю $64 \times 32 = 2048$ кбіт/с. З них 2 канали є службовими (передача сигналів циклової синхронізації, передача команд для управління комутуючими приладами на АТС та ін), а решта 30 – користувачькими.

Кожен з вищих рівнів (апаратура ІКМ-120, ІКМ-480, ІКМ-1920) розрахований на ущільнення чотирьох цифрових потоків нижчого рівня методом часового побітового мультиплексування.

Інший важливий стандарт носить назву **синхронної цифрової ієрархії** (СЦІ) - Synchronous Digital Hierarchy (SDH) – і визначає спосіб мультиплексування цифрових потоків для їх передачі оптоволоконним кабелем. Апаратура SDH не призначена для мультиплексування дрібних цифрових потоків, що надходять від користувачів (для цього використовується апаратура PDH, наприклад, ІКМ-30), вона застосовується для транспортування більших потоків між вузлами мережі. Синхронна цифрова ієрархія створювалася для роботи в синхронних мережах зв'язку, де є синхронізація всіх вузлів мережі з єдиного центру. Завдяки цьому швидкості передачі всіх однотипних потоків строго однакові. Передача цифрових сигналів на кожному рівні мультиплексування здійснюється блоками тривалістю 125 мкс.

Роботою передавальної частини багатоканальної СПІ у синхронному режимі передачі керує досить стабільний первинний генератор тактових імпульсів. Період поступання цих імпульсів дорівнює тривалості найкоротшого елемента групового сигналу. У цифровій СПІ з часовим розподілом каналів це один біт, у СПІ з кодовим розподілом це один з чіпів інформаційного біта. Наприклад, в апаратурі ІКМ-30 тактові імпульси мають частоту 2048 кГц.

Такий самий комплект генераторного обладнання використовується для управління роботою всіх пристроїв у приймальній частині СПІ. Система синхронізації приймального пристрою забезпечує умови, за яких тактові імпульси всіх видів у приймачі формуються одночасно з відповідними тактовими імпульсами в передавачі (точніше, з однією і тією ж затримкою, що дорівнює поточному часу поширення сигналу від передавача до приймача).

Щоб система синхронізації могла виконувати свої функції, крім сигналів, що несуть інформацію про повідомлення, необхідно ще передавати спеціальні синхросигнали різних видів.

У приймальній частині СПІ є селектори синхросигналів, які постійно порівнюють часове положення інформаційних і синхронізуючих сигналів, з одного боку, і тактових імпульсів від місцевого генераторного обладнання, з іншого боку, і в міру необхідності подають команди на корекцію останніх.

Спеціальні синхросигнали в момент початку чергового найдрібнішого елемента групового сигналу, наприклад біта, як правило, не передають. Достатньо передавати кадровий синхросигнал на початку кожного кадру (як сигнали точного часу достатньо передавати раз на годину, а не щомиті). Для передачі кадрового синхросигналу виділяється окремий канал, і в ньому в цифровий СПП передається спеціальна кодова комбінація тієї ж довжини, що і в будь-якому іншому каналі.

Наприклад, в апаратурі ІКМ-30 на початку кожного парного циклу на 2-8 бітових позиціях передається синхросигнал виду 0011011. Єдина відмінність циклового синхросигналу від комбінацій в інформаційних каналах полягає в тому, що він завжди один і той самий, а комбінації в інформаційних каналах змінюються. При збої синхронізації в селекторі він переходить у режим пошуку, тобто перевіряє поспіль всі слоти доти, доки не зустрінеться комбінація 0011011. Після цього переходить в нормальний режим, тобто режим стеження.

Природно, що вимоги до стійкості перешкод селектора синхросигналів істотно вищі, ніж вимоги до стійкості перешкод інших вирішальних пристроїв.

Вирішення проблеми синхронізації в системах радіозв'язку з багатостанційним доступом має низку особливостей. Якщо в багатоканальній СПП формування цифрового сигналу, кодування, модуляція та ущільнення каналів здійснюються в пункті передачі під управлінням загального генераторного обладнання, то в системах радіозв'язку передавач, пристрій ущільнення каналних сигналів (базова станція в системах стільникового зв'язку, ретранслятор в системах супутникового зв'язку тощо) та приймач перебувають у різних пунктах. Тут доводиться налагоджувати спільну синхронну роботу не двох, а трьох об'єктів за умов, коли відстані між ними змінюються у часі (отже, змінюються і затримки сигналу).

У таких системах, як правило, здійснюється двостороння передача інформації, тобто на обох ділянках є два рівноцінні канали: у прямому та зворотному напрямках. Тут провідна роль належить ретранслятору, тому в ньому знаходиться генераторне обладнання, що задає часовий графік роботи всієї системи, та формувачі синхросигналів. Приймачі всіх абонентів синхронізують своє генераторне обладнання за сигналами, що передаються ретранслятором, і це обладнання керує роботою свого передавача.

Залишається лише одне питання: з яким випередженням повинен працювати передавач, щоб його сигнал досяг ретранслятора на той момент, коли почнеться потрібний часовий слот? Для оцінки поточного часу запізнення по дорозі від станції до ретранслятора кожна станція періодично приймає свій сигнал, що повернувся від ретранслятора, і вимірює його запізнення щодо переданого сигналу.

Режим асинхронної передачі.

В ряді випадків здійснити точну синхронізацію важко. Із подібними ситуаціями доводиться зустрічатися, наприклад, при організації оперативного зв'язку великої кількості рухливих об'єктів (літаків, автомобілів, людей). Така задача виникає, наприклад, при організації оперативного зв'язку із використанням штучних супутників Землі в якості ретрансляторів. У всіх цих випадках стають кращими системи асинхронного багатоканального зв'язку, коли сигнали всіх абонентів передаються одночасно в загальній смузі частот, а канали не синхронізовані між собою.

Оскільки в таких системах за каналами не закріплені ні частотні, ні часові інтервали, а час роботи кожного каналу довільний, то такі системи називають **системами із вільним доступом** до лінії зв'язку або системами із незакріпленими каналами. У системах з вільним доступом кожному каналу (абонентові) привласнюється певна форма сигналу, яка і є відмітною ознакою («адресою»-кодом) даного абонента. На відміну від звичайного поділу за формою, де умова ортогональності сигналів виконується лише тоді, коли тактові інтервали всіх каналів жорстко синхронізовані, для можливості повного лінійного поділу сигналів у системі із вільним доступом ортогональність (або, принаймні, лінійна незалежність) між будь-якою парою сигналів повинна зберігатися при будь-якому зсуві за часом.

Можна показати, що ця умова виконується тільки в тому випадку, коли сигнали $s_k(t)$ являють собою білий шум, тобто мають необмежену ширину спектра і нескінченну дисперсію. Для реальних же сигналів така умова нездійсненна. Можна, однак, побудувати сигнали, для яких ця умова виконується приблизно в тому сенсі, що тобто скалярні добутки сигналів при будь-якому зсуві за часом будуть значно меншими від енергії елемента сигналу. Такі системи сигналів можна назвати «майже ортогональними». Такі сигнали повинні мати дуже широкий спектр, або, точніше, велику базу (добуток FT).

Найпоширенішим прикладом технічного використання майже ортогональних сигналів може служити мобільний стільниковий зв'язок, де використовуються певним чином підібрані послідовності дискретних, зокрема двійкових, імпульсів. Кожному каналу привласнюється одна із множини майже ортогональних і, отже, розділюваних дискретних послідовностей, яка є «адресою» (кодом) каналу. Використання майже ортогональних сигналів є подальшим розвитком способів поділу сигналів за формою.

Важливою перевагою таких систем є відсутність необхідності у центральній комутаційній станції; всі абоненти мають прямий доступ один до одного без частотної перебудови приймального і передавального обладнань. Подібно системам автоматичного телефонного зв'язку (АТС) тут досить набрати адресу викликуваного абонента, тобто змінити «форму» (код)

послідовності імпульсів, або, що те ж саме, змінити закон чергування імпульсів у послідовності.

У системах із закріпленими каналами додавання хоча б одного нового абонента виявляється можливим лише при виключенні одного із наявних у системі. Значно простіше ця задача вирішується в кодових системах. Тут внаслідок вільного доступу до лінії зв'язку можуть вести передачу будь-які K_0 абонентів із загального числа K абонентів системи зв'язку. При визначенні числа K_0 потрібно враховувати, що внаслідок неповної ортогональності сигналів у системі неминучі перехідні завади («шуми неортогональності»), рівень яких швидко росте зі збільшенням K_0 . Тому число одночасно працюючих абонентів повинно бути обмежене. Залежно від ступеня «активності» абонентів (тобто від частки часу, займаного k -м каналом для передачі повідомлень) можна організувати, наприклад, 1000-канальну систему зв'язку, в якій одночасно ведуть передачу будь-які 50 абонентів із 1000. У таких системах легко реалізуються резерви пропускну здатності, що виникають за рахунок «мало активних» абонентів.

Вивчивши статистику повідомлень, переданих по кожному каналу, можна встановити припустиме число каналів у системі K , при якому забезпечується нормальна робота K_0 «активних» каналів.

Тема 14. Ефективність систем електрозв'язку

Критерії ефективності. Інформаційна, енергетична, частотна ефективність. Границя Шеннона. Ефективність аналогових та цифрових систем. Способи підвищення ефективності. Енергетичний виграш завадостійкого кодування.

14.1 Оцінка ефективності та оптимізація параметрів телекомунікаційних систем (ТКС)

ТКС представляє складний комплекс, що характеризується ієрархічністю структури, наявністю прямих, зворотних і перехресних зв'язків між елементами. У попередніх дисциплінах і темах даного предмета розглядалися окремі пристрої (елементи) ТКС та процеси, що відбуваються в них. Тепер необхідно розглянути роботу ТКС у цілому, визначити алгоритм її функціонування з урахуванням взаємодії та властивостей елементів, що входять до неї, дати оцінку ефективності роботи системи в цілому та визначити шляхи її оптимізації. Для вирішення таких завдань застосовується системний підхід (системний аналіз).

Системний підхід базується на поданні об'єкта як складної системи з урахуванням її специфічних зв'язків та властивостей. Система визначається як цілісне утворення, що складається з пов'язаних між собою елементів, що володіє власними властивостями, що не впливають безпосередньо з властивостей елементів, що входять до неї.

Властивості системи передусім визначаються її цільовим призначенням, яке сприймається як сукупність завдань, розв'язуваних цією системою.

Для дослідження ТКС створюється її модель, у якій відображені найістотніші властивості та ознаки. Математична модель є сукупністю математичних співвідношень, що відображають структуру системи, алгоритми її функціонування, статистичні характеристики каналу, сигналу та перешкод, технічні та економічні показники системи.

Критерії ефективності.

Вирішення питань вибору найдоцільніших варіантів ТКС у кінцевому підсумку зводиться до вирішення завдань оптимізації цих систем за обраними критеріями якості.

Узагальненою характеристикою ефективності систем зв'язку є коефіцієнт використання каналу (**інформаційна ефективність**), який характеризує реальну швидкість передачі інформації R по відношенню до пропускнув здатності C каналу зв'язку:

$$\eta = R/C. \quad (14.1)$$

Інформаційна ефективність η завжди менше одиниці; що ближче η до одиниці, то досконаліша система.

Для оцінки ефективності систем зв'язку вводяться також коефіцієнт використання каналу потужності (**енергетична ефективність**):

$$\beta = R/(P_c/N_0), \quad (14.2)$$

та коефіцієнт використання каналу по смузі частот (**частотна ефективність**):

$$\gamma = R/\Delta F. \quad (14.3)$$

У цих формулах P_c – потужність сигналу; N_0 – спектральна густина шуму; ΔF – ширина смуги частот, яку займає сигнал.

Граничні можливості системи передачі інформації можна оцінити за допомогою виразу для пропускної здатності гаусівського неперервного каналу зв'язку зі смугою частот ΔF :

$$C = F \log_2(1 + P_c/P_w). \quad (14.4)$$

Тут $P_c = E_b B$ – середня потужність сигналу: E_b – енергія, що витрачається на передачу одного біта інформації; $B = 1/T_b$ – швидкість передачі джерела; T_b – час передачі джерелом одного біта інформації; $P_w = N_0/\Delta F$ – середня потужність шуму в смузі частот.

У реальних ТКС швидкість передачі інформації B [Біт/с], менша за пропускну здатність неперервного каналу: $B \leq C$. Можна показати, що після елементарних перетворень ця нерівність наводиться до виду:

$$\beta \leq \gamma/(2^\gamma - 1), \quad (14.5)$$

де
$$\beta = 1/h_2^2 = N_0/E_b \quad (14.6)$$

Тоді інформаційну ефективність для гаусівського неперервного каналу можна знайти за формулою:

$$\eta = \gamma/\log_2(\gamma/(\beta + 1)). \quad (14.7)$$

Відповідно до теореми Шеннона, при відповідних способах передачі та прийому величина η може бути як завгодно близькою до одиниці. При $\eta = 1$ отримуємо граничну залежність між β та γ :

$$\beta = \gamma/(2^\gamma - 1). \quad (14.8)$$

Наочно дана залежність представляється у вигляді кривої $\beta(\gamma)$ на площині (рис. 14.1).

Ця залежність часто називається границею (межею) Шеннона: вона відображає найкращий обмін між β і γ в неперервному каналі.

Аналіз співвідношення (14.6) та межі Шеннона показує, що підвищення частотної ефективності (тобто зниження витрат смуги $1/\gamma$) потребує збільшення енергетичних витрат (зниження енергетичної ефективності). Для неперервного каналу частотна ефективність змінюється в межах від 0 до ∞ , тоді як енергетична ефективність обмежена зверху:

$$\beta_{\max} = \lim_{\gamma \rightarrow 0} \beta = \lim_{\gamma \rightarrow 0} \frac{\gamma}{2^{\gamma} - 1} = \frac{1}{\ln 2} \approx 1,443.$$

Аналогічні граничні залежності $\beta = f(\gamma)$ можна отримати і для інших моделей каналу, якщо в (14.2) і (14.3) замість швидкості R підставити вираз для пропускної здатності відповідного каналу. Граничні залежності β -номограми дозволяють визначити системи, що задовольняють заданим вимогам щодо енергетичної та частотної ефективності, та встановити, наскільки ці показники близькі до граничних.

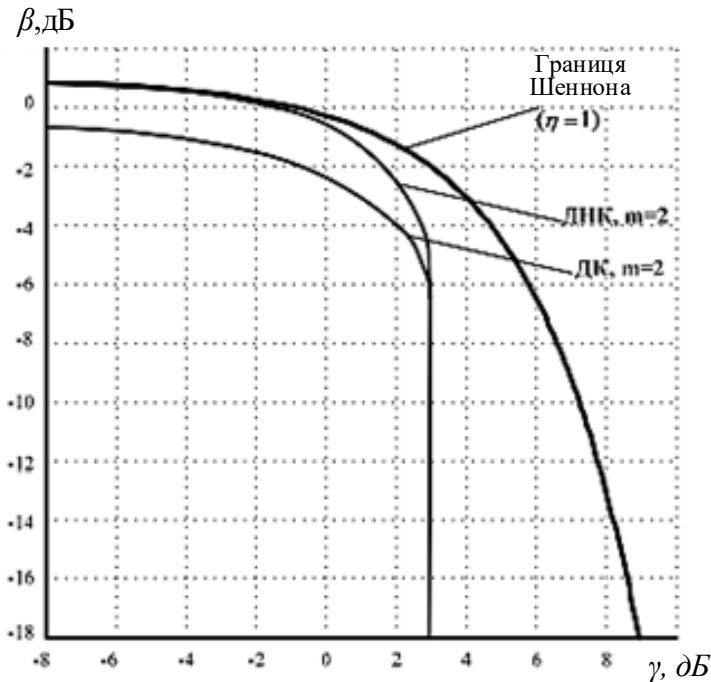


Рисунок 14.1 – Зв'язок частотної (γ) та енергетичної (β) ефективності

14.2 Ефективність аналогових та цифрових систем

У системах передачі дискретних повідомлень сигнал формується за допомогою кодування та модуляції. При цьому кодування здійснюється зазвичай у два етапи: кодування джерела з метою скорочення його надмірності та кодування каналу з метою зменшення ймовірності помилки за рахунок введення надмірності коду. При цьому вираз (14.1) для інформаційної ефективності системи передачі дискретних повідомлень можна подати у вигляді добутку:

$$\eta = R/C = \eta_{\text{ки}} \cdot \eta_{\text{кк}} \cdot \eta_{\text{м}}, \quad (14.9)$$

де $\eta_{\text{ки}}$ – ефективність кодера джерела; $\eta_{\text{кк}}$ – ефективність кодера каналу; $\eta_{\text{м}}$ – ефективність модему, що залежить від виду модуляції та способу обробки сигналу в каналі.

Середня швидкість передачі у системі під час використання багатопозиційних сигналів тривалістю T дорівнює $R = R_{\text{кк}} \cdot (\log_2 m)/T$ (біт/с), де $R_{\text{кк}} = k/n$ – швидкість перешкодостійкого коду. Тоді енергетична ефективність:

$$\beta = \frac{R}{P_c/N_0} = \frac{R_{\text{кк}} \log_2 m}{E_0/N_0} = \frac{1}{E_b/N_0}, \quad (14.10)$$

частотна ефективність може бути знайдена за формулою:

$$\gamma = R/\Delta F = \log_2 m / T \Delta F, \quad (14.11)$$

де $E_0 = P_c T = E_b R_{\text{кк}} \log_2 m$ – енергія сигналу; $E_b = E_0 / R_{\text{кк}} \log_2 m$ – енергія, що витрачається на передачу одного біта інформації.

Для визначення β та γ можуть використовуватися наближені формули:

$$\gamma \approx (\log_2 m)/n; \quad \beta \approx 1/(E_b / N_0), \quad (14.12)$$

де n – розмірність сигналу, в m -позиційній системі.

У таблиці 14.1 наведено значення m та формули для наближених розрахунків γ деяких ансамблів сигналів.

Формули для наближених розрахунків частотної ефективності деяких ансамблів сигналів

Ансамбль сигналів	Ортогональний	Біортогональний	Симплексний
m	n	$2n$	$n + 1$
γ	$(\log_2 m)/m$	$2(\log_2 m)/m$	$(\log_2 m)/(m - 1)$

У реальних системах ймовірність помилки завжди має ненульове значення і $\eta < 1$. У цих випадках при заданому значенні $p_{\text{пом}} = \text{const}$ можна визначити окремо β та γ та побудувати криві $\beta = f(\gamma)$. (рис. 14.2).

$\beta, \text{дБ}$

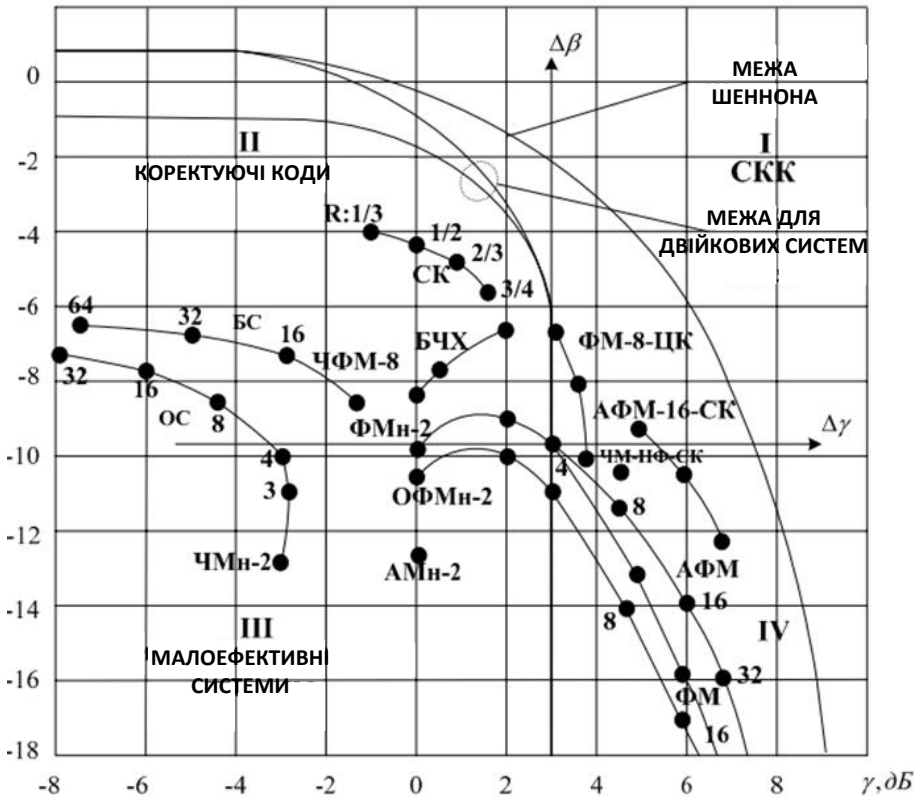


Рисунок 14.2 – Криві енергетичної та частотної ефективності цифрових систем зв'язку

У координатах β і γ кожному варіанту реальної системи відповідатиме точка на площині. Всі ці точки розташовуються нижче за граничну криву

Шеннона і нижче за граничну криву відповідного каналу. Хід цих кривих залежить від виду модуляції, методу кодування та способу обробки сигналів. Біля графіків рисунку 14.2. вказано число позицій дискретного сигналу m . Криві розраховані на підставі формул оцінки завадостійкості різних методів модуляції для оптимального прийому сигналів за ймовірності помилки на біт $p_{\text{пом}} = 10^{-5}$.

Смуга частот, що займається для ЧМн $\Delta F = m/(T \cdot \log_2 m)$, а для ФМн (АМн) $\Delta F = 1/(T \cdot \log_2 m)$.

Аналіз рис. 14.2. показує, що у системах з ЧМн зі збільшенням числа позицій m енергетична ефективність β збільшується, а частотна ефективність γ зменшується. У системах з ФМн та ОФМн, навпаки, зі збільшенням m коефіцієнт β зменшується, а γ – збільшується. Таким чином, умови обміну на γ за рахунок зміни числа позицій сигналів в системах зв'язку з ЧМн і ФМн різні.

Подані на рис. 14.2. результати дозволяють визначити системи, що задовольняють заданим вимогам щодо енергетичної та частотної ефективності, та встановити, наскільки ці показники близькі до граничних.

Після вибору системи за показниками β та γ інформаційна ефективність обчислюється з використанням формули (14.7).

Наприклад, для сигналів АМн-2 показник інформаційної ефективності $\eta \approx 0,228$, а для ЧМн-2 $\eta \approx 0,145$; для ФМн-2 $\eta \approx 0,25$, а для ФМн-4 $\eta \approx 0,47$.

Аналіз граничних кривих показує, що ефективність дискретних систем передачі можна значно підвищити, якщо замість двійкових застосовувати багатопозиційні сигнали ($m > 2$).

Отже, **ефективність передачі дискретних повідомлень** можна підвищити шляхом використання:

- рознесеного прийому сигналів – передане повідомлення у приймачі відтворюється не за одним прийнятим сигналом, а за двома чи декількома, які несуть одне й те саме повідомлення; рознесення може бути частотним, часовим або просторовим;

- прийому в цілому – замість двох вирішуючих схем для кодеків джерела і каналу зв'язку в приймачі використовують одну, а рішення приймають по сигналу, що відповідає всій кодовій комбінації, а не її частині. При цьому в роботі приймача використовується та частина корисної інформації з виходу першої вирішуючої схеми, яка у випадку поелементного прийому губиться після того, як схема прийняла рішення. Поелементний прийом і прийом в цілому є еквівалентними лише у випадку кодування без надмірності;

- каналів зворотного зв'язку – пропускну здатність каналів обох напрямків частково використовують для передачі додаткової інформації, яка сприяє підвищенню ефективності зв'язку;

– шумоподібних несучих – дозволяє не лише наблизити швидкість передачі інформації до теоретичної межі (пропускної здатності каналів), але й істотно підвищити вірність передаваних повідомлень;

– адаптивної корекції характеристик каналу зв'язку – періодичне зондування каналу випробувальним імпульсом і корегування за необхідності коефіцієнту передачі КЗ.

Ефективність передачі неперервних повідомлень значною мірою залежить від виду модуляції. Для порівняльного аналізу різних видів модуляції зазвичай використовують вигравш по відношенню сигнал/шум ($h_{\text{вих}}$) та коефіцієнт використання пропускної спроможності каналів зв'язку (ν):

$$\eta = \frac{\log h_{\text{вих}}}{\nu \cdot \log(h_{\text{вих}}/g + 1)}. \quad (14.13)$$

У таблиці 14.2 наведено дані порівняльного аналізу ефективності різних видів модуляції, отримані при $h_{\text{вих}} = 40$ дБ та пік-факторі $\Pi = 3$ для гаусівського каналу при оптимальній обробці сигналів.

Таблиця 14.2

Значення вигравшу та інформаційної ефективності деяких систем передачі неперервних повідомлень

Вид модуляції	$\nu = F/F_c$	$g = h_{\text{вих}} / h_{\text{вх}}$	$g' = g/\nu$	$\eta = R/C$
АМ	2	0,2	0,1	0,42
БМ	2	2	1	0,50
ОМ	1	1	1	1
ФМ	20	222	11,1	0,12
ЧМ	20	666	33,3	0,17
ФІМ-АМ	20	666	33,3	0,17
ІКМ-АМ	20	250	12,5	0,23
ІКМ-ЧМ	20	500	25	0,32
ІКМ-ФМ	20	1000	50	0,48
ІС	20	6310	315	1

Аналіз показує, що найбільша інформаційна ефективність досягається при односмуговій модуляції, проте значення узагальненого вигравшу для цього виду модуляції ($g'=1$) свідчить про те, що в системі відсутній вигравш за стійкістю до перешкод.

Одноканальні системи ЧС та ФІМ приблизно рівноцінні. У цих системах, а також у цифрових системах з ІКМ, висока стійкість до перешкод може бути досягнута за допомогою збільшення ширини спектра сигналу, тобто за рахунок частотної надмірності. При великих індексах ФМ та ЧС

наближаються за перешкодою до ідеальної системи (виграш становить десятки і сотні разів), але інформаційна ефективність таких систем мала (0,12 - 0,17) через велику частотну надмірність. Основними способами підвищення ефективності передачі неперервних повідомлень є усунення надмірності, статистичне ущільнення та застосування цифрових видів модуляції.

Аналогові системи ОМ, АМ та вузькосмугові ЧС забезпечують високу частотну ефективність за порівняно низької енергетичної ефективності. Застосування цих систем доцільно в каналах з гарною енергетикою (при великих значеннях $h_{вх}$) або у випадках, коли необхідне значення $h_{вих}$ мале. Цифрові системи забезпечують високу β – ефективність за досить хорошої γ – ефективності. У каналах з обмеженою енергетикою (за малих значень $h_{вх}$) переваги цифрових систем особливо помітні. При високій якості передачі, коли необхідні значення $h_{вих}$ великі, широкосмугова ЧС і цифрові системи забезпечують приблизно однакову ефективність.

У багатоканальних системах ефективність зв'язку знижується з допомогою недосконалості системи поділу сигналів. Показники частотної, енергетичної та інформаційної ефективності для систем з множинним доступом визначаються на підставі сумарної швидкості передачі, яка залежить від методів формування та обробки інформаційних сигналів у парціальних каналах та методів доступу.

Основними способами **підвищення ефективності** передачі неперервних повідомлень є:

- усунення надмірності неперервних повідомлень – шляхом дискретної передачі не власне сигналу, а його найбільш характерних параметрів, зміна яких у часі відбувається набагато повільніше, порівняно із зміною самого сигналу;
- статистичне ущільнення, суть якого полягає в тому, що паузи в передачі і вільні смуги частот, зумовлені статистичними особливостями джерел інформації, використовують для організації додаткових каналів зв'язку і передачі додаткової інформації;
- керування динамічним навантаженням каналів – контроль і вирівнювання реального навантаження різноманітних каналів, в результаті чого коефіцієнти використання пропускних здатностей всіх каналів зв'язку стають приблизно однаковими;
- використання цифрових методів модуляції тощо.

14.3 Вибір сигналів та перешкодостійких кодів

Ефективність систем передачі дискретних повідомлень можна суттєво підвищити шляхом застосування багатопозиційних сигналів та коригувальних кодів.

На рис. 14.2. наведено $\beta\gamma$ -діаграми для деяких ансамблів багатопозиційних сигналів. Центральне місце рис. 14.2. займають криві для

систем із сигналами ФМн-4, які належать до класу багатопозиційних при $m = 4$. У цифрових мережах система ФМн-4 є найбільш поширеною і прийнята як стандарт, тому при порівняльній оцінці ефективності систем вона прийнята за еталон. Якщо початок координат перенести в точку, що відповідає ФМн-4, то в новій системі координат по вертикальній осі буде відраховуватись енергетичний вигравш $\Delta\beta$ аналізованих систем порівняно з ФМн-4, а по горизонтальній осі – вигравш $\Delta\gamma$ по питомій швидкості. У цій системі координат усі можливі системи зв'язку можна умовно розділити на чотири групи, що відповідають чотирьом квадрантам на площині.

Малоефективні системи (III квадрант), що мають щодо ФМн-4 програш по β і γ , наприклад, АМн-2, ЧМн-2. Системи з високою енергетичною ефективністю (II квадрант), що забезпечують вигравш по β і програш по γ (системи з коригуючими кодами). Системи з високою частотною ефективністю (IV квадрант), що забезпечують вигравш γ і програш β (системи з багатопозиційними ФМн та АФМ сигналами). Високоефективні системи (I квадрант), що дозволяють отримати одночасно вигравш за обома показниками β та γ на основі застосування складних сигнально-кодових конструкцій.

Поряд з багатопозиційними сигналами для підвищення ефективності широко використовуються задові коди. Застосування коригувальних кодів дозволяє підвищити вірність передачі повідомлень або за заданої вірності підвищити енергетичну ефективність системи. Це особливо важливо для систем із малою енергетикою, наприклад, систем супутникового зв'язку.

На практиці використовуються як блокові, і безперервні коди. На рис 14.2. наведені криві ефективності для циклічного коду Боуза-Чоудхурі-Хоквінгема (БЧХ) та для коду згортки (СК) з декодуванням за алгоритмом Вітербі.

Застосування циклічного коду дозволяє отримати енергетичний вигравш $\Delta\beta = 2 \dots 4$ дБ, а згорткового коду $\Delta\beta = 4 \dots 6$ дБ в обмін на зниження частотної ефективності приблизно в 2 рази (3 дБ).

Енергетичний вигравш $\Delta\beta$ від застосування задовітійкого кодування тим більше, чим вище потрібна вірність передачі. Для безперервного каналу з білим гауссівським шумом при ймовірності помилки 10^{-5} граничний енергетичний вигравш кодування порівняно з ФМн-2 без кодування при оптимальному когерентному прийомі становить приблизно 10 дБ.

Застосування циклічного коду в каналі з ФМн або згорткового коду в каналі з АФМ дозволяє підвищити як енергетичну, так і частотну ефективності.

Побудова високоефективних систем з урахуванням сигнально-кодових конструкцій веде до неминучого збільшення складності системи. Не пропускна здатність, а складність є обмежуючим фактором при побудові високоефективних систем. Завдання полягає в тому, щоб побудувати систему, яка задовольняє високі показники ефективності, при допустимій складності.

При сучасній елементній базі витрати на реалізацію пристроїв, що кодують і декодують, значно скоротилися. Водночас, вартість енергетики каналу практично не змінилася. Таким чином, «ціна» виграшу $\Delta\beta$ за рахунок кодування може бути істотно меншою за ціну того ж виграшу, отриманого за рахунок збільшення енергетики каналу (потужності сигналу або розмірів антен).

Зазначимо, що вибір способів кодування та модуляції залежить від характеристик каналу. Поліпшення цих характеристик, наприклад, шляхом адаптації до перешкод та оцінювання спотворень сигналу та їх подальшої компенсації, знижує втрати в каналі та створює кращі умови для застосування коригувальних кодів.

Порівняння між собою різних ТКС здійснюється за ступенем використання ними основних ресурсів каналу зв'язку (пропускної здатності, потужності, займаної смуги частот), що виражається через показники інформаційної, енергетичної та частотної ефективності. Створення ТКС, у яких досягаються близькі до граничних показники ефективності, потребує узгодження методів модуляції та кодування з урахуванням статистичних властивостей неперервного каналу.

ІНФОРМАЦІЙНІ ДЖЕРЕЛА

1. Вакуленко О. В., Голь В. Д., Ірха М. С., Хахлюк О. А. Лінії передачі: підручник. Київ : ІСЗЗІ КПІ ім. Ігоря Сікорського, 2021. 218 с.
2. Головін Ю. О. Основи радіозв'язку: підручник. Київ: КПІ ім. Ігоря Сікорського, Політехніка, 2021. 234 с.
3. Головін Ю. О., Могилевич Д. І. Основи теорії радіозв'язку. Теоретичні основи та практичні аспекти. К.: КПІ ім. Ігоря Сікорського, 2023. 248 с.
4. Голь В. Д., Ірха М. С. Системи передачі даних: конспект лекцій. Київ: ІСЗЗІ КПІ ім. Ігоря Сікорського, 2021. 126с. URL: <https://ela.kpi.ua/handle/123456789/45443> (дата звернення 26.12.2023 р).
5. Гусев О. Ю., Конахович Г. Ф., Корнієнко В. І. Теорія електричного зв'язку. Львів: «Магнолія», 2021. 364 с.
6. Івашко А. В., Крилова В. А. Теорія інформації та кодування в прикладах і задачах: навч.-метод. посіб. Харків: НТУ «ХПІ», 2022. 317 с.
7. Пархомей І., Цюпа Н. Основи теорії інформаційних процесів. Частина 2. Системи обробки сигналів. навч. посіб. Київ: КПІ ім. Ігоря Сікорського, 2020. URL: <https://ela.kpi.ua/handle/123456789/39133> (дата звернення 26.012.2023 р)
8. Alencar M. S., da Rocha V. C. Jr. Communication Systems. Springer Cham, 2020. 401 p.
9. Bhooshan S. Fundamentals of Analogue and Digital Communication Systems. Springer Singapore. 2022. 645 p.
10. Harris J. F. Multirate Signal Processing For Communication Systems. Prentice Hall, 2021. 600p.
11. Plonus M. Electronics and Communications for Scientists and Engineers. Elsevier Inc. 2020. 498p.

Теорія електричного зв'язку. Конспект лекцій для здобувачів першого (бакалаврського) рівня вищої освіти освітньої програми «Комп'ютеризовані телекомунікаційні мережі» галузі знань 17 Електроніка, автоматизація та електронні комунікації, спеціальності 172 Електронні комунікації та радіотехніка, всіх форм навчання / уклад. Н. М. Якимчук. Луцьк: ЛНТУ, 2025. 190 с.

Комп'ютерний набір
Редактор

Наталія ЯКИМЧУК
Наталія ЯКИМЧУК

Підп. до друку «__»_____2025 р.
Формат 60x84/16. Папір офс.
Гарн. Таймс. Ум. друк. арк. 6.
Тираж 50 прим.

Відділ іміджу та промоції
Луцького національного технічного університету
43018 м. Луцьк, вул. Львівська, 75
Друк – ВІП ЛНТУ